International Dairy Journal 149 (2024) 105817

Contents lists available at ScienceDirect

International Dairy Journal

journal homepage: www.elsevier.com/locate/idairyj

# Whole-genome sequencing reveals virulence and antibiotic resistance determinants in *Enterococcus faecium* strains isolated from the dairy industry in Mexico

Jean Pierre González-Gómez [a], María Guadalupe Avila-Novoa [b], Berenice González-Torres [a], Pedro Javier Guerrero-Medina [b], Bruno Gomez-Gil [c], Cristobal Chaidez [a], Melesio Gutiérrez-Lomelí [b, *]

[a] Laboratorio Nacional para la Investigación en Inocuidad Alimentaria (LANIIA), Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD), Carretera a El dorado Km 5.5, Campo El Diez, 80110, Culiacan, Sinaloa, Mexico
[b] Centro de Investigación en Biotecnología Microbiana y Alimentaria, Departamento de Ciencias Básicas, Centro Universitario de la Ciénega, Universidad de Guadalajara, Av. Universidad 1115, Ocotlan, Jalisco, Mexico
[c] Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD), Unidad Mazatlán en Acuicultura y Manejo Ambiental, AP 711, Mazatlán, Sinaloa, Mexico

ARTICLE INFO

*Article history:*
Received 15 June 2023
Received in revised form
25 September 2023
Accepted 25 September 2023
Available online 29 October 2023

ABSTRACT

*Enterococcus faecium* is a part of the native microbiota in fermented and dairy products. However, the acquisition of virulence determinants turns this bacterium into a pathogen. To evaluate the genetic relatedness and virulence profile, we performed whole-genome sequencing of four *E. faecium* strains isolated from the dairy industry and compared them with clinical and environmental isolates reported in Mexico. Our findings revealed close genetic relationships between certain dairy-associated and clinical isolates, with shared antibiotic-resistance determinants and prophages. Notably, the *tetL*, *tetM*, *lnuA*, and *lsaE* genes, associated with tetracycline and lincosamide resistance, were present in the two dairy-related strains. Clinical isolates exhibited a higher prevalence of adhesion and biofilm-related genes, including *acm*, *ecbA*, *fss3*, and *sgrA*. Additionally, the identification of prophages in dairy-related strains suggests their potential for genetic exchange. Thus, it is crucial to assess the risk posed by *E. faecium* as a potential pathogen in the dairy plant environment.



## 1. Introduction

The *Enterococcus* genus is a widespread group of Gram-positive, facultatively anaerobic, and catalase-negative cocci that are commonly found in large numbers in the human and animal gastrointestinal tract, as well as in animal-derived foods such as dairy products (Ben Braiek & Smaoui, 2019; Giraffa, 2003). Within the dairy industry, nonstarter lactic acid bacteria, including *Enterococcus faecium*, *Enterococcus faecalis*, and *Enterococcus durans* are frequently present in artisanal cheeses made with either raw or pasteurised milk (Citak, Yucel, & Orhan, 2004; Gelsomino, Vancanneyt, Condon, Swings, & Cogan, 2001; Nieto-Arribas et al., 2011) and exhibit high levels of thermal resistance during high-temperature short-term pasteurisation (Chajecka-Wierzchowska, Zadernowska, & Garcia-Solache, 2020). Additionally, species within this genus are utilised as starter cultures or cocultures in Mediterranean cheeses due to their advantageous biotechnological properties (Foulquie Moreno, Sarantinopoulos, Tsakalidou, & De Vuyst, 2006). Their role in cheese ripening and flavor development is mediated by citrate metabolism and proteolytic and esterolytic activity; furthermore, enterococci are also capable of producing bacteriocins and acting as probiotics (Sarantinopoulos et al., 2001; van Kranenburg et al., 2002). However, despite the substantial importance of the Mexican dairy industry and its production of unique and traditional cheeses, there is a lack of information on the *Enterococcus* species present in this context.

On the other hand, the acquisition of virulence factors has positioned *E. faecium* as a member of the ESKAPE group. This acronym represents the species *E. faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp., all of which are notorious for their

* Corresponding author.
*E-mail address:* melesio.gutierrez@academicos.udg.mx (M. Gutiérrez-Lomelí).

https://doi.org/10.1016/j.idairyj.2023.105817
0958-6946/

antibiotic resistance and ability to cause infections in immunocompromised individuals in health care settings and residential environments (Rice, 2008). As a pathogen, *E. faecium* can cause a variety of complex infections, such as abdominal, skin, urinary tract, endocarditis, and bacteremia (Srinivasan & Evans, 2018), which can be particularly challenging to treat due to its high levels of antibiotic resistance (Agudelo Higuita & Huycke, 2014). This pathogen is intrinsically resistant to various antimicrobial agents and can also acquire resistance via horizontal gene transfer and single nucleotide mutations (Arias & Murray, 2012; Hollenbeck & Rice, 2012). Among infectious isolates, vancomycin-resistant *E. faecium* (VREF) is a high-priority pathogen on the World Health Organisation (WHO) global list of antibiotic-resistant bacteria because it is associated with higher mortality rates, longer hospital stays, and higher health care costs (Tacconelli et al., 2018). Additionally, it can harbour genes related to virulence factors, such as adhesins and invasion, or prophages that provide fitness, virulence, and/or resistance to secondary infections (Matos et al., 2013; Ogier & Serror, 2008). Compared with clinical strains, the incidence of virulence factors in enterococci isolated from food and other natural sources has been shown to be low (Franz et al., 2001). However, it is essential to monitor the genetic content and evaluate the relatedness of strains found in the dairy industry with clinical isolates to avoid potential risks to dairy product consumers.

In this study, we performed a comprehensive genome analysis of four *E. faecium* strains obtained from food-contact surfaces from the dairy industry, coupled with five previously described clinical and environmental isolates, using whole-genome sequencing and comparative genome analysis. The assessment of the genetic distances between food-related strains and clinical isolates in Mexico and their genetic content related to virulence factors provides important insights into the potential risk of strains present in dairy products with undesirable and dangerous characteristics for the consumer.

## 2. Materials and methods

### 2.1. DNA extraction and whole-genome sequencing of E. faecium strains

Four strains of *E. faecium* isolated in a previous study from a dairy industry were provided by the Centre for Research in Microbial and Food Biotechnology of the University of Guadalajara. The strains were cultured in TSB under aerobic conditions at 37 °C for 24 h. DNA was extracted using the DNeasy Blood & Tissue kit (QIAGEN, Mexico City, Mexico) according to the manufacturer's instructions, and quantification was performed using a NanoDrop 2000c Spectrophotometer. Following the extraction process, the libraries were created using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) following the manufacturer's guidelines. The libraries were quantified using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and whole-genome sequencing was conducted with the Illumina MiniSeq platform (2 × 150 bp paired-end protocol, 300 cycles).

### 2.2. De novo genome assembly and recovery of reported genomes

The quality of the reads was assessed through FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and then trimmed by fastp v0.22.0 (Chen, Zhou, Chen, & Gu, 2018). After cleaning, SPAdes v3.15.3 (Bankevich et al., 2012) was utilised for assembling the reads into contigs, with default settings used for both tools. The characteristics of the assemblies, such as lengths, number of contigs, GC percentage, N50, and L50, were obtained by QUAST v5.2.0 (Gurevich, Saveliev, Vyahhi, & Tesler, 2013).

Additionally, we performed a search of the *E. faecium* genomes previously reported from Mexico in the GenBank repository of the National Center for Biotechnology Information (NCBI). The assemblies and SRA projects were downloaded and the fastq sequences were cleaned and assembled as previously mentioned. Assemblies that showed 37.5–38% GC and genome lengths of 2.5 kbp–3.2 kbp, typical for *E. faecium* strains, were selected for further analyses (Table 1).

### 2.3. Pangenome analysis

The assemblies of nine *E. faecium* isolates, including three clinical isolates (ERV275, ERV279, and MGRG), were used to conduct a pangenomic analysis with anvi'o v7.1 (Eren et al., 2015, 2021). The analysis was carried out following the comprehensive pangenomic workflow available at (https://merenlab.org/2016/11/08/pangenomics-v2/). First, the anvi-gen-contigs-database script was utilised to construct a database, and Prodigal (Hyatt et al., 2010) was applied to identify open reading frames in the contigs. Second, the anvi-run-ncbi-cogs and anvi-run-kegg-kofams scripts were used for gene annotations from NCBI's Clusters of Orthologous Groups database (Tatusov, Galperin, Natale, & Koonin, 2000) and KOfam profiles (Aramaki et al., 2020). The genome database was built using anvi-gen-genomes-storage, and an anvi-pan-genome was employed for visualisation. This last program utilises BLASTP from NCBI to assess the similarity of each amino acid sequence in all genomes to each other and subsequently applies the MCL algorithm (van Dongen & Abreu-Goodger, 2012) to recognise clusters in the amino acid sequence similarity outcomes. Finally, the average nucleotide identity (ANI) across genomes was calculated with PyANI (Pritchard, Glover, Humphris, Elphinstone, & Toth, 2016) through anvi-compute-ani.

### 2.4. Phylogenomic analysis

To estimate the evolutionary distance between genomes, we selected the single copy core genes showing a maximum functional

**Table 1**
Characteristics of the *Enterococcus faecium* genomes included in this study.

| Strain | GB/SRA accession | Length (bp) | Contigs | N50 | L50 | %GC | Coverage | Isolation source |
|---|---|---|---|---|---|---|---|---|
| CIBMA12 | JASAUC000000000 | 2,935,203 | 757 | 30,728 | 26 | 37.92 | 157.8 | Dairy factory |
| CIBMA13 | JASAUB000000000 | 2,806,502 | 194 | 54,252 | 15 | 37.86 | 60.4 | Dairy factory |
| CIBMA14 | JASAUA000000000 | 2,993,710 | 799 | 57,185 | 15 | 37.88 | 80.6 | Dairy factory |
| CIBMA17 | JASATZ000000000 | 2,670,103 | 119 | 163,283 | 7 | 37.89 | 49.6 | Dairy factory |
| D | MEGX00000000.1 | 2,772,228 | 81 | 88,852 | 11 | 37.92 | 326.0 | Cotija cheese |
| EF-IBT-2022 | CP101669.1 | 2,686,950 | 1 | 2,686,950 | 1 | 38.00 | 300.0 | Shrimp gut |
| ERV275 | MJFS00000000.1 | 3,088,802 | 277 | 31,906 | 31 | 37.60 | 159.0 | *Homo sapiens* |
| ERV279 | MJFT00000000.1 | 3,136,967 | 415 | 18,556 | 415 | 37.70 | 89.0 | *Homo sapiens* |
| MGRG | SRS7879195 | 2,704,315 | 118 | 81,568 | 12 | 37.91 | 94.2 | Vaginal discharge |

homogeneity index of 0.95 and minimum geometric index of 0.95 occurring in all nine genomes in the visualisation tool implemented in the anvi-display-pan program in anvi'o v7.1. The aligned and concatenated gene clusters were exported through the anvi-get-sequences-for-gene-clusters program and this fasta file was used to generate a Newick file with PhyML v2.2.3 (Guindon et al., 2010) with 100 bootstrap replicates and a Le Gascuel substitution model. The resulting phylogenomic tree was edited using iTOL v6 (Letunic & Bork, 2021).

### 2.5. Prophages, antimicrobial resistance and virulence genes

The presence of prophage in the *E. faecium* genomes was assessed through PHASTEST v3.0 (Arndt, Marcu, Liang, & Wishart, 2019). To evaluate the presence of genes related to antimicrobial resistance and virulence, the program ABRicate v0.8.13 (https://github.com/tseemann/abricate) was employed to screen the genomes against the databases CARD (https://card.mcmaster.ca/home), Resfinder (Florensa, Kaas, Clausen, Aytan-Aktug, & Aarestrup, 2022), NCBI AMRFinderPlus (PRJNA313047), and VFDB (Liu, Zheng, Zhou, Chen, & Yang, 2022). A map of the presence or absence of the evaluated genes was obtained through the ggplot2 library in R v4.2.1 (R Core Team, 2019).

## 3. Results and discussion

### 3.1. Genome characterisation

In total, nine genomes of *E. faecium* isolated from diverse environments in Mexico were included in this study. The mean coverage of four novel genomes assembled de novo was 87×, and the number of contigs ranged between 119 and 757 (Table 1). The average nucleotide identity among all genomes was higher than 98% (Fig. 1); therefore, the novel genomes belong to *E. faecium* since those previously reported in GenBank were assigned to this taxonomic group since the criteria for bacterial species demarcation are 95–96% ANI values (Kim, Oh, Park, & Chun, 2014).

This *Enterococcus* species, together with *E. faecalis* and *E. durans*, is among the most common in food products (Ogier & Serror, 2008). *E. faecium* strain D, which was isolated in Mexico, has been reported

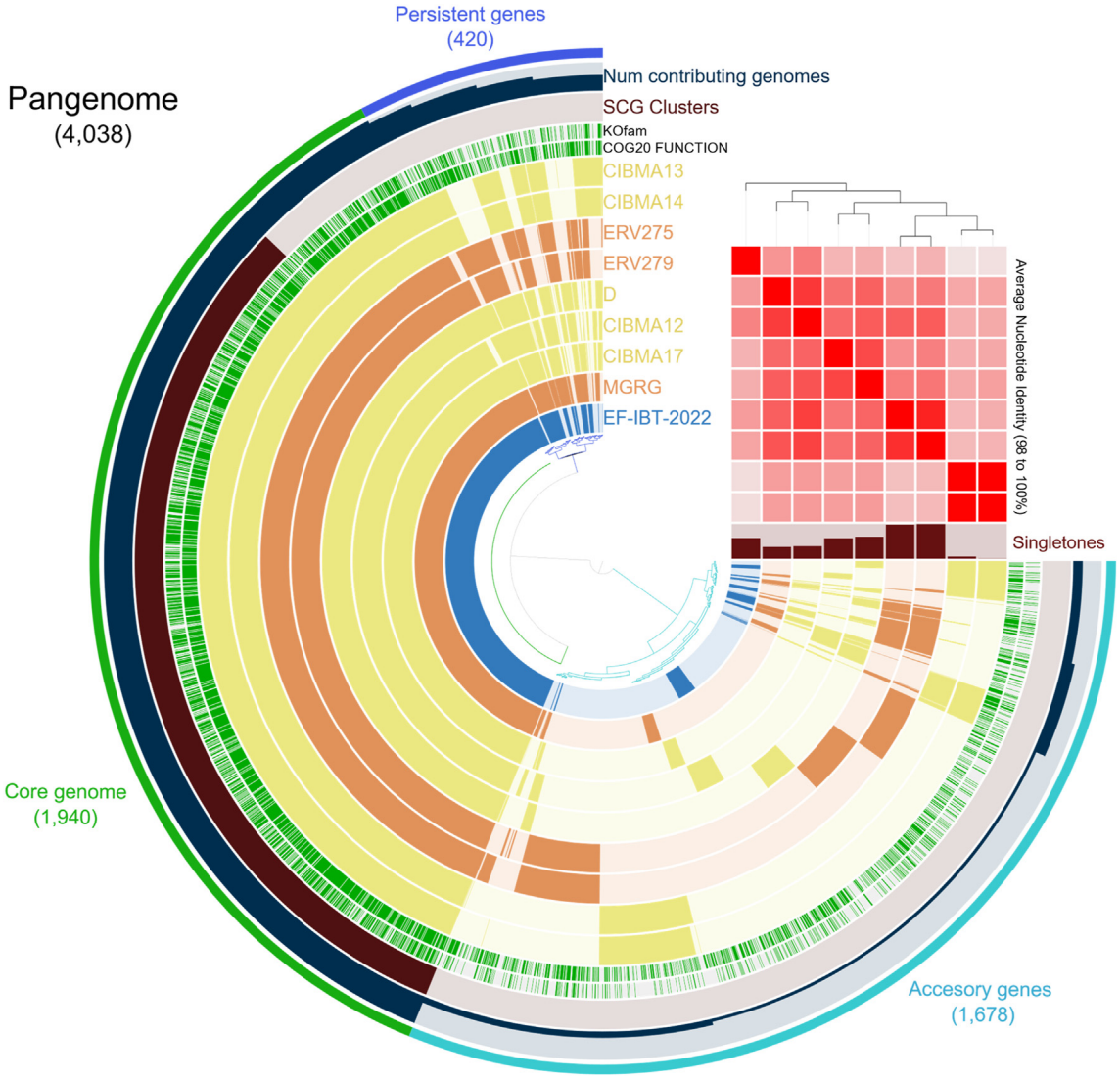


**Fig. 1.** Pangenome of nine *Enterococcus faecium* strains isolated from Mexico. The layers represent individual genomes organised by cluster frequency. Dark colours in the layers indicate the presence of a gene cluster, and light colours indicate its absence. ANI values across different genomes are represented by a heatmap with a gradient colour from 98% identity (white) to 100% identity (red).

to have potential for the development of aromas and flavors in the dairy industry, as well as genes for producing antimicrobial compounds and lipolytic activity that are likely associated with the cheese ripening process (Olvera-Garcia, Sanchez-Flores, & Quirasco Baruch, 2018). However, this strain also harbours genes associated with antimicrobial resistance and exhibits phenotypic resistance to gentamicin, kanamycin, streptomycin, and erythromycin, which suggests that careful evaluation of its biotechnological potential is necessary. Although this species is prevalent in the dairy industry, there is also significant intraspecies genetic variability (Kim & Marco, 2014), meaning that its presence in this type of environment cannot be ignored due to the high likelihood of the emergence of strains with highly virulent profiles, which could pose a risk to consumers.

### 3.2. Pangenome analysis

The pangenome analysis revealed a total of 4038 gene clusters (GCs) among the nine *E. faecium* genomes investigated (Fig. 1), which were classified into three main groups based on their prevalence: the core genome (100% occurrence) encompassing 1940 GCs, persistent genes (90–99%) including 420 GCs, and accessory genes (less than 90% occurrence) that comprised 1678 GCs. The core/pangenome ratio was calculated as 48.04%, indicating an open pangenome for organisms with an allopatric lifestyle, which typically have a core/pangenome ratio of less than 89% (Rouli, Merhej, Fournier, & Raoult, 2015). The pangenome size and the core/pangenome ratio are influenced by various factors, such as horizontal gene transfer, gene loss events, pseudogene formation, and the range of environments that an organism inhabits (Rouli et al., 2015). Therefore, new virulence genes, including those associated with antimicrobial resistance, are likely to be present in novel *E. faecium* strains isolated from the dairy industry.

Most of the functionally annotated genes by the KOfam and COG20 databases were identified in the core genome, which mainly consists of essential genes related to primary functions, such as DNA replication and repair, transcription and translation, transport and metabolism, and cell structure and division. Singleton genes are those that are detected in a single genome and may be acquired by genetic recombination or may be the result of a pseudogenisation process. Pseudogenes can serve as a historical record of phenotypic characteristics that have been lost during the course of prokaryotic evolution, as they are the remnants of genes that have undergone inactivating nucleotide substitutions or insertions/deletions in comparison with their original coding sequences (Goodhead & Darby, 2015; Ochman & Davalos, 2006). The clinical strains ERV279 and ERV275 displayed a larger number of pseudogenes compared with environmental and dairy-related strains, possibly attributed to the stronger selective pressures clinical strains undergo compared with environmental strains, such as exposure to antibiotics (Bengtsson-Palme, Kristiansson, & Larsson, 2018). This exposure heightens the probability of accumulating deleterious mutations in nonessential genes.

### 3.3. Phylogenomic tree

The phylogenomic analysis showed two main clades (Fig. 2). The first consists of two *E. faecium* strains isolated from the dairy industry, and the second consists of strains from different environments, even though they are closely related to each other. These results are consistent with those obtained in the ANI test and confirm the distinction between the two main groups.

de Been, van Schaik, Cheng, Corander, and Willems (2013) evaluated the phylogenomic relationship based on core genome single-nucleotide polymorphisms of 34 *E. faecium* genomes and reported two main clades grouped according to the isolation source, which has been previously identified and designated hospital-associated and commensal-associated clades (Galloway-Pena, Roh, Latorre, Qin, & Murray, 2012). These results differ from those observed in our study, which may be due to the increased dispersion of *E. faecium* strains between isolation environments and, as inferred from their open pangenome, their ability to acquire new virulence factors that allow them to colonise new environments.

One of the most remarkable aspects is the lack of studies on genomic surveillance of emerging pathogens in Mexico. Multiple population studies of *Enterococcus* spp. have reported a high number of genomes in a given geographical location (Holman et al., 2021; Lebreton et al., 2013), even in developing countries (Mbanga et al., 2021). With the genomes reported in this study, we almost doubled the *E. faecium* genomes available in GenBank isolated from Mexico; however, it is extremely important to increase the number of strains that are sequenced and reported, both from clinical cases and from environments related to the dairy industry, to make a more robust analysis of the distribution of populations and their virulence profiles.

### 3.4. Distribution of prophages, antimicrobial resistance, and virulence genes

To date, lysogenic bacteriophages have been insufficiently investigated in *E. faecium*; however, they play a crucial role in the genetic variability of this species. The presence of lysogenic bacteriophages may provide benefits to the host organism, as they are not expected to persist if they do not confer any advantages to the host organism in the face of competition with noninfected strains.

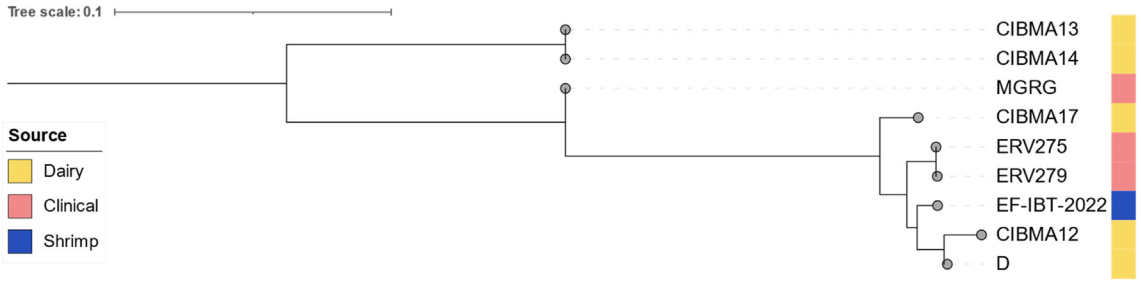


**Fig. 2.** Phylogenomic tree based on the best seven single-copy core genes of *Enterococcus faecium* isolated from Mexico.

Additionally, the presence of lysogenic bacteriophages can increase the environmental fitness of the lysogenic host, known as lysogenic conversion. This phenomenon has been previously described in several gram-positive pathogens, including *Streptococcus pyogenes* and *S. aureus*, where it has been shown to enhance virulence (Brüssow, 2014).

Intact prophages were found in 7/9 (77%) of the *E. faecium* genomes (Table 2). Interestingly, some strains that were closely related, such as CIBMA13 and CIBMA14, showed an identical prophage profile, and others, such as ERV275 and ERV279, showed different prophages, even though they were indistinguishable through the phylogenomic analysis. van Schaik et al. (2010), also identified a large variety of phages in this relatively small sample of strains, demonstrating that they play an important role in the intraspecies diversity of *E. faecium*; however, they raised the prospect of determining the contribution of bacteriophages in niche adaptation and an infectious phenotype in this species. A total of 10 intact prophages were identified across all the isolates, and their distribution was independent of their location in the phylogenomic tree. Entero_vB_IME197 was the most common prophage (n = 3) and occurred only in dairy-related strains. This phage belongs to the *Caudoviricetes* class and contains genes associated with both lysogenic and lytic cycles, such as integrases and lytic enzymes, but no antimicrobial resistance or virulence-associated determinants from which the bacterium can benefit were found; thus, their contribution may be recurrent in the protection against further infections by similar phages or the lytic activity of their lysins (Zhang & Stevens, 2021).

The virulence profiles observed in the strains were diverse (Fig. 3). All genomes harbour *aac(6′)-Ii*, *msrC*, and *efmA* genes associated with resistance to aminoglycosides and macrolides, consistent with previously reported findings (Holman et al., 2021). The clinical isolates showed a larger arsenal of virulence factors and antibiotic resistance-associated genes. However, the dairy-related strains CIBMA13 and CIBMA14 showed resistance genes to lincosamides and tetracycline that were not present in the two clinical strains that are resistant to vancomycin. Notably, strain D, which has been reported to have biotechnological properties for use as a starter culture in fermented dairy products, did not exhibit virulence factors in addition to the species' inherent antibiotic resistance genes (Olvera-Garcia et al., 2018). Genes associated with vancomycin resistance were detected only in clinical-associated ERV275 and ERV279 strains, in accordance with Holman et al. (2021), who reported that these genes were absent in *E. faecium* strains isolated in food-related environments, and neither displayed phenotypic resistance to vancomycin. However, *tetL* and *tetM* genes, which encode ribosome protection proteins and efflux pumps (Tao, Ying, Su, Zhou, & Sidhu, 2010), were detected only in dairy isolates, highlighting their ability to acquire virulence factors and genes that, thus far, have been found only in clinical isolates and may be disseminated in strains with biotechnological potential, compromising their suitability.

**Table 2**
Prophages detected in *Enterococcus faecium* assemblies.

| Strain | Best match | Accession | Length (kbp) |
|---|---|---|---|
| CIBMA12 | Entero_IME_EFm5 | NC_028826 | 19.5 |
| CIBMA13 | Lactob_phig1e | NC_004305 | 39.3 |
| | Entero_vB_IME197 | NC_028671 | 48.3 |
| | Lister_B025 | NC_009812 | 43.0 |
| CIBMA14 | Lister_B025 | NC_009812 | 45.2 |
| | Lactob_phig1e | NC_004305 | 39.3 |
| | Entero_vB_IME197 | NC_028671 | 43.9 |
| D | Bacill_phBC6A52 | NC_004821 | 36.8 |
| | Entero_vB_IME197 | NC_028671 | 48.1 |
| ERV275 | Entero_IME_EFm5 | NC_028826 | 19.9 |
| | Lactoc_TP901_1 | NC_002747 | 31.1 |
| ERV279 | Lister_2389 | NC_003291 | 25.1 |
| | Lactoc_ul36 | NC_004066 | 17.7 |
| | Bacill_BCJA1c | NC_006557 | 16.7 |
| MGRG | Entero_phiEf11 | NC_013696 | 34.7 |

The virulence factors harboured in the *E. faecium* genomes were mainly found in the clinical isolates and were associated with functions related to adhesion and biofilm formation. Interestingly, these attributes, often associated with pathogenicity, are also considered advantageous in probiotic strains of this species due to their potential to enhance persistence within the intestinal environment (Trunk, Khalil, & Leo, 2018). For instance, the *acm* gene, recognised for its role in adhesion, facilitates binding to type I and type IV collagen, a characteristic observed in strains isolated from diverse settings including dairy production, shrimp, and clinical contexts (Nallapareddy, Weinstock, & Murray, 2003). It is noteworthy that *Enterococcus* spp. strains, despite harbouring this gene, frequently exhibit limited binding capabilities, particularly in nonclinical isolates where the gene is typically inactive (Fu et al., 2022; Nallapareddy, Singh, Okhuysen, & Murray, 2008). Similarly, genes such as *sgrA*, *ecbA*, and *fss3*, encoding surface proteins that interact with fibrinogen's alpha, beta, and gamma chains, respectively, contribute to biofilm formation and the capacity to adhere to both abiotic surfaces and the host's extracellular matrix (Donlan, 2001; Hendrickx et al., 2009). It is important to recognise that the significance of these genes may vary widely among strains, as many within the *Enterococcus* genus exhibit beneficial characteristics. Consequently, the desirability of adhesion and biofilm formation traits depends on the overall profile and intended purpose of the specific strain under investigation (Scardaci et al., 2021).

Although *E. faecium* was previously regarded as an innocuous commensal microorganism, certain strains have been utilised as efficacious probiotics owing to their classification within the lactic acid bacteria group (Hanchi, Mottawea, Sebei, & Hammami, 2018). These strains can be detected in select cheeses and have been shown to play a role in the elimination of foodborne pathogens through the production of bacteriocins (Khan, Flint, & Yu, 2010). Nevertheless, it is pertinent to note that *E. faecium* has not been granted the Generally Regarded as Safe (GRAS) or Qualified Presumption of Safety (QPS) status (Rodríguez-Lucas & Ladero, 2023), indicating the need for further investigations to evaluate its safety and efficacy. Moreover, their presence in dairy products may not be desirable due to the production of biogenic amines (BAs), which are toxic compounds that can induce food poisoning (Ladero, Calles-Enriquez, Fernandez, & Alvarez, 2010). In addition, there have been documented instances of antibiotic-resistant enterococci with virulence factors that can contaminate the food supply, leading to direct or indirect transmission to humans (Gaglio et al., 2016). These antibiotic-resistant strains can act as reservoirs of resistance genes that have the potential to transfer to other pathogenic bacteria or strains better suited for infecting humans (Hernando-Amado, Coque, Baquero, & Martinez, 2019).

Our findings reveal that *E. faecium* strains are not restricted to their original location, as seen in their varied distribution among different subclades in the phylogenomic analysis but possess the ability to adapt to new environments. These results underscore the need for stringent monitoring of *E. faecium* strains and their genetic determinants in different settings, including the dairy industry, to mitigate the spread of antibiotic resistance and virulence genes and their attendant risks to public health.

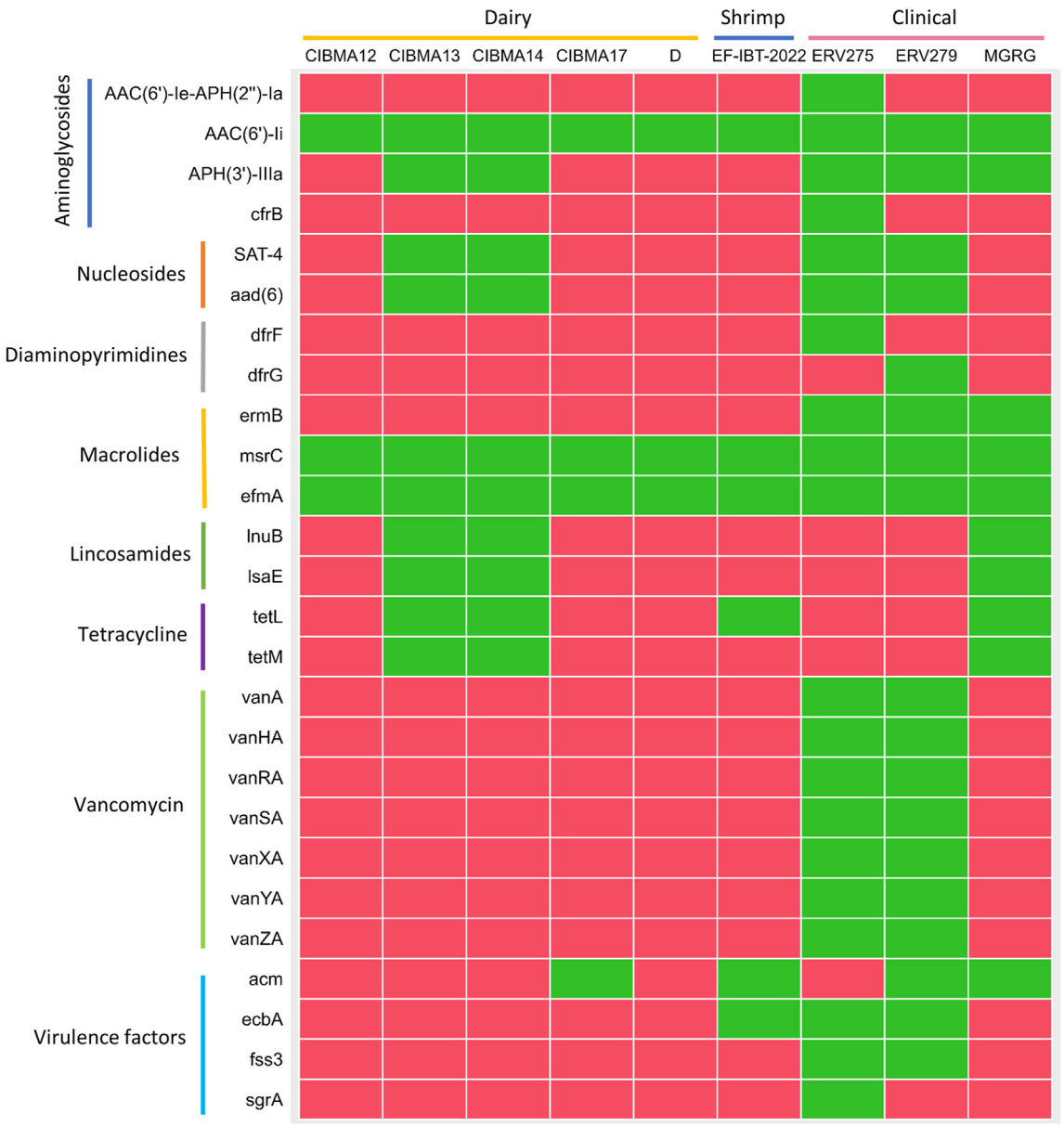


**Fig. 3.** Distribution of genes associated with antimicrobial resistance and virulence factors across *Enterococcus faecium* genomes. The presence of genes is represented with green and the absence of genes with red.

## 4. Conclusions

In conclusion, this study is the first comparative genomics study of *E. faecium* isolates from Mexico and demonstrates the importance of understanding the genetic relatedness and virulence profile of the strains from the dairy industry and highlights the importance of utilising WGS as a valuable tool for these purposes. The results revealed that some dairy-related strains shared genetic determinants of antibiotic resistance and prophages with the clinical isolates, indicating the potential for the transfer of virulence and resistance genes. This suggests that the dairy industry may serve as a reservoir for pathogenic strains of *E. faecium*, which could pose a risk to consumers of dairy products. Therefore, there is a need to implement effective control measures to avoid the occurrence of pathogenic strains of *E. faecium* in dairy products and to monitor the genetic content of strains present in the dairy plant environment to minimise potential risks.

## CRediT author statement

**Jean Pierre González-Gómez**: Conceptualization, Formal analysis, Methodology, Writing – original draft. **Maria Guadalupe Avila-Novoa**: Formal analysis, Validation, Writing – review & editing. **Berenice González-Tores**: Formal analysis, Methodology, Writing – review & editing. **Pedro Javier Guerrero-Medina**: Validation, Writing – review & editing. **Bruno Gomez-Gil**: Methodology, Writing – review & editing. **Cristobal Chaidez**: Conceptualization, Writing – review & editing. **Melesio Gutiérrez-Lomelí**: Conceptualization, Funding acquisition, Validation, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## Acknowledgements

We thank Julissa Enciso-Ibarra, Célida Isabel Martínez-Rodríguez, and Miriam Vega-Rodríguez for their technical support. The research presented in this article was conducted without any external funding or financial support.

## References

Agudelo Higuita, N. I., & Huycke, M. M. (2014). Enterococcal disease, epidemiology, and implications for treatment. In M. S. Gilmore, D. B. Clewell, Y. Ike, & N. Shankar (Eds.), *Enterococci: From commensals to leading causes of drug resistant infection*. Boston, MA, USA: Eye and Ear Infirmary.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics, 36*, 2251–2252.

Arias, C. A., & Murray, B. E. (2012). The rise of the *Enterococcus*: Beyond vancomycin resistance. *Nature Reviews Microbiology, 10*, 266–278.

Arndt, D., Marcu, A., Liang, Y., & Wishart, D. S. (2019). PHAST, PHASTER and PHASTEST: Tools for finding prophage in bacterial genomes. *Briefings in Bioinformatics, 20*, 1560–1567.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology, 19*, 455–477.

Ben Braiek, O., & Smaoui, S. (2019). Enterococci: Between emerging pathogens and potential probiotics. *Biomedical Research International, 2019*, Article 5938210.

Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2018). Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiology Reviews, 42*. Article fux053.

Brüssow, H. (2014). Impact of phages on evolution of bacterial pathogenicity. In M. J. Pallen, K. E. Nelson, & G. M. Preston (Eds.), *Bacterial pathogenomics* (pp. 267–300). Washington, DC, USA: ASM Press.

Chajecka-Wierzchowska, W., Zadernowska, A., & Garcia-Solache, M. (2020). Ready-to-eat dairy products as a source of multidrug-resistant *Enterococcus* strains: Phenotypic and genotypic characteristics. *Journal of Dairy Science, 103*, 4068–4077.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics, 34*, i884–i890.

Citak, S., Yucel, N., & Orhan, S. (2004). Antibiotic resistance and incidence of *Enterococcus* species in Turkish white cheese. *International Journal of Dairy Technology, 57*, 27–31.

de Been, M., van Schaik, W., Cheng, L., Corander, J., & Willems, R. J. (2013). Recent recombination events in the core genome are associated with adaptive evolution in *Enterococcus faecium*. *Genome Biology and Evolution, 5*, 1524–1535.

Donlan, R. M. (2001). Biofilm formation: A clinically relevant microbiological process. *Clinical Infectious Diseases, 33*, 1387–1392.

Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ, 3*, Article e1319.

Eren, A. M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S. E., Schechter, M. S., et al. (2021). Community-led, integrated, reproducible multi-omics with anvi'o. *Nature Microbiology, 6*, 3–6.

Florensa, A. F., Kaas, R. S., Clausen, P., Aytan-Aktug, D., & Aarestrup, F. M. (2022). ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics, 8*, Article 748.

Foulquie Moreno, M. R., Sarantinopoulos, P., Tsakalidou, E., & De Vuyst, L. (2006). The role and application of enterococci in food and health. *International Journal of Food Microbiology, 106*, 1–24.

Franz, C. M., Muscholl-Silberhorn, A. B., Yousif, N. M., Vancanneyt, M., Swings, J., & Holzapfel, W. H. (2001). Incidence of virulence factors and antibiotic resistance among Enterococci isolated from food. *Applied and Environmental Microbiology, 67*, 4385–4389.

Fu, X., Lyu, L., Wang, Y., Zhang, Y., Guo, X., Chen, Q., et al. (2022). Safety assessment and probiotic characteristics of *Enterococcus lactis* JDM1. *Microbial Pathogenesis, 163*, Article 105380.

Gaglio, R., Couto, N., Marques, C., de Fatima Silva Lopes, M., Moschetti, G., Pomba, C., et al. (2016). Evaluation of antimicrobial resistance and virulence of enterococci from equipment surfaces, raw materials, and traditional cheeses. *International Journal of Food Microbiology, 236*, 107–114.

Galloway-Pena, J., Roh, J. H., Latorre, M., Qin, X., & Murray, B. E. (2012). Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of *Enterococcus faecium*. *PLoS One, 7*, Article e30187.

Gelsomino, R., Vancanneyt, M., Condon, S., Swings, J., & Cogan, T. M. (2001). Enterococcal diversity in the environment of an Irish Cheddar-type cheese-making factory. *International Journal of Food Microbiology, 71*, 177–188.

Giraffa, G. (2003). Functionality of enterococci in dairy products. *International Journal of Food Microbiology, 88*, 215–222.

Goodhead, I., & Darby, A. C. (2015). Taking the pseudo out of pseudogenes. *Current Opinion in Microbiology, 23*, 102–109.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology, 59*, 307–321.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics, 29*, 1072–1075.

Hanchi, H., Mottawea, W., Sebei, K., & Hammami, R. (2018). The genus *Enterococcus*: Between probiotic potential and safety concerns – An update. *Frontiers in Microbiology, 9*, 1791.

Hendrickx, A. P., van Luit-Asbroek, M., Schapendonk, C. M., van Wamel, W. J., Braat, J. C., Wijnands, L. M., et al. (2009). *SgrA*, a nidogen-binding LPXTG surface adhesin implicated in biofilm formation, and *EcbA*, a collagen binding MSCRAMM, are two novel adhesins of hospital-acquired *Enterococcus faecium*. *Infection and Immunity, 77*, 5097–5106.

Hernando-Amado, S., Coque, T. M., Baquero, F., & Martinez, J. L. (2019). Defining and combating antibiotic resistance from one health and global health perspectives. *Nature Microbiology, 4*, 1432–1442.

Hollenbeck, B. L., & Rice, L. B. (2012). Intrinsic and acquired resistance mechanisms in *Enterococcus*. *Virulence, 3*, 421–433.

Holman, D. B., Klima, C. L., Gzyl, K. E., Zaheer, R., Service, C., Jones, T. H., et al. (2021). Antimicrobial resistance in *Enterococcus* spp. isolated from a beef processing plant and retail ground beef. *Microbiol Spectrum, 9*, Article e0198021.

Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics, 11*, Article 119.

Khan, H., Flint, S., & Yu, P.-L. (2010). Enterocins in food preservation. *International Journal of Food Microbiology, 141*, 1–10.

Kim, E. B., & Marco, M. L. (2014). Nonclinical and clinical *Enterococcus faecium* strains, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. *Applied and Environmental Microbiology, 80*, 154–165.

Kim, M., Oh, H. S., Park, S. C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology, 64*, 346–351.

Ladero, V., Calles-Enriquez, M., Fernandez, M., & Alvarez, M. A. (2010). Toxicological effects of dietary biogenic amines. *Current Nutrition & Food Science, 6*, 145–156.

Lebreton, F., van Schaik, W., McGuire, A. M., Godfrey, P., Griggs, A., Mazumdar, V., et al. (2013). Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio, 4*, Article 10-1128.

Letunic, I., & Bork, P. (2021). Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research, 49*, W293–W296.

Liu, B., Zheng, D., Zhou, S., Chen, L., & Yang, J. (2022). VFDB 2022: A general classification scheme for bacterial virulence factors. *Nucleic Acids Research, 50*, D912–D917.

Matos, R. C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., et al. (2013). *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. *PLoS Genetics, 9*, Article e1003539.

Mbanga, J., Amoako, D. G., Abia, A. L. K., Allam, M., Ismail, A., & Essack, S. Y. (2021). Genomic analysis of *Enterococcus* spp. isolated from a wastewater treatment plant and its associated waters in Umgungundlovu District, South Africa. *Frontiers in Microbiology, 12*, Article 648454.

Nallapareddy, S. R., Singh, K. V., Okhuysen, P. C., & Murray, B. E. (2008). A functional collagen adhesin gene, *acm*, in clinical isolates of *Enterococcus faecium* correlates with the recent success of this emerging nosocomial pathogen. *Infection and Immunity, 76*, 4110–4119.

Nallapareddy, S. R., Weinstock, G. M., & Murray, B. E. (2003). Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Molecular Microbiology, 47*, 1733–1747.

Nieto-Arribas, P., Sesena, S., Poveda, J. M., Chicon, R., Cabezas, L., & Palop, L. (2011). *Enterococcus* populations in artisanal Manchego cheese: Biodiversity, technological and safety aspects. *Food Microbiology, 28*, 891–899.

Ochman, H., & Davalos, L. M. (2006). The nature and dynamics of bacterial genomes. *Science, 311*, 1730–1733.

Ogier, J. C., & Serror, P. (2008). Safety assessment of dairy microorganisms: The *Enterococcus* genus. *International Journal of Food Microbiology, 126*, 291–301.

Olvera-Garcia, M., Sanchez-Flores, A., & Quirasco Baruch, M. (2018). Genomic and functional characterisation of two *Enterococcus* strains isolated from Cotija cheese and their potential role in ripening. *Applied Microbiology and Biotechnology, 102*, 2251–2267.

Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., & Toth, I. K. (2016). Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Analytical Methods, 8*, 12–24.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rice, L. B. (2008). Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *Journal of Infectious Diseases, 197*, 1079–1081.

Rodríguez-Lucas, C., & Ladero, V. (2023). Enterococcal phages: Food and health applications. *Antibiotics, 12*, Article 842.

Rouli, L., Merhej, V., Fournier, P.-E., & Raoult, D. (2015). The bacterial pangenome as a new tool for analysing pathogenic bacteria. *New Microbes New Infections, 7*, 72–85.

Sarantinopoulos, P., Andrighetto, C., Georgalaki, M. D., Rea, M. C., Lombardi, A., Cogan, T. M., et al. (2001). Biochemical properties of enterococci relevant to their technological performance. *International Dairy Journal, 11*, 621–647.

Scardaci, R., Manfredi, M., Barberis, E., Scutera, S., Marengo, E., & Pessione, E. (2021). Serotonin exposure improves stress resistance, aggregation, and biofilm formation in the probiotic *Enterococcus faecium* NCIMB10415. *Microbiology Research, 12*, 606–625.

van Dongen, S., & Abreu-Goodger, C. (2012). Using MCL to extract clusters from networks. *Methods in Molecular Biology, 804*, 281–295.

van Kranenburg, R., Kleerebezem, M., van Hylckama Vlieg, J., Ursing, B. M., Boekhorst, J., Smit, B. A., et al. (2002). Flavour formation from amino acids by lactic acid bacteria: Predictions from genome sequence analysis. *International Dairy Journal, 12*, 111–121.

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

Srinivasan, L., & Evans, J. R. (2018). Health care-associated infections. In C. A. Gleason, & S. E. Juul (Eds.), *Avery's diseases of the newborn* (pp. 566–580). Amstrdam, the Netherlands: Elsevier.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research, and development of new antibiotics: The WHO priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet Infectious Diseases, 18*, 318–327.

Tao, R., Ying, G. G., Su, H. C., Zhou, H. W., & Sidhu, J. P. (2010). Detection of antibiotic resistance and tetracycline resistance genes in *Enterobacteriaceae* isolated from the Pearl rivers in South China. *Environmental Pollution, 158*, 2101–2109.

Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research, 28*, 33–36.

Trunk, T., Khalil, H. S., & Leo, J. C. (2018). Bacterial autoaggregation. *AIMS Microbiology, 4*, 140–164.

Zhang, H., & Stevens, R. H. (2021). Intrinsic resistance of *Enterococcus faecalis* strains to PhiEf11 phage endolysin is associated with the presence of PhiEf11 prophage. *Archives of Virology, 166*, 249–258.