Gut Microbes

ISSN: 1949-0976 (Print) 1949-0984 (Online) Journal homepage: www.tandfonline.com/journals/kgmi20

Comprehensive analysis of *Enterococcus* spp. from two European healthy infant cohorts shows stable genomic traits including antimicrobial resistance (AMR)

Muhammad Hassan Saeed, Magdalena Kujawska, Kristiana Ellen, Antia Acuna-Gonzalez, Manuel Bernabeu Lorenzo, Maria Carmen Collado, Monique Mommers, Niels van Best, Lindsay J. Hall & Klaus Neuhaus

To cite this article: Muhammad Hassan Saeed, Magdalena Kujawska, Kristiana Ellen, Antia Acuna-Gonzalez, Manuel Bernabeu Lorenzo, Maria Carmen Collado, Monique Mommers, Niels van Best, Lindsay J. Hall & Klaus Neuhaus (2025) Comprehensive analysis of *Enterococcus* spp. from two European healthy infant cohorts shows stable genomic traits including antimicrobial resistance (AMR), Gut Microbes, 17:1, 2516699, DOI: 10.1080/19490976.2025.2516699

To link to this article: https://doi.org/10.1080/19490976.2025.2516699

GUT MICROBES
2025, VOL. 17, NO. 1, 2516699
https://doi.org/10.1080/19490976.2025.2516699

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

OPEN ACCESS Check for updates

Comprehensive analysis of *Enterococcus* spp. from two European healthy infant cohorts shows stable genomic traits including antimicrobial resistance (AMR)

Muhammad Hassan Saeed [a,b], Magdalena Kujawska [b,c], Kristiana Ellen [a], Antia Acuna-Gonzalez [d], Manuel Bernabeu Lorenzo [e], Maria Carmen Collado [e], Monique Mommers [f], Niels van Best [g,h], Lindsay J. Hall [b,c,d], and Klaus Neuhaus [a]

[a]Core Facility Microbiome, ZIEL – Institute for Food & Health, Freising, Germany; [b]Intestinal Microbiome, TUM School of Life Sciences, Freising, Germany; [c]Microbes, Infection, and Microbiomes, College of Medicine and Health, University of Birmingham, Birmingham, UK; [d]Food, Microbiome and Health, Quadram Institute Bioscience, Norwich Research Park, Norwich, UK; [e]Department of Biotechnology, Institute of Agrochemistry and Food Technology-National Research Council (IATA-CSIC), Valencia, Spain; [f]Care and Public Health Research Institute (CAPHRI), Department of Epidemiology, Maastricht University, Maastricht, The Netherlands; [g]School of Nutrition and Translational Research in Metabolism (NUTRIM), Department of Medical Microbiology, Infectious Diseases and Infection Prevention, Maastricht University Medical Centre, Maastricht, The Netherlands; [h]Institute of Medical Microbiology, RWTH Aachen University Hospital, Aachen, Germany

ABSTRACT

Enterococcus spp. some of which are pathogenic, are common gut microbiota members, including also infants. Infants may be more susceptible to *Enterococcus* due to their developing gut ecosystems. It is unclear whether antibiotic resistance genes (ARGs) and certain genomic traits in enterococci are restricted to the human subpopulation or more widespread. Furthermore, the correlation between these traits and geographic variation is poorly understood. Therefore, we sequenced 100 strains isolated from full-term healthy infants' fecal samples from two geographically distant European cohorts (MAMI in Spain and LucKi from the Netherlands) to explore the diversity of *Enterococcus* spp. within the infant's gut microbiome and assess cohort-specific traits such as ARGs. Most isolates were *E. faecalis* and *E. gallinarum*, with a total of 11 species identified. We found a rich reservoir of ARGs, plasmids, prophages and virulence factors in the infant strains, with minimal cohort-specific differences in resistome profiles. In addition, Epx, a pore-forming toxin associated with pathogenicity, was found in *E. hirae* strains. While metabolic profiles were similar across cohorts, *E. faecalis* strains harbored more virulence genes and prophages compared to other species. An analysis of public *Enterococcus* genomes revealed that multi-drug resistant (MDR) strains exist without any significant geographic or temporal pattern. Phenotypic resistance analysis indicated that 28% of MAMI strains were gentamicin resistant, compared to 5% of the strains from the LucKi cohort, though LucKi isolates were also resistant to other antibiotics. We also selected ten *E. faecalis* isolates with varying virulence gene repertoires for phenotypic virulence testing in *Caenorhabditis elegans* and found them killing at various rates, however no clear pattern emerged in correlation with any specific genetic determinant. Overall, our results suggest that *Enterococcus* spp. including ARGs, are highly mobile across Europe and beyond. Their adaptability likely facilitates long-distance dissemination, with strains being acquired early in life from community environments.

ARTICLE HISTORY
Received 19 February 2025
Revised 12 May 2025
Accepted 22 May 2025

KEYWORDS
Enterococcus; AMR; resistome; healthy infants; phylogenomics; genomics

Introduction

The *Enterococcus* genus is highly versatile, found in a wide range of habitats, and known for its resilience in surviving harsh conditions.[1] *Enterococcus* is also one of the first colonizers of the newborn gut, and remains a 'core' member of the human gut microbiota across life; typically representing ~1% of the fecal microbial community and accounting for between 10^4 to 10^6 microorganisms per gram wet weight.[2,3] While some *Enterococcus* strains have beneficial properties and are even used as probiotics,[4] many are opportunistic pathogens, commonly found in both human and animal guts. Currently, 63 valid *Enterococcus* species have been published, with 14 additional species awaiting validation.[5]

Globally, antimicrobial resistance (AMR) is recognized as one of the major 'One Health'

CONTACT Klaus Neuhaus neuhaus@tum.de Core Facility Microbiome, ZIEL – Institute for Food & Health, Weihenstephaner Berg 3, Freising 85354, Germany

Supplemental data for this article can be accessed online at https://doi.org/10.1080/19490976.2025.2516699

issues.[6] *Enterococcus*, being widely distributed in nature, with its persistence, occurrence, and prevalence, is an important target to study AMR from the One Health perspective. Often, it is considered as an indicator of human and animal fecal contamination in the environment.[7,8] The rise of enterococci as nosocomial pathogens is likely linked to their increase in virulence determinants.[9] Brooks et al.[10] emphasized determinants of the early gut microbiome as 'hospital rooms' along with diet, mode of delivery, and antibiotics use. For instance, they found some strains in hospitalized infants across cohorts and years, which were also present in sinks and surfaces in the hospitals. Thus, besides others, significant risk factors for enterococcal infections include hospital stays, antibiotic exposure and immunocompromised patients and patients with cancer, diabetes, urinary tract infections, abdominal surgery and chronic kidney disease.[11] Once the gastrointestinal tract (GIT) is compromised, the risk of enterococcal infection is increased.[12]

The ecological adaptation of the enterococci causes their presence in various other niches, for example including the human bladder,[13] wild birds,[14] wastewater treatments,[15] and in foods as well,[16] to name a few. Some enterococci tolerate high salt concentrations[7] and they are found for instance in Turkey's traditional cheeses.[17] For spreading genetic determinants in the genus (and beyond), plasmid, and other mobile genetic elements (MGEs) play an important role.[18] Thus, exemplary, many different virulence determinants are seen in clinical *E. faecalis* and *E. faecium* in Bulgaria[19] and elsewhere. Various *Enterococcus* ssp. have been reported in Australian poultry, which has been suggested as reservoir.[20] Similarly, on the other side of the globe, *Enterococcus* were found associated to disease in French poultry.[21] Taken together, generalist to specialist enterococcal species were described, including 18 previously unknown species in an analysis of *Enterococcus* across wide range of hosts within different ecosystems and geographies. These findings conclusively demonstrate the high diversity and the huge transmissibility of this genus and its genes.[22]

Pregnancy and early life represent periods of particularly high antibiotic use, particularly for treating infections in premature infants. This antibiotic exposure disrupts the developing gut microbiome, sometimes leading to overgrowth of potentially pathogenic strains such as *Enterococcus*,[23] and increasing the risk of secondary infections.[24] Antibiotic treatment can also influence the overall antimicrobial 'resistome' of the gut microbiome, which represents an ideal 'melting pot' for genetic exchange,[25] where antimicrobial resistance genes (ARGs) can persist even after cessation of antibiotics, including in *Enterococcus* spp.[26–28] Previous work indicates that some factors, such as birth mode, influence the resistome of full term infants, and multi-drug resistant (MDR) bacteria from the environment, particularly hospitals, may further contribute to colonization with these bacteria.[29] Besides AMR, these bacterial genomes may also encode phages and virulence factors like adhesions, toxins and capsule genes. These genes contribute to enhanced colonization of such strains in the gut, since they allow a more effective competition against their commensal rivals.[30–32] A recent study identified *Enterococcus'* pore-forming toxins (Epxs): Epx1 and Epx3 in *E. faecalis*, Epx2 and Epx7 in *E. faecium*, and Epx4, Epx5, Epx6, and Epx8 in basically all *E. hirae* strains. Such pore-forming toxins are aggregated into the common class of bacterial toxins. While Epx2 and Epx3 recognize human's MHC-I, but also from equine, bovine, and porcine, the MHC-I of mice is not.[33] In the past, the nematode *C. elegans* was found to be a suitable model organism to study host-microbe interactions including virulence[4,34,35] and we also use this model to obtain information about the overall virulence of selected strains.

To combat the growing threat of AMR, it is important to explore the resistome of particular gut microbes in infants, and understand how ARGs are acquired and shared. While, many studies have focused on preterm infants with high antibiotic exposure, healthy full term infants, who typically receive minimal antibiotics, are understudied in the context of *Enterococcus*. Given the role of this genus as a reservoir for MDR strains and its involvement in serious infections, we sought to determine strain level diversity, the wider genomic landscape, AMR genotypes and phenotypes in two healthy infant populations. For this study, we analyzed *Enterococcus* isolates from the fecal samples of two geographically distinct European cohorts:

MAMI (Spain) and LucKi (the Netherlands). We conducted metabolic estimation and predicted enriched functions for *E. faecalis*, the most prominent and pathogenic species, to understand determinants and mechanisms of adaptation, persistence, and resistance. In addition, to understand 'global' trends, we incorporated public *Enterococcus* genomes from various locations and hosts. Finally, since AMR and virulence factors should be tested phenotypically, we assessed antibiotic resistance via broth microdilution and virulence using *C. elegans*.

Methods

Cohort description

Two European birth cohorts were used in this study (Supplementary Table S1). The MAMI birth cohort focused on mother-infant microbiota during early life and located in the Spanish-Mediterranean area.[36] We obtained a subset of 18 infants' fecal samples, which were collected at a single time point, i.e., 4 months. Infants' data was collected using questionnaires on health and medication. The MAMI study was approved by all participating hospital's Hospital ethics committees (HECs) and Atención Primaria – Generalitat Valenciana (CEIC-APCV) and registered on ClinicalTrial.gov under NCT03552939.

The second cohort, the LucKi-Gut Study, is an ongoing longitudinal birth cohort study aiming at determinants of early-life microbiome development and the association with childhood health outcome. Questionnaires were used to collect, among others, data on infant's lifestyle, health, medication use, as well as on maternal health (during pregnancy), medication use, and diet. For the present study, we included 58 stool samples from 35 infants, collected at 8 weeks, 6 months, and 11 months. The LucKi-Gut Study was approved by the Maastricht University Medical Centre's Medical Ethical Committee (METC 15–4–237).

Samples and bacterial isolation

All infant fecal samples were used to isolate bacteria (Suppl. Table S1). A 10-µl loop of the fecal material was suspended in phosphate buffer saline (PBS) and used to produce ten-fold serial dilutions. An aliquot of 100 µL of the diluted samples from 10^{-3} and 10^{-4} were plated on Brain Heart Infusion (BHI) agar plates and incubated aerobically at 37 °C for 24 h. From those plates, colonies were randomly picked and streaked to purity on BHI. Pure cultures were stored at −80 °C as glycerol stocks.

DNA extraction and sequencing

DNA was extracted from pure bacterial cultures using MP Bio FastDNA™ SPIN Kit according to manufacturer's guidelines for 16S rRNA gene sequencing. PCR was conducted using a mix of universal forward primer FD1 (5'-AGA GTT TGA TCC TGG CTC AG-3'), FD2 (5'-AGA GTT TCA TGG CTC AG-3') together with reverse primer RP1 (5'-ACG GTT ACC TTG TTA CGA CTT-3').[37] These primers are comparable to 27F and 1492 R,[38] producing a near full-length 16S rRNA gene amplicon. PCR products were sent to Eurofins (Ebersberg, Germany) for Sanger sequencing. Sequences were aligned against the nucleotide sequence database nr in GenBank using BLASTn, to confirm whether the isolated strain belonged to *Enterococcus*. A total number of 21 and 79 *Enterococcus* spp. isolates from MAMI and LucKi were obtained, respectively. DNA of confirmed *Enterococcus* spp. was subjected to shotgun whole genome sequencing on Illumina NextSeq platform at the Quadram Institute Bioscience (Norwich, UK).[39]

Genome assembly, annotation, and search for prophages

Fastp v0.23.2 was used for initial raw data cleanup using raw sequencing reads.[40] Bacterial genomes were assembled using UniCycler v0.4.9.[41] CheckM v1.2.0 was used for information on contamination and completeness.[42] Following Bowers et al.,[43] we kept genomes with a completeness >99%, contamination <5% and coverage ≥55×. GTDB-Tk v2.3.2 provided bacterial taxonomic classification.[44] *Enterococcus* genomes were annotated by Prokka v1.14.6 for identifying features of interest.[45] The genomic features were determined using a shell script.[46] PHASTEST was used to detect prophages. For this, a "completeness score" is assigned based on

proportion of phage genes in identified region. The score above 90 is taken as "complete" or "intact", while 70–90 is taken as "questionable" (if intact) and <70 as "incomplete" phages.[47–49] FastANI v1.33 is used for average nucleotide identity (ANI) analysis[50] (Suppl. Table S2). Isolates were considered identical above 99.99% threshold values and duplicate strains were removed prior to further analysis. During the submission of genome sequences to NCBI, the Prokaryotic Genome Annotation Pipeline (PGAP) was run and revealed that three of our unidentified strains were *E. entomosocium*.[51]

Sequence typing using multi-locus sequence typing (MLST)

Sequence typing was conducted for *E. faecalis and E. faecium*. Here, we used MLST v2.0.9 from the Center for Genomic Epidemiology (CGE), run online (19 Oct 2023) on each individual genomes. This tool uses BLAST-based ranking method for best matching MLST alleles of specified MLST scheme. The scheme is based on combination of the following seven genes with dispersed locations on the chromosome (i.e., minimal distance between loci, 137 kb) for *E. faecalis* as following: *gdh* (glucose-6-phosphate dehydrogenase), *gyd* (glyceraldehyde-3-phosphate dehydrogenase), *pstS* (phosphate ATP binding cassette transporter), *gki* (putative glucokinase), *aroE* (shikimate 5-dehydrogenase), *xpt* (shikimate 5-dehydrogenase), and *yiqL* (acetyl-coenzyme A acetyltransferase). For *E. faecium*, it includes *gdh* and *gyd*, as before in *E. faecalis*, but also other genes as *adk* (adenosine kinase), *atpA* (ATP synthase, subunit alpha), *ddl* (D-alanine – ligase), *pstS* (phosphate-binding protein), and *purK* (N5-carboxyaminoimidazole ribonucleotide synthase).[52]

Genomics analysis

A tree was created using Mashtree v1.2.0 based on mash distances (i.e., based on distances between any two genomes) using the fasta assembled genomes as input.[53] Further, as outgroup, we added a strain from same taxonomic order, but of a different genus (i.e., *Clostridium perfringens*). Whole genome sequencing data was used for prediction of putative AMR profiles. Here, ABRicate v1.0.0 was used for screening of ARGs.[54] For the AMR detection, minimum coverage and identity were set at 95% each. ABRicate uses pre-downloaded databases and we selected CARD (Comprehensive Antibiotic Resistance Database), which is ontology based and provides information on AMR genes and their classes.[55] Virulence factors were searched using Virulence Factors Database (VFDB) at 90% identity and 80% coverage.[56,57] Finally, plasmids were detected using ABRicate's PlasmidFinder at 80% identity and 50% coverage. Sequences, which were indicated to contain a plasmid replicon by PlasmidFinder in our isolates were verified using BLAST with the nr database of Genbank. Almost all hits were plasmids indeed (Suppl. Table S3). The presence of Epxs (Exp1 to Exp8) were detected using BLAST using their reference sequences from NCBI against all infant and public genomes (Supplementary Table S4a).

Comparative genomics for public genomes

The same approach as before was executed to study all available complete and high-quality genomes at NCBI, using the same parameters for the tree as above (Suppl. Table S4b). About 30,000 records were available as of October 2023. These were quality-screened with CheckM and > 90% completeness and < 5% contamination criteria were used, which left 622 entries with complete and high quality genomes. Genomes with an ANI above 99.99% were considered identical and removed (Suppl. Table S5). After that, 587 genomes were retained that were considered to be of sufficiently high quality.[43] This dataset (designated "public") was used for phylogenomic analysis concerning ARGs, plasmids and virulence factors. A tree was built using both our 100 isolates' genomes and the 587 public *Enterococcus* genomes. Of note, for the public genomes, the associated metadata are sometimes limited concerning the health status of the hosts from which the isolates were obtained. Otherwise, genomes stored in RefSeq and their meta data were downloaded (19 Oct 2023) using the packages ncbi-genome-download v0.3.3 and datasets v15.24.0 from NCBI respectively.[58,59]

Pangenomic analysis of E. faecalis (our infant strains)

The pipeline anvi'o v8 was used to generate the pangenome of *E. faecalis*.[60] A text file was created

on the information of our sets of genomes to generate a genome storage database. We then computed the pangenome using the program 'anvi-pan-genome' (parameters – minbit 0.5 –mcl-inflation 10 –use-ncbi-blast).[61–63] Later we used 'anvi-display-pan' to visualize the pangenome along with an ANI dendogram. To calculate the number of accessory and core genes, we applied 'anvi-script-compute-bayesian-pan-core'.[64] The contigs and genome database generated using anvi'o was enriched with several annotation sources including COG (Cluster of Orthologous Genes), KEGG (Kyoto Encyclopedia of Genes and Genomes), KOfam (a database of KEGG Orthologs; KO), and CAZyme (Carbohydrate Active enZymes).

Metabolic estimation and functional enrichment of E. faecalis *(our infant strains)*

As basis for this analysis, we firstly used the dataset of all *E. faecalis* genomes available in this study. We calculated the path-wise completeness for a given KEGG module[65,66] in using 'anvi-estimate-metabolism'.[67] A heat map was made using visualization module of anvi'o with a path-wise completeness matrix. First, a Newick tree was generated with 'anvi-matrix-to-newick' and name, category, subcategory, and class of each module with groups (of the genomes) were added as an additional layer. This was imported to our profile database and later used for visualization in interactive mode. Next, 'anvi-compute-functional-enrichment' determined whether a given module is present in a particular genome or group (i.e., MAMI *vs*. LucKi) by fitting a binomial generalized linear model (GLM). For finding enriched genes or gene functions, the unadjusted p-values were used. Later, the enrichment score is displayed along with p-values.[68] Afterward, 'anvi-script-gen-function-matrix-across-genomes' was used to analyze the presence/absence of functions in each of the genomes across groups. Both, metabolic estimations and functional matrices were visualized using 'anvi-interactive' and 'anvi-display' functions, respectively. Only 30 enriched functions that had p-values less than 0.05 were considered and visualized.

Phenotypic screening of antibiotic resistance

For resistance screening, the broth microdilution method was employed using gentamicin, vancomycin, amoxicillin, and linezolid.[69] Briefly, minimum inhibitory concentration (MIC) testing is carried out in 96-well microtiter plate using twofold method. Diluting an overnight broth culture of the required organisms accordingly to 0.5× McFarland Standard results in about 1.5×10^8 CFU/ml. The antibiotic stock was diluted from 64 µl to 0.25 µl for all antibiotics, but gentamicin; this was diluted from 256 µl to 1 µl. A positive control, without antibiotic, and a negative control, without inoculum, were included. MICs were calculated and interpreted according to European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines.[70]

Phenotypic virulence testing of selected E. faecalis *strains using nematodes*

Based on virulence gene profiles, ten *E. faecalis* providing a diverse range of different virulence genes were selected. *C. elegans* N2 were synchronized as follows. Ten milliliters of a freshly prepared alkaline hypochlorite solution (2:5 12% NaOCl to 2 M NaOH) was added to a suspension of gravid worms in 10 ml sterile water. The mixture was vortexed for about 30 s and left on a rocking platform at room temperature. After 6 min, 30 ml of M9 buffer was added. The tubes were then centrifuged for 1 min at 12,500×*g*, and the resulting egg pellet was washed in M9 buffer twice. The egg pellet is taken up in 1 ml M9 buffer and amounts are transferred to an empty 35-mm NGM agar plate and incubated at 22°C for 24 h until L1 larvae hatch. These larvae enter L1 diapause due to the absence of food. Arrested L1 are transferred onto NGM agar plates with *E. coli* OP50 and incubated at 22°C for 3 d until the nematodes reach stage L4. Ten 35-mm NGM agar plates were prepared with the bacteria and ten worms per plate were added (i.e., in total 100 worms per *Enterococcus* strain). Worms were transferred each day to freshly prepared plates until they were dead. Worms were considered dead if they did not respond to touch. Missing worms or worms killed by handling were not considered. The TD50 was determined by calculating the daily percentage of dead worms and fitting a sigmoidal curve using drc[71] in R v4.3. Using the Rhea R scripts,

correlations between the number of virulence factors, plasmids, and prophages were computed against observed TD50.[72]

Results

Isolation of *Enterococcus* spp. from the infant cohorts

Of all 100 isolates, eleven distinct species, i.e., *E. faecalis* (60), *E. gallinarum* (9), *E. casseliflavus* (7), *E. avium* (5), *E. faecium* (3), *E. gilvus* (3), *E. entomosocium* (3), *E. durans* (3), *E. lactics* (3), *E. hirae* (2), and *E. raffinosus* (2) were obtained from the healthy infants' fecal samples. All assembled *Enterococcus* genomes had an average of 58 contigs and genome sizes range from an estimated 2.6 Mbp to 4.3 Mbp. Genomes displayed an average GC content of 38.4%, with a minimum of 36.5% for *E. hirae* and the highest observed value at 42.7% for *E. casseliflavus*. However, only three species were isolated from the MAMI cohort, namely *E. faecalis, E. gallinarum* and E. *faecium*, but all eleven above mentioned species were found in the LucKi cohort. Of note, the samples of the LucKi cohort comprised more time points compared to the MAMI cohort (Suppl. Table S6).

Only a small number of infants from MAMI, e.g., 3 out of 18 (either at day 7 or 15) and only 6 from 35 LucKi received antibiotics (either in 9th or 11th month); all unrelated to gut issues. The socio-economic background and geographical locations of both cohorts differ significantly,[36,73] leading us to expect substantial differences in strains isolated from each cohort.

In both cohorts, *E. faecalis* was the prevalent species with a total of 60 isolates, followed by *E. gallinarum* (9 strains) and *E. casseliflavus* (7 strains, but only present in LucKi). The sequence type was determined for *E. faecalis* and *E. faecium* using multi-locus sequence typing (MLST). The 60 *E. faecalis* isolates were classified into 24 sequence types (ST) including some unknown ST types not previously observed. The types most frequently found were ST179 (22 isolates), ST191 (5 isolates), ST16 (5 isolates), and further 3 isolates were indicated as 'unknown ST'. In addition, the three isolated *E. faecium* strains, each belonged to a different ST type, namely ST80, ST214, and a further unknown ST (Suppl. Table S6).

Phylogenetic relatedness of the infant strains

We conducted a genome analysis of all infant *Enterococcus* isolates to understand their phylogenetic relationships, and to get an overview of resistome, virulence factors, plasmids, and prophages diversity, all of which contribute to the pathogenicity of this genus. We compared genomic sequences of the 100 infant strains, with the addition of the 11 type strains for each species present (Suppl. Table S7). A distance-based association tree (Figure 1), enriched with the datasets concerning resistance genes, plasmids, virulence genes, and the number of prophages, showed no clear distinction between traits, cohorts, or reference genomes. Interestingly, we observed that across both cohorts, all vaginally born infants harbored had the same species. In contrast, two infants from LucKi, which were born by Cesarean-section (C-section) displayed a different set of species, but *E. gilvus* was common for the C-section children.

The most prevalent species, *E. faecalis*, showed a number of subclades, which was different compared to the other species, which mainly showed a single branch in the tree (e.g., *E. gallinarum, E. casseliflavus, E. faecium, E. lactis, E. hirae, E. durans, E. gilvus*, and *E. avium)*, but this could be due to the lower number of strains isolated. Only *E. raffinosus* seems to have diverged less from the common ancestor.

Prediction of AMR in infant strains

Genomic analysis for resistance determinants predicted a range of different AMR genes for the majority of isolates (83%), across both cohorts. Most AMR genes conferred predicted resistance to aminoglycosides, glycopeptides, macrolides, tetracylines, and fluoroquinolones. *E. faecalis* had the highest abundance of AMR genes compared to any other species (Suppl. Table S6).

The *in silico* prediction of *aac(6')-aph(2")* genes validated the phenotypically confirmed gentamicin-resistance in ten *E. faecalis* strains (see below). Notably, the most abundant plasmid, DOp1, was present in most *E. faecalis* strains, including all phenotypically gentamicin-resistant strains, alongside several ARGs (specifically *tet(M)*) in resistant strains. However, the DOp1 plasmid appears to have no

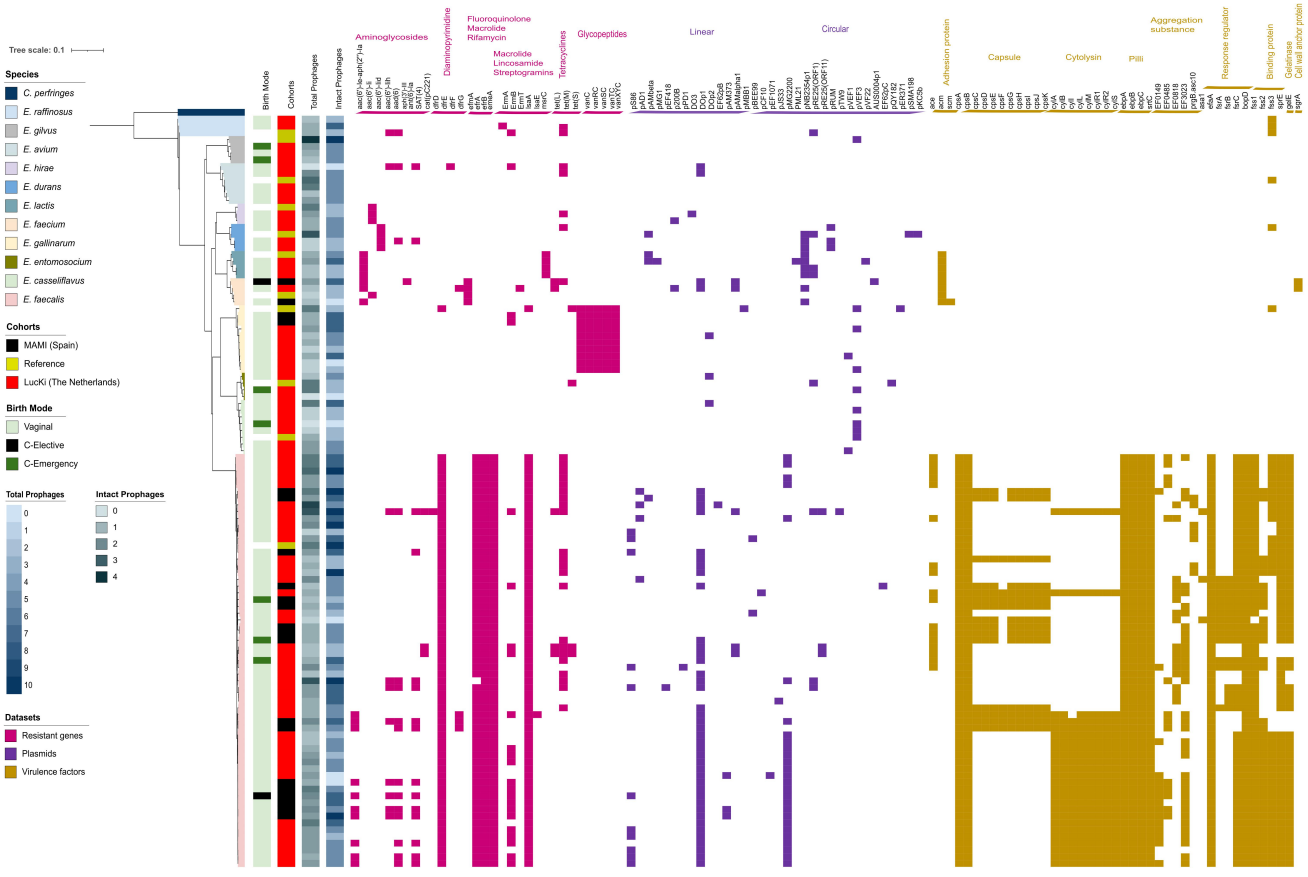

Figure 1. Left, cladogram (rooted at midpoint) using mash distances, based on distances between any two genomes of *Enterococcus*, using *C. perfringens* as outgroup. The tree is color coded for the species. Next, columns indicate birth mode – vaginal (light green), C-emergency (dark green) and C-elective (black); the cohort genomes for MAMI (black), LucKi (red) and reference genomes (yellow); the number of total and intact prophages (darker color, more prophages). The matrix to the right shows, presence (color) or absence (no color) for antimicrobial resistance genes, according to CARD (pink); plasmids, according to PlasmidFinder (purple); and virulence genes according to VFDB (brown).

direct relationship with gentamicin resistance, but with the presence of the *tet(M)* gene. No amoxicillin resistance genes were detected *in silico*, all *E. gilvus* and *E. faecium* strains displayed phenotypic resistance to this antibiotic (Figure 2). Vancomycin Resistant Enterococci (VRE) associated with severe outbreaks normally carry *vanA* and *vanB*, but these genes were absent.[74] However, we detected the *vanC* gene and its variants (*vanC1XY*, *vanC2XY*, *vanC4XY*, but also *vanRC*, *vanSC*, and *vanTC*) in nine *E. gallinarum* strains (Figure 1), all of which exhibited phenotypic vancomycin resistance (see below).

Virulence factors in infant strains

All *Enterococcus* genomes were examined for known virulence genes. In total, we detected 43 different virulence factors across 9 categories including capsule, adhesin protein, aggregation, response proteins, gelatinase, cytolysin, pilus proteins, and cell wall anchor proteins. The presence of *ebpA-C*, *acm*, *scm*, *cpsA-K*, *efaA* and *sgrA* indicates a potential pathogenic capacity in strains carrying these determinants, as these genes are associated with functions like collagen adhesion, endocarditis specific antigen, and biofilm-associated proteins. The number of virulence factors varied between species and strains. Interestingly, *E. faecalis* had the highest number of virulence genes, ranging from 30 to 36, which were largely absent from other species. Since *E. faecalis* contained an extensive array of virulence factors, including *sprE* and *gelE* – both known to increase killing in nematodes[75] these strains may possess a significant, albeit opportunistic, virulence

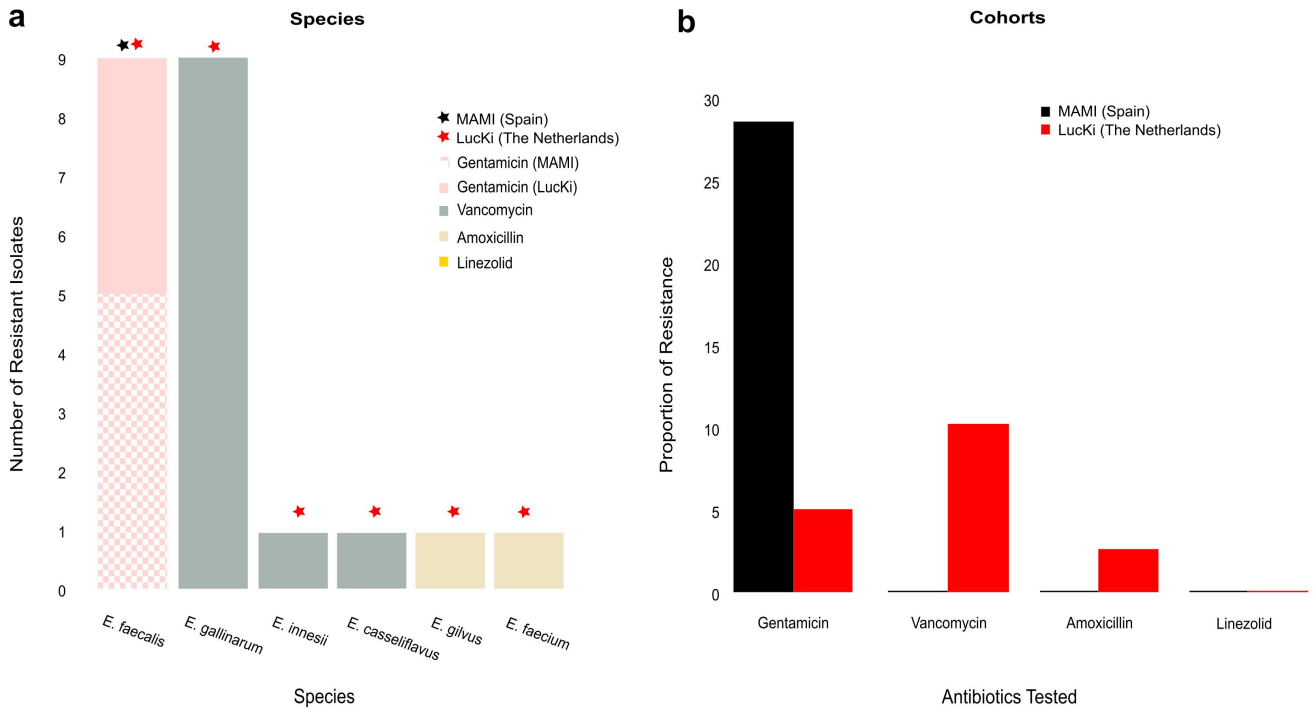

Figure 2. Phenotypic antibiotics resistance profiles of enterococci from the two cohorts. All strains were sensitive to linezolid. (a) Number of resistant isolates within each species, broken down into antibiotics (non-susceptible species against each antibiotics are not shown); black asterisk to indicate MAMI and red asterisk for LucKi; for gentamicin – its MAMI (pink) and LucKi (checkered pink). (b) Comparison of resistant strains concerning cohorts for tested antibiotics. Black; MAMI and Red; LucKi.

potential, despite the generally healthy status of infants in the cohorts. Among all genomes analyzed, 31 strains did not possess any virulence genes. These strains belonged to species such as *E. gallinarum*, *E. casseliflavus*, *E. avium*, *E. hirae, E. entomosocium*, and *E. gilvus*. When comparing isolates from both cohorts, 41 virulence factors were common between them, while 1 was unique to each group (*scm* in MAMI and *asa1* in LucKi). The virulence gene *acm* (collagen binding adhesion), which is considered to be of lesser importance in infections, was present in only about 6% of *E. faecium* and *E. lactis* strains. The adhesion-conferring gene *scm*, was only found in one *E. faecium* strain. In addition, we detected two Epxs that belong to a family of the recently detected novel *Enterococcus* pore-forming toxins. We found Epx7 and Epx8 in one of our *E. hirae* strains (i.e., LH-NE-35). This strain, however, does not have any other known virulence factors. However, it contains few ARGs, phages, and a plasmid.

Plasmid carriage in infant strains

Plasmids play a key role in bacterial ecology, often acting as vectors for the transfer of ARGs. We found 28 different plasmids across 76 out of 100 strains. Interestingly, eight of these plasmids were present in both cohorts. DOp1 was the most abundant plasmid, followed by pMG2200 and pS86. In the MAMI cohort, 10 plasmids were detected in 15 MAMI strains, accounting for 31 occurrences. In contrast, LucKi cohort strains had 26 different plasmids, with 118 occurrences in 61 strains. Thus, when a strain carries one plasmid, there is a 50% chance it carries a second plasmid; a pattern observed in both cohorts. Four species were found to be notorious for carrying a second plasmid: *E. faecalis*, *E. faecium*, *E. durans*, and *E. lactis*. The plasmids frequently carried AMR-associated genes, for instance, *lsa(A)*, *aac(6')-aph(2")*, and *tet-(M)*, which confer resistance to lincomycin, clindamycin, gentamicin, amikacin, tetracycline, doxycycline, and minocycline. In addition, some pheromone-responsive plasmids were detected,

which trigger expression of virulence determinants like cytolysin and pili, such as pAM373, pAD1, pTEF2, and pCF10, thereby contributing to the pathogenicity of the bacterial host.[76] Figure 1 summarizes the plasmid findings.

Two isolates (*E. lactis* and *E. faecalis*) carried five plasmids each, including a shared plasmid pAMbeta, the highest number of plasmids identified in a single strain. Only two isolates from the LucKi cohort (*E. casseliflavus* and *E. entomosocium*) shared the same plasmid (pVEF3), which was associated with glycopeptide resistance. Interestingly, *E. entomosocium*, from three different LucKi infants also exhibited glycopeptide resistance, with two of these strains showing phenotypic vancomycin resistance. The plasmid DOp1 confers resistance to tetracycline, while pVEF provides vancomycin resistance with pVEF3 detected in four out of six phenotypic-vancomycin resistant enterococci. Additionally, pVEF3 was found in *E. entomosocium* and pDOp1 in *E. faecalis*, both isolated from infants born by C-section. Notably, 24 strains did not carry any plasmids (8 *E. faecalis*, 5 *E. gallinarum*, 3 each of *E. gilvus* and *E. avium*, 2 each of *E. casseliflavus* and *E. raffinosus*, lastly 1 *E. entomosocium*). From these 24 strains without any plasmid, 14 also lacked any known virulence genes, suggesting they are likely nonpathogenic.

Prophages in infant strains

Phages play a crucial role in bacterial evolution and are often associated with the presence of resistance genes and virulence factors. We identified prophages in nearly all *Enterococcus* isolates (97 of 100). Over time, prophages may lose some or all of their functionality, which is why phages were characterized as either intact, questionable (if intact), or incomplete (see Methods). Only seven genomes were found to carry all their prophages in an intact state, with six and one strains possessing three and four intact phages, respectively (Suppl. Table S6). Overall, intact prophages were found in 92% of the genomes. Interestingly, a minority of strains (~20%) carried 75% of the different intact prophages identified. Additionally, five genomes were found to harbor only partial prophages, despite having at least one prophage identified (Figure 1).

Phenotypic antibiotic resistance profiling of infant strains

All isolated *Enterococcus* strains were tested phenotypically against four important commonly used antibiotics, namely gentamicin, vancomycin, amoxicillin, and linezolid (Figure 2). These antibiotics are commonly prescribed in Europe by clinicians in neonatal care.[77–79] Pre-term infants often receive antibiotics due to the likely cause of infection, however full-terms also receive it as a prevention measure.[79] As a part of ESKAPE group and under WHO pathogen priority list, *Enterococcus faecium* makes vancomycin very relevant.[80] There have been VRE outbreaks in all over the Europe and we checked for VRE in our collection.[81] Among aminoglycosides, gentamicin is most widely used in neonatal care and high-level resistance is quite common in *Enterococcus*.[79] Resistance to β-lactams antibiotics is another major concern as these antibiotics are used either alone or in combination for enterococcal infections.[82] Lastly, there is an increasing emergence of LRE among children's infected with *Enterococcus*, which is then difficult to treat.[83,84] Hence, besides detection AMR genetically, we focused on phenotypic verification using the above mentioned.

Among the 21 MAMI isolates, 29% were resistant to gentamicin, while resistance in the 79 LucKi strains was observed for amoxicillin, vancomycin, and gentamicin at rates of about 2%, 11%, and 5%, respectively. All vancomycin-resistant isolates were from the LucKi cohort, and belonged to three different species (5 *E. gallinarum*, 2 *E. entomosocium*, and 1 *E. casseliflavus*) with MIC values between 8 μg/ml for most strains and reaching 16 μg/ml for two strains. However, no resistance genes were found *in silico* for *E. entomosocium* and *E. casseliflavus*, despite their phenotypic resistance when stringent parameters were used for minimum coverage and minimum identity (i.e., 95%). However, we found *vanC* genes for both species when loose parameters were used (Suppl. Table S8). Notably, all gentamicin-resistant strains were *E. faecalis*, with MIC values at or above 256 μg/ml. Additionally, E. *gilvus* and *E. faecium* each were resistant to amoxicillin, although no corresponding resistance genes were identified *in silico*. All strains were sensitive to linezolid , but no resistance genes were detected (Table 1).

Phenotypic virulence profiles of selected E. faecalis *on the nematode* Caenorhabditis elegans

The nematode *C. elegans* emerged as a versatile *in vivo*, yet less expensive and free-of-ethics model.[85] This nematode is used successfully for many years as model organism for bacterial pathogenesis,[86] e.g., studying the interaction of probiotic *Enterococcus* with pathogenic *E. coli*.[4] Also commensal microorganisms, such as *E. faecalis*, possess antibiotic-resistant genes and virulence factors that enhance their ability to induce pathogenicity in nematodes. Such determinants are linked to nosocomial epidemics and capable of causing infections,[87] which we sought to test in some of our strains using nematodes. As said, *E. faecalis* was the most prominent species, possessing a high number of virulence genes, plasmids, and prophages. We aimed to determine whether the mere presence of these genetic markers could predict the pathogenic potential of these strains. Of note, all infants from which the enterococci were isolated, were generally healthy. Thus, we selected ten different strains (five from each cohort), each encoding a greater variety of virulence factors, phages, and plasmids, and tested them using nematode killing assays (Table 2). In these assays, nematodes were fed with the selected strains until death was observed. One strain, *E. faecalis* LH-Sp-12, particularly shortened the lifespan for the nematodes, with a time-to-death for 50% of the population (TD50) of just 2.4 days. In contrast, nematodes fed with LH-NE -78 had a TD50 of 8.3 days, while the control worms fed with *E. coli* OP50 lived an average of 11.4 days. Thus, all selected *Enterococcus* strains killed nematodes more quickly than the normal bacterial feed, with strains from the MAMI cohort showing a slightly shorter mean TD50 (4.8 days) compared to the LucKi strains (5.9 days), though

Table 1. Resistant isolates with their MIC values. Strains not listed had no resistance.

Classification	Isolate No.	Antibiotic	MIC values [µg/ml]
E. faecalis	LH-Sp-1	Gentamicin	>512
E. faecalis	LH-Sp-26	Gentamicin	256
E. faecalis	LH-Sp-29	Gentamicin	256
E. faecalis	LH-Sp-38	Gentamicin	>128
E. faecalis	LH-Sp-42	Gentamicin	>256
E. faecalis	LH-Sp-50	Gentamicin	>128
E. faecalis	LH-NE-98	Gentamicin	>512
E. faecalis	LH-NE-114	Gentamicin	512
E. faecalis	LH-NE-168	Gentamicin	512
E. faecalis	LH-NE-192	Gentamicin	512
E. gilvus	LH-NE-22	Amoxicillin	>64
E. faecium	LH-NE-49	Amoxicillin	>64
E. casseliflavus	LH-NE-5	Vancomycin	8
E. entomosocium	LH-NE-39	Vancomycin	8
E. entomosocium	LH-NE-44	Vancomycin	8
E. gallinarum	LH-NE-42	Vancomycin	16
E. gallinarum	LH-NE-63	Vancomycin	8
E. gallinarum	LH-NE-136	Vancomycin	16
E. gallinarum	LH-NE-158	Vancomycin	8
E. gallinarum	LH-NE-164	Vancomycin	8

Table 2. TD50 with number of virulence genes, prophages, plasmids, and virulence factors corresponding to each of ten *E. faecalis* strains.

Strain	TD50 (days)	Prophages		Plasmids	Virulence factors	Total
		Intact	All			
LH-Sp-01	5.6	2	6	2	29	37
LH-Sp-09	6.3	3	5	2	24	31
LH-Sp-12	2.4	1	6	2	25	33
LH-Sp-33	5.8	3	4	2	24	30
LH-Sp-44	4.0	2	3	2	13	18
LH-NE-61	4.6	0	2	0	24	26
LH-NE-78	8.3	0	2	2	16	20
LH-NE-116	4.4	1	4	1	35	40
LH-NE-182	7.2	1	5	4	22	31
LH-NE-192	5.0	1	7	3	24	34

this difference was not statistically significant (p-value >0.05). Despite this general finding, we did not observe any significant correlation between the observed TD50 and the total number of virulence genes, plasmids, or prophages, nor any specific genetic determinant. A weak association was observed with the presence of some prophages (data not shown). Thus, the combination of virulence genes, plasmids, and prophages, contributing to pathogenicity and nematode killing remains to be defined.

Pangenome and metabolic estimation of E. faecalis

Since *E. faecalis* was the most prevalent infant-associated species, with 60 out of 100 strains identified, we performed a detailed genomic analysis for this species. We determined the pangenome including the type strain. In total, 5640 gene clusters were found across the 61 genomes, of which 2238 were considered 'core genes' and 3402 were 'accessory genes'. All genomes shared 1845 gene clusters, representing a sum of 112,545 genes (Figure 3).

We examined the metabolic capacity of the 60 *E. faecalis* strains in this study to gain a deeper understanding of their potential pathogenicity and impact on the human host. We concentrated on differences between strains from the two geographically distinct cohorts to assess regional influences. Overall, we found limited differences in metabolic pathways (related to heme biosynthesis, carbohydrate metabolism, and vitamin/cofactor metabolism; Figure 4). Notably, distinct variations emerged only in specific pathway modules, such as methionine degradation, D-galactonate degradation, glucuronate pathways, and the malonate semialdehyde pathway. Furthermore, certain signature modules, including tetracycline resistance-efflux pump Tet38, multidrug resistance-efflux pump AbcA and vancomycin resistance of the D-Ala-D-Lac type, varied between cohort strains. Despite these differences, most strains exhibited highly similar metabolic capacities, with only a few pathways differing, such as dermatan and chondroitin sulfate degradation absent in some strains.

Interestingly, while *Enterococcus* spp. typically encode a rich repertoire of ARGs, modules related to "drug resistance", "aromatics degradation" and "glycosaminoglycan metabolism" were either absent or displayed a low completeness score. This suggests that these genomes share only a limited KOs (KEGG Orthologues), which contribute to several pathways (Suppl. Table S9). Taken together, the metabolic profiles of *E. faecalis* from both cohorts share major metabolic pathways, including pathways for lysine, pyrimidine desoxyribonucleotide and coenzyme A biosynthesis, as well as glycolysis (Supplementary Figure S1). These pathways were not species-specific, underscoring the metabolic consistency of *E. faecalis* across different geographic regions.

Functional enrichment for genomes of E. faecalis

To further investigate potential differences between strains from the two cohorts, we analyzed the presence and absence of metabolic functions and pathways. We found a higher occurrence of different functions in LucKi strains, which correlates with the higher number of isolates. However, the number of enriched functions was conversely. Here we found 25 enriched functions in MAMI strains, while only 5 functions were enriched in LucKi strains (Figure 5). In summary, while we expected more pronounced differences between the strains from each cohort, our findings indicate only subtle variations in metabolic functions between all strains, suggesting that the geographic influence on strain-specific metabolism is relatively minor.

Examining *Enterococcus*' global diversity using public genomes

We were surprised by the relatively low regional variation observed among our cohort isolates. To explore this further, we expanded our analysis to identify broader geographic and temporal patterns. We obtained high-quality complete genome sequences of *Enterococcus* spp. from public datasets, focusing on complete genomes. While a large number of *Enterococcus* strains had been sequenced, only 587 genomes (representing 15 known species and 1

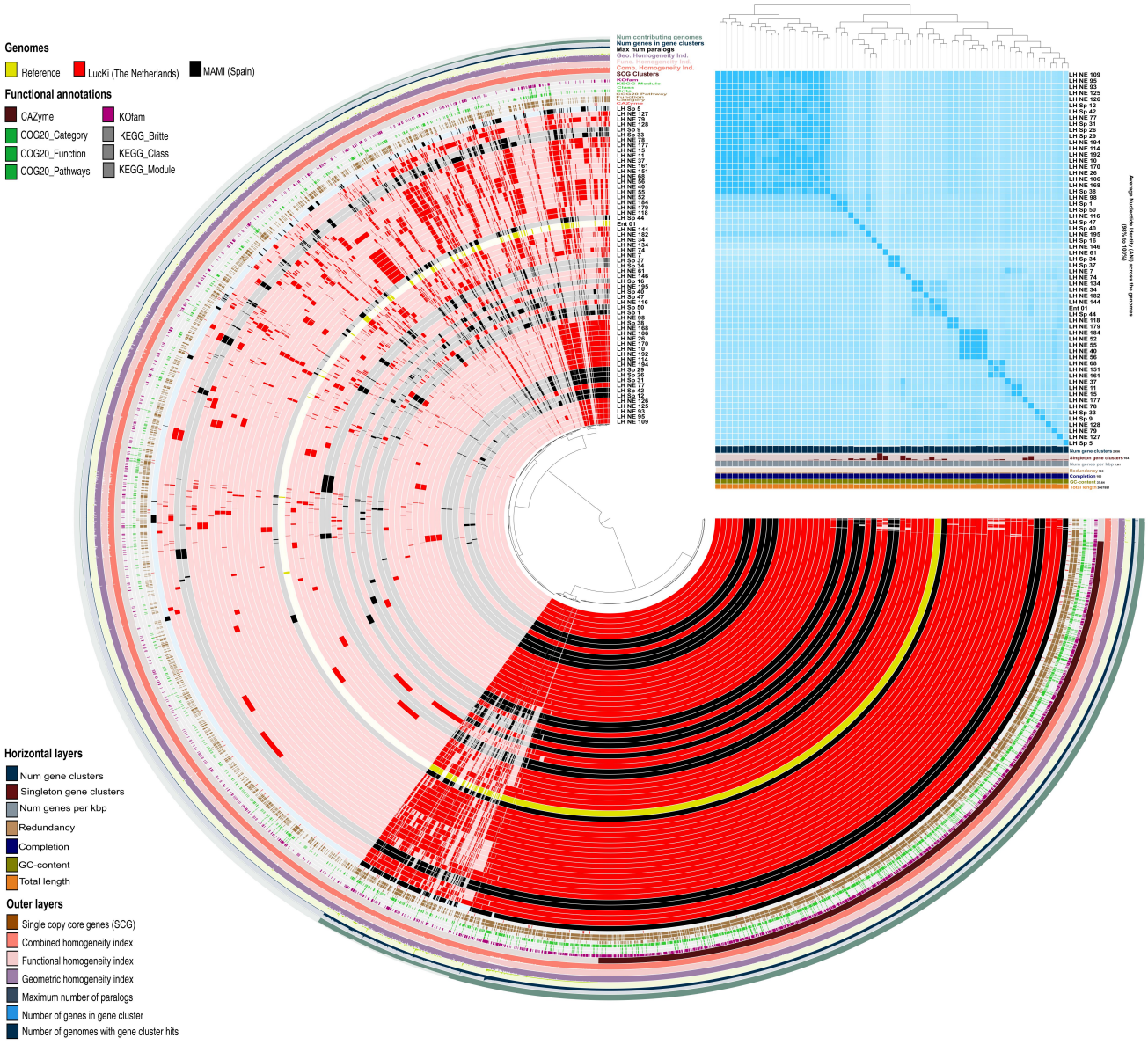

Figure 3. Pangenome of 61 genomes of *E. faecalis* (showing gene presence/absence) revealing 5,640 gene clusters of 3402 accessory and 2238 core genes in our isolates plus the type strain (*E. faecalis* DSM 20478, GCF_000392875.1). The blue heatmap with dendogram displayed above the pangenome presents the average nucleotide identity (ANI); below this are layers representing number of gene clusters, singletons, redundancy, GC content and total length. Outer rings represent core genes, total genes in gene cluster, combined, functional and geometric homogeneity index and functional annotations (CAZyme, COG20, KOfam and KEGG). Functional homogeneity indicates how conserved aligned amino acid residues across genes are. Geometric homogeneity compares the positions of gaps in the aligned residues without considering specific amino acids.

unidentified) met sufficient quality thresholds (see Methods). Unfortunately, for many of these genomes, detailed metadata (such as host type, disease status, and age of host) were lacking, with only a small number apparently isolated from healthy individuals. Most strains originated from North America, Asia, and Europe, with fewer from Australia, the Middle East, and South America – likely reflecting sampling and sequencing effort biases. Notably, 120 isolates lacked any geographic region data but were distributed throughout the phylogenetic tree without forming distinct clusters. We constructed a phylogenetic tree from the 587 public genomes and our 100 infant isolates. Seven species were common to both (Figure 6a). This expanded dataset was the analyzed for plasmids, phages, and virulence genes.

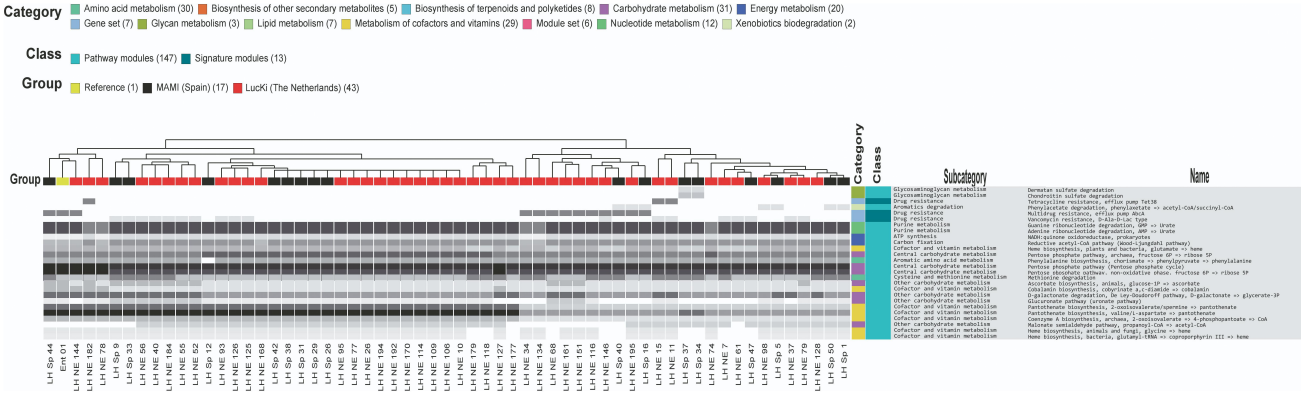

Figure 4. Heatmap indicating presence of functional modules along with their completeness concerning each biochemical pathway in the *E. faecalis* metabolism (excerpt, full heatmap in supplementary Figure S1). On top, the cohort is indicated: black, MAMI; red, LucKi, while the yellow labeled genome indicates the reference genome of *E. faecalis*. The class of the pathway is also indicated by shades of teal: lighter teal refers to 'pathway modules' (functional units of gene sets in metabolic pathways, including molecular complexes), while darker teal refers to 'signature modules' (functional units of gene sets that characterize phenotypic features). In the heatmap itself, darker shades indicate complete or near complete pathways, while lighter shades indicate a low percent pathway completeness. Thus, only a few genes are shared concerning a specific pathway found in the KEGG database (KO).

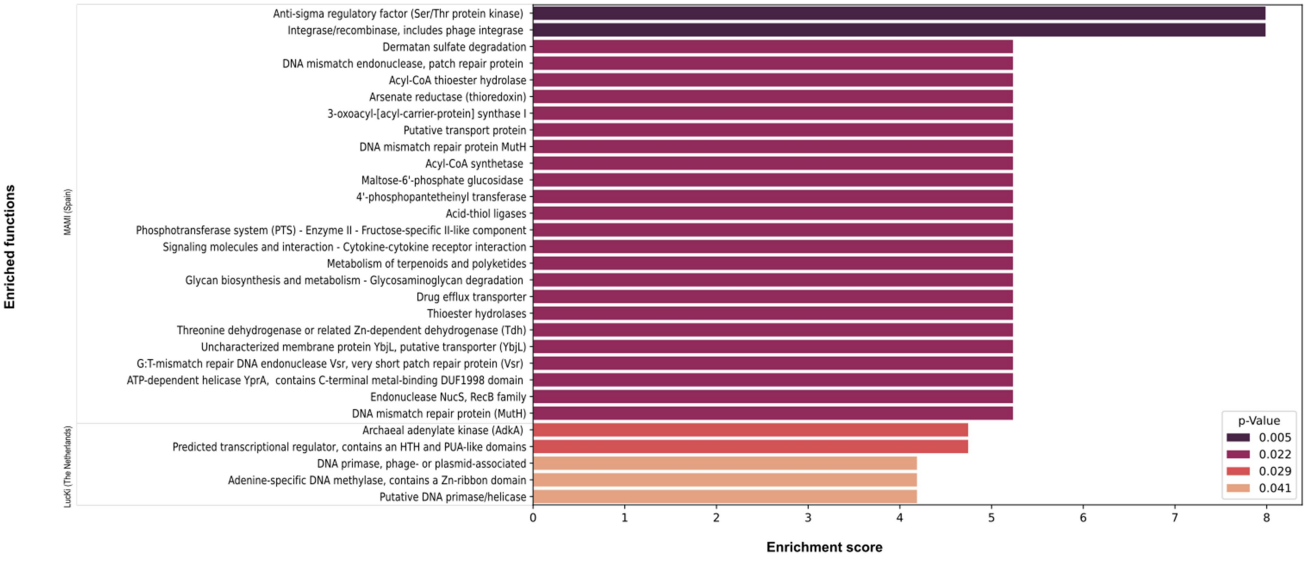

Figure 5. The bar plot illustrates the enriched functions identified by anvi-compute-functional-enrichment program using four different sources: cog20_function, KEGG_Module, KEGG_Britte, and KOfam. Each bar represents a specific function (terms of molecular function, cellular component and biological process) with the bar length indicating the enrichment score and the significance denoted by the color-coded p-value; <0.05 was considered to be significant.

Since plasmids play a crucial role in disseminating ARGs and virulence genes, we focused on their prevalence first. We identified 82 different plasmids across the 515 out of 587 public isolates, 28 of which were also common in our isolates. These plasmids were found in 11 species (including 1 unidentified spp.) from 22 hosts, with 6 species with plasmids shared between the public dataset and our cohort. The top ten plasmids, analyzed in a presence/absence matrix (Figure 6a, second inner ring), encoded ARGs conferring resistance to a range of antibiotics. Glycopeptide resistance was the most prevalent (28%), followed by aminoglycoside (23%) and tetracycline resistance (14%; Figure 6b) (Suppl. Table S10). Despite biases in regional sampling, Asia, Europe, and North America had the highest carriage of different ARGs, with the most common conferring resistance to glycopeptides, aminoglycosides, fluoroquinolone/macrolide/rifamycin (FMR), macrolide/

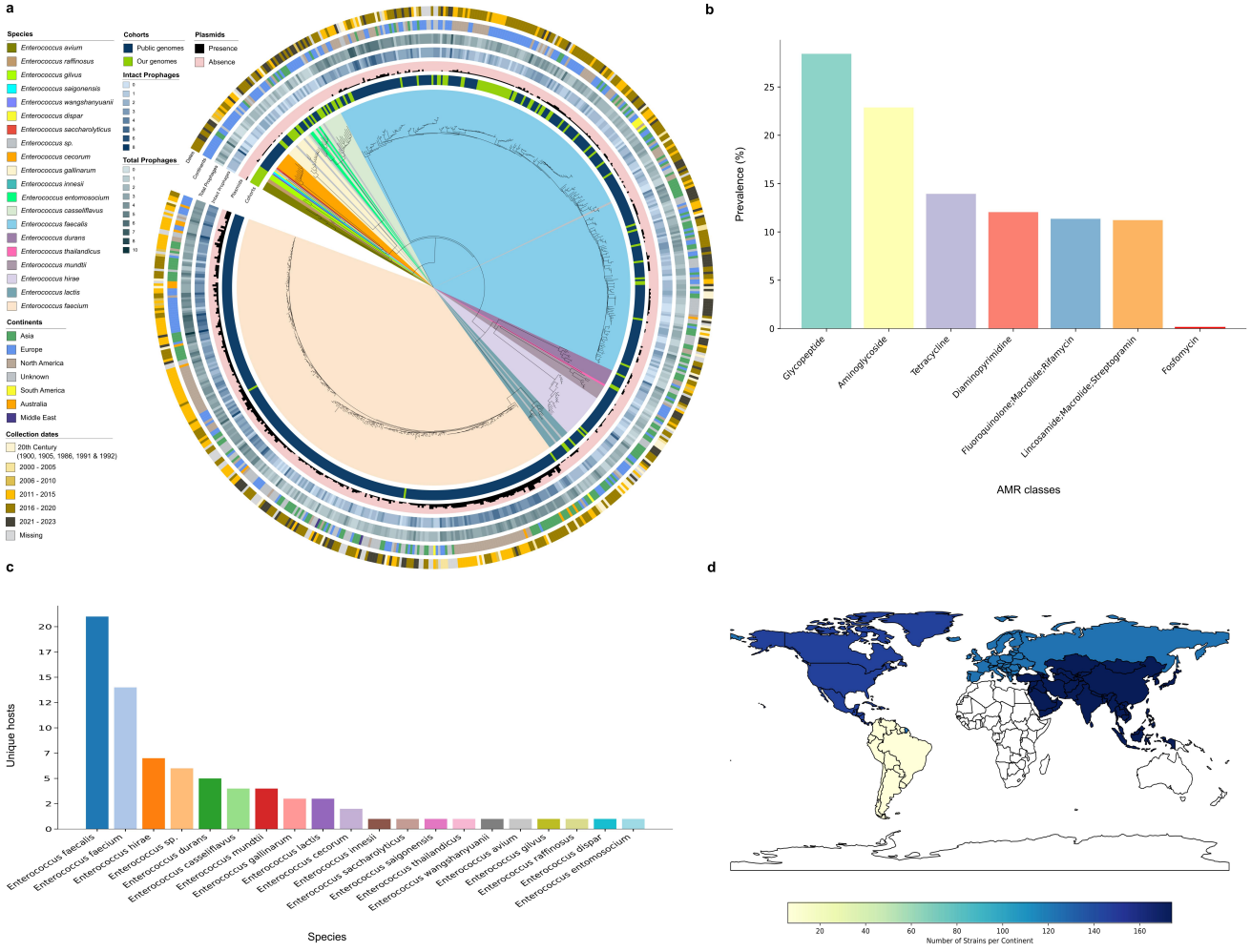

Figure 6. Global *Enterococcus* diversity and distribution (a) midpoint rooted tree of 687 *Enterococcus* genomes (587 public genomes and 100 genomes from this study). A total of 15 species are found in public genomes from over 30 countries, with the first isolate from 1900 until 2023. Ten highly prevalent plasmids (DOp1, pNb2354p1, pRE25, pRUM, pB82, pE1p13, pAmalpha1, pAD1, p200B, and pQY003) were selected and a presence/absence matrix is shown as inner circle next to cohorts (black, presence; pink, absence); the number of total and intact prophages (darker color, more prophages; next bar is of continents followed by dates of collection. (b) Resistance of all public *Enterococcus* from panel a sorted for prevalence. (c) Host distribution among different *Enterococcus* species (public and infants) with highest number of hosts occupied by *E. faecalis* and *E. faecium*. (d) A world map shows the prevalence of 587 public strains in each region.

lincosamide/streptomycin (MLS) and tetracycline. Among the most frequently detected genes were AAC(6')-li, *efmA*, *dfrE*, *IsaA*, *efrA*, *efrB*, *emeA*, *tetM*, *dfrF*, and *dfrG*. In both the public and our infant cohorts, *E. faecalis* was the most resistant species, harboring the highest number of ARGs.

Phages were found in 93.2% of isolates, with 6.8% containing more than five prophages. All 15 species from the public dataset contained prophages. A total of 127 isolates had intact prophages (ranging from 1 to 8 per genome), with those carrying more than 5 mostly isolated from humans. Interestingly, three isolates, namely two *E. mundtii* and one *E. casseliflavus*, lacked any genetic determinants of ARGs, virulence factors and prophages. In terms of virulence, three species, namely *E. faecalis, E. faecium*, and *E. lactis*, encoded the highest abundance of virulence genes across datasets (Suppl. Table S11). In contrast, *E. raffinosus*, *E. durans*, and another unidentified *Enterococcus* strain, had fewest virulence factors. We also found three different Epxs in two public *E. hirae* strains, namely Epx4 and Epx7 in genome GCF_002278015.2, which was isolated from a human, and Epx6 in GCF_016727265.1 from an unknown source. However, comparable to our *E. hirae* strains from the infants, they also do not carry any known virulence factors. In contrast,

ARGs, phage, and plasmids are present. Here, the latter public strain carries Epx6 on a large conjugative plasmid. Of the 41 virulence genes found in the public isolates, 39 were shared with our cohort enterococci. Only two genes (*esp* and *ecbA*) were unique to public strains, with the strains being isolated from 31 different hosts. This indicates that virulence factors in *Enterococcus* are not highly specific to particular hosts.

When examining host diversity, *Homo sapiens* had the highest number of isolates (294 strains), with *E. faecium*, and *E. faecalis* being the most prevalent (173 and 108 strains, respectively) (Figure 6c). While this reflects a sampling bias toward human infections, *E. faecalis* demonstrated the ability to thrive in a wide range of hosts, including mammals, fish, birds, and even marine algae. *E. faecium* was also versatile, found in 14 different host organisms, including insects and fish (see Figure 6c).

Discussion

We investigated *Enterococcus* spp. diversity from two European infant cohorts, MAMI from Spain and LucKi from the Netherlands, to understand what distinguishes these commensal enterococci across regions in terms of metabolic features, virulence, and antimicrobial resistance. While our study focused on healthy full term infants, very few received antibiotics for non-gut-related conditions (e.g., conjunctivitis, or within an intensive care unit). Despite this, we found these healthy infants carried a surprisingly large resistome and a wide variety of *Enterococcus* spp., which was similar across both cohorts. In order to understand the global picture, we performed a comparative analysis using public datasets. Our analysis revealed *Enterococcus* spp. are panmictic, meaning that they move freely across regions and niches, exchanging genetic material. Notably, LucKi cohort strains exhibited greater species diversity than MAMI strains, likely due to the larger sample size and more time points from LucKi. Definitely, age is involved while geographic region could also have contributed to these differences.

Even among antibiotics naïve infants, we detected substantial antibiotic resistant *Enterococcus* strains, particularly gentamicin-resistant strains, which were more common in MAMI than LucKi. Aminoglycosides, including gentamicin, are frequently prescribed for neonates, both in the Netherlands and Spain.[78] This could suggest that resistant strains may have been acquired from the community or hospital environments. Similarly, genetic determinants for resistance to the FMR group of antibiotics (i.e., fluoroquinolone, macrolide, and rifamycin), followed by diaminopyrimidine and LMS (lincosamide, macrolide, oxazolidinone, phenicol, pleuromutilin, streptogramin, and tetracycline), were highly prevalent, potentially spreading via plasmids or phages. Conversely, we identified only one amoxicillin resistant *E. faecium* in LucKi, likely due to point mutations in penicillin-binding proteins (PBPs) rather than ARG uptake.[88] Thus, even though amoxicillin is commonly prescribed in the Netherlands,[89] the selective pressure exerted by this antibiotic appears to differ to ARG-mediated resistance. Since only very few infants were given antibiotics at all, we could not analyze for significant differences in antibiotic resistance and virulence potential for our strains.

The interplay between ecological niches, genetic elements, and environmental factors has implications for gut microbiota. As the gut microbiome acquires ARGs from the wider microbial community, it may contribute to increased AMR, which can lead to treatment failures. This, in turn, facilitates the emergence of MDR strains, increasing the risk of nosocomial infections and posing a growing challenge for public health systems.[90] This risk is particularly high in neonates with immature microbiomes, often exposed to antibiotics, and even more so in those born by C-section. Indeed, studies have found high prevalence of *Enterococcus* in infants born by C-section compared to vaginally born infants.[91,92] Our study found five different *Enterococcus* species in nine C-section infants, including *E. gilvus*, which was uncommon in other samples. Moreover, strains from C-section infants encoded unique genetic determinants; however, given that we had such few samples, we cannot draw wider conclusions.

A key question in *Enterococcus* research is whether commensal strains can become pathogenic under certain conditions. Many of our strains encoded

a large array of virulence genes, some up to 36 genes. It is reported that the presence of the virulence genes *sprE* (serine protease) and *gelE* (gelatinase) increases the pathogenicity by increasing rate of killing of nematodes.[75] Commensal *Enterococcus* ssp. were found to use these factors toward colonization and proliferation in the gut system.[87] Thus, *C. elegans* has emerged as powerful model for virulence screenings of pathogenic bacteria.[86] Previously, it has been used with *E. faecalis*, *Streptococcus pneumoniae*, *Staphylococcus aureus*,[93] extra-intestinal pathogenic *Escherichia coli* (ExPEC), *Pseudomonas aeruginosa*,[94] *Streptococcus pyogenes*,[95] *Salmonella typhimurium*,[96] and others. Using *C. elegans* as a model, we tested the virulence of selected *E. faecalis* strains, of which some indeed encoded *sprE* and *gelE*. While one *E. faecalis* strain from the MAMI cohort exhibited significant pathogenicity, killing 50% of nematodes in just about 2.4 days, no clear correlation between the number or any specific virulence genes and pathogenicity was observed. This may link to our metabolic capacity analysis, which was highly similar across all genomes. It therefore remains unclear under which circumstances, environmental conditions and genetic conditions a strain behaves as a commensal, a protective organism, or as a pathogen.[35]

Overall, our findings suggest *Enterococcus* species and strains are highly mobile across Europe, sharing genetic and functional traits across wide geographic regions, given cohort sites were more than 1000 km beeline distant. This led us to expand our analysis by assembling high-quality genomes of *Enterococcus* species from publicly available databases. These datasets included strains collected over the past century, with a noticeable increase in samples from the last two decades. Unlike our isolates, which came from healthy infants, these public datasets included strains from clinical disease cases, offering a broader perspective on *Enterococcus* diversity. The public strain collection encompassed a range of species, each isolated at different time points, highlighting the increasing diversity of *Enterococcus* species over time. We found a significant number of isolates originating from North America and Europe, possibly reflecting a research bias toward these regions. Of all species examined, both *E. faecalis* and *E. faecium*, exhibited the widest host ranges, consistent with their well-documented ecological adaptability and versatility.[97] Interestingly, whilst *E. faecium* is classified as an ESKAPE pathogen, the majority of (nosocomial) infections still involve *E. faecalis*. This discrepancy may be somewhat arbitrary and highlights the need for more nuanced classifications of these species. Our findings indicate that *Enterococcus* spp. disseminate easily across different geographical regions, with their metabolic profiles, resistome, and host range often showing similarities across large distances and, in some cases, even across species-boundaries. This ease of transmission, especially for ARGs, has been shown in studies on tourists, who acquire and carry resistant strains back to their home countries.[98] It is therefore unsurprising that Europe has seen several large VRE outbreaks in recent years.[99–101] Comparing public *Enterococcus* genomes, we observed that ARGs, virulence factors and plasmids often co-occurred, while strains lacking ARGs tended to also lack virulence factors and plasmids. Despite this, some isolates with virulence factors still carried prophages, indicating that such other genetic elements may play a role in acquiring these traits. For our own isolates, a very weak correlation ($R^2 = 0.12$) was observed for plasmid carriage and presence of virulence factors. However, a moderate correlation ($R^2 = 0.36$) was found for the presence DOp1 plasmid, which is the most abundant in our study, and possession of virulence determinants. Interestingly, despite the absence of known virulence factors in some of the analyzed genomes, the presence of the recently described *epx* genes was a notable finding. Its toxin-mediated virulence was demonstrated using a toxin-carrying *E. faecium* strain. This strain induces death of peripheral blood mononuclear cells and damages intestinal organoids during co-culture. The authors who had discovered the novel family of Epx toxins emphasized their widespread distribution. The host diversity, i.e. binding of MHC-I in different hosts, suggests that the acquisition of such toxins is not a rare event and probably confers competitive advantages.[33]

Previously, a rich repertoire of ARGs and virulence factors was found in *Enterococcus* ssp. isolated from wastewater and associated waters in South Africa.[15] The authors found aminoglycosides genes in *E. faecalis*, *E. hirae* and *E. durans*. Similarly, we detected various aminoglycoside-

resistance genes in some *E. faecalis* strains, a few genes in *E. faecium, E. casseliflavus, E. durans, E. gilvus*, and *E. lactis*. In the mentioned study from South Africa, *E. faecium* strains were enriched with tetracycline, erythromycin, and tetracycline genes, while we found tetracycline genes in *E. faecium,* but erythromycin resistance in *E. lactis* strains. While such findings cause quite negative attention, *Enterococcus* ssp. seem to be two-faced. Not only pathogens exist, but many harmless commensals and, as already mentioned, probiotics. Probiotic strains for *E. faecalis* such as Symbioflor® 1 and SI-FC-01 showed positive properties when tested in *C. elegans* by down-regulating virulence genes in enterohemorrhagic *E. coli* O157:H7 (EHEC)[4] or promoting the healthspan and neuroprotection, respectively.[102] Thus, Tadesse *et al.*[103] emphasized that, besides the potential pathogenic strains, they found 19 *Enterococcus* strains among 44 environmental isolates, which appear to be safe and could be utilized in food fermentations. These strains seem to lack genes encoding antibiotic resistance and important virulence factors. Therefore, the positive potential of *Enterococcus* strains should not be dismissed, but rather exploited.

On closing, the fact that *Enterococcus* has even been found as a prominent microbial inhabitant on the International Space Station[104] highlights its ability to thrive in diverse and extreme environments. The 'One Health' approach, which includes zoonotic transmission,[105] is increasingly relevant, as MDR enterococci are highly prevalent in non-hospital and non-human environments.[97] Human activities, including travel, agriculture, and global trade, have turned the planet into a 'mixing vessel' for *Enterococcus* strains, posing a potential threat to global public health.[106] Future work should expand on our findings by longitudinally studying larger, more diverse cohorts, combining strain isolation with metagenomics to better understand early life *Enterococcus* acquisition and persistence, particularly in relation to ARGs and AMR. In addition, the expression of Epxs should be confirmed phenotypically. Here, it would be interesting to understand their recognition toward MHC-1 of different hosts, especially for those yet not tested in, e.g., cell lines. Finally, investigating whether probiotic *Enterococcus* strains could mitigate acquisition of potential pathogenic *Enterococcus* strains could offer new strategies for improving infant health.

Conclusion

Our comprehensive, comparative analysis of *Enterococcus*, particularly within infants, has revealed a complex landscape of diversity, resistance, and mobility across geographic regions. *Enterococcus* strains, while often harmless commensals under steady-state conditions, may become opportunistic pathogens, particularly when the wider gut microbiota is perturbed (e.g., by antibiotics). Understanding the diversity and ARG profiles of *Enterococcus* globally is essential for developing effective strategies to combat infections. Given the high mobility of these bacteria, a global effort is needed to track colonization, monitor resistance, and implement preventive measures, including in healthy individuals. The development of potent probiotics that safely exclude virulent strains may offer a promising approach to mitigate the growing threat of *Enterococcus*-related MDR infections, especially in vulnerable populations like infants. Further research and global data-sharing are critical for addressing the increasing challenges posed by these resilient bacteria.[35]

Acknowledgments

We would like to thank the QIB sequencing team for the whole-genome sequencing data. We would like to thank John Penders for providing us with a subset of samples from the LucKi cohort.

M.H.S. isolated the *Enterococcus* strains, performed phenotypic assays, genome analysis, and wrote the paper draft. M.B and M.C.C. were lead of the MAMI cohort and provided infant fecal samples. M.M. was lead of the LucKi cohort and N.V.B. provided infant fecal samples. A.A.G. helped with sequencing. M.K. helped with data preprocessing and assembly. K.E. conducted nematode assays. L.J.H. supervised and conceived study. K.N. supervised study, provided material, and finalized draft. All authors read and approved the final draft.

Author contributions

CRediT: **Muhammad Hassan Saeed:** Conceptualization, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft; **Magdalena Kujawska:** Data curation, Methodology, Writing – review &

editing; **Kristiana Ellen:** Investigation, Writing – review & editing; **Antia Acuna-Gonzalez:** Methodology, Writing – review & editing; **Manuel Bernabeu Lorenzo:** Investigation, Project administration, Resources, Writing – review & editing; **Maria Carmen Collado:** Funding acquisition, Investigation, Resources, Writing – review & editing; **Monique Mommers:** Investigation, Writing – review & editing; **Niels van Best:** Investigation, Resources, Writing – review & editing; **Lindsay J. Hall:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing; **Klaus Neuhaus:** Conceptualization, Project administration, Resources, Supervision, Writing – review & editing.

Data availability

Publicly available genomes were downloaded from NCBI (19 Oct 2023; Suppl. Table S4). All genomic data from our isolates are submitted in GenBank under JBDKAF000000000 to JBDKEA000000000 [https://www.ncbi.nlm.nih.gov/nuccore/JBDKAF000000000. to https://www.ncbi.nlm.nih.gov/nuccore/JBDKEA000000000.] within BioProject PRJNA1110724 [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1110724]. All isolates were submitted to the culture collection Weihenstephan Strain Collection (WS; Freising, Germany; CCInfo #1163 [https://ccinfo.wdcm.org/details?regnum=1163]). Strain numbers are WS 5642 to WS 5738 (Suppl. Table S6).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

M.H.S. is supported by Higher Education Commission, Pakistan (HEC), and German Academic Exchange Service (DAAD) scholarship (K.N. as supervisor). L.J.H. is supported by a Wellcome Investigator Award [No. 220876/Z/20/Z}, and a BBSRC Institute Strategic Programme Food, Microbiome and Health BB/X011054/1 and its constituent project BBS/E/QU/230001B. M.C.C. would like to acknowledge the support from H2020-ERC Starting Grant (MAMI, ref. 639226), Spanish Ministry of Science and Innovation (MCIN) research grant [MAMI plus, ref. PID2022-139475OB-I00] and from Generalitat Valenciana PROMETEO grant [NEOHEALTH ref. 2020/012, MICROGLOCAL ref. CIPROM/2023/3]). M.B. is grateful to have received a post-doctoral "Juan de la Cierva" grant [ref. JDC2022- 050269-I] MCIU/AEI and European Union "Next Generation EU"/PRTR grant. M.C.C. and M.B. would also like to acknowledge the award of the Spanish Government MCIN/AEI to the IATA-CSIC as Centre of Excellence Severo Ochoa [CEX2021-001189-S MCIN/AEI/10.13039/501100011033].

ORCID

Muhammad Hassan Saeed http://orcid.org/0000-0001-5188-4668
Magdalena Kujawska http://orcid.org/0000-0002-8842-3512
Kristiana Ellen http://orcid.org/0009-0005-5497-0277
Antia Acuna-Gonzalez http://orcid.org/0000-0002-8062-925X
Manuel Bernabeu Lorenzo http://orcid.org/0000-0002-3011-5373
Maria Carmen Collado http://orcid.org/0000-0002-6204-4864
Monique Mommers http://orcid.org/0000-0002-6341-3596
Niels van Best http://orcid.org/0000-0003-3577-7285
Lindsay J. Hall http://orcid.org/0000-0001-8938-5709
Klaus Neuhaus http://orcid.org/0000-0002-6020-2814

References

1. García-Solache M, Rice LB. The enterococcus: a model of adaptability to its environment. Clin Microbiol Rev. 2019;32(2). doi: 10.1128/CMR.00058-18.
2. Boehm AB, Sassoubre LM. Enterococci as indicators of environmental fecal contamination. Enterococci: From Commensals to Leading Causes of Drug Resistant Infection. Massachusetts Eye and Ear Infirmary. Boston; 2014. https://www.ncbi.nlm.nih.gov/books/NBK190421/
3. Dubin K, and Pamer EG, Britton RA, Cani PD. Enterococci and their interactions with the intestinal microbiome. Microbiol Spectr. 2014;5(6). doi: 10.1128/microbiolspec.BAD-0014-2016.
4. Neuhaus K, Lamparter MC, Zölch B, Landstorfer R, Simon S, Spanier B, Ehrmann MA, Vogel RF. Probiotic enterococcus faecalis Symbioflor® down regulates virulence genes of EHEC in vitro and decrease pathogenicity in a Caenorhabditis elegans model. Arch Microbiol. 2017;199(2):203–213. doi: 10.1007/s00203-016-1291-8.
5. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. Internat J Systemat Evolut Microbiol. 2020;70(11):5607–5612. doi: 10.1099/ijsem.0.004332.
6. Zaheer R, Cook SR, Barbieri R, Goji N, Cameron A, Petkau A, Polo RO, Tymensen L, Stamm C, Song J. Surveillance of enterococcus spp. reveals distinct species and antimicrobial resistance diversity across a one-health continuum. Sci Rep. 2020;10(1):3937. doi: 10.1038/s41598-020-61002-5.
7. Zaidi S-E-Z, Zaheer R, Poulin-Laprade D, Scott A, Rehman MA, Diarra M, Topp E, Domselaar GV, Zovoilis A, McAllister TA. Comparative genomic analysis of enterococci across sectors of the one health

continuum. Microorganisms. 2023;11(3):727. doi: 10.3390/microorganisms11030727.
8. Zaidi S, Zaheer R, Zovoilis A, McAllister T. Enterococci as a one health indicator of antimicrobial resistance. Can J Microbiol. 2024;70(8):303–335. doi: 10.1139/cjm-2024-0024.
9. Arias CA, Murray BE. The rise of the enterococcus: beyond vancomycin resistance. Nat Rev Microbiol. 2012;10(4):266–278. doi: 10.1038/nrmicro2761.
10. Brooks B, Olm MR, Firek BA, Baker R, Thomas BC, Morowitz MJ, Banfield JF. Strain-resolved analysis of hospital rooms and infants reveals overlap between the human and room microbiome. Nat Commun. 2017;8(1):1814. doi: 10.1038/s41467-017-02018-w.
11. Kajihara T, Nakamura S, Iwanaga N, Oshima K, Takazono T, Miyazaki T, Izumikawa K, Yanagihara K, Kohno N, Kohno S. Clinical characteristics and risk factors of enterococcal infections in Nagasaki, Japan: a retrospective study. BMC Infect Dis. 2015;15(1):426. doi: 10.1186/s12879-015-1175-6.
12. Luo X, Li L, Xuan J, Zeng Z, Zhao H, Cai S, Huang Q, Guo X, Chen Z. Risk factors for enterococcal intra-abdominal infections and outcomes in intensive care unit patients. Surg Infect (Larchmt). 2021;22(8):845–853. doi: 10.1089/sur.2020.417.
13. Hochstedler-Kramer BR, Ene A, Putonti C, Wolfe AJ. Comparative genomic analysis of clinical enterococcus faecalis distinguishes strains isolated from the bladder. BMC Genom. 2023;24(1):752. doi: 10.1186/s12864-023-09818-z.
14. Kwit R, Zając M, Śmiałowska-Węglińska A, Skarżyńska M, Bomba A, Lalak A, Skrzypiec E, Wojdat D, Koza W, Mikos-Wojewoda E, et al. Prevalence of enterococcus spp. and the whole-genome characteristics of enterococcus faecium and enterococcus faecalis strains isolated from free-living birds in Poland. Pathogens. 2023;12(6):836. doi: 10.3390/pathogens12060836.
15. Mbanga J, Amoako DG, Abia ALK, Allam M, Ismail A, Essack SY. Genomic analysis of enterococcus spp isolated from a wastewater treatment plant and its associated waters in Umgungundlovu District, South Africa. Front Microbiol. 2021;12:2021. doi: 10.3389/fmicb.2021.648454.
16. Giraffa G. Enterococci from foods. FEMS Microbiol Rev. 2002;26(2):163–171. doi: 10.1111/j.1574-6976.2002.tb00608.x.
17. Yuksel FN, Akcelik N, Akcelik M. Incidence of antibiotic resistance and virulence determinants in enterococcus faecium and enterococcus faecalis strains, isolated from traditional cheeses in Turkey. Mol Genet Microbiol Virol. 2015;30(4):206–215. doi: 10.3103/S089141681504014X.
18. Sanderson H, Gray KL, Manuele A, Maguire F, Khan A, Liu C, Navanekere Rudrappa C, Nash JHE, Robertson J, Bessonov K, et al. Exploring the mobilome and resistome of enterococcus faecium in a one health context across two continents. Microb Genom. 2022;8(9). doi: 10.1099/mgen.0.000880.
19. Strateva T, Atanasova D, Savov E, Petrova G, Mitov I. Incidence of virulence determinants in clinical enterococcus faecalis and enterococcus faecium isolates collected in Bulgaria. Brazilian J Infectious Dis. 2016;20(2):127–133. doi: 10.1016/j.bjid.2015.11.011.
20. Wigmore SM, Greenhill AR, Bean DC. Isolation and characterization of enterococci from poultry reveals high incidence of enterococcus thailandicus in Victoria, Australia. J Appl Microbiol. 2024;135(8). doi: 10.1093/jambio/lxae194.
21. Souillard R, Laurentie J, Kempf I, Le Caër V, Le Bouquin S, Serror P, Allain V. Increasing incidence of enterococcus-associated diseases in poultry in France over the past 15 years. Vet Microbiol. 2022;269:109426. doi: 10.1016/j.vetmic.2022.109426.
22. Schwartzman JA, Lebreton F, Salamzade R, Shea T, Martin MJ, Schaufler K, Urhan A, Abeel T, Camargo I, Sgardioli BF, et al. Global diversity of enterococci and description of 18 previously unknown species. Proc Natl Acad Sci USA. 2024;121(10):e2310852121. doi: 10.1073/pnas.2310852121.
23. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. Nat Med. 2016;22(7):713–722. doi: 10.1038/nm.4142.
24. Lugli GA, Mancabelli L, Milani C, Fontana F, Tarracchini C, Alessandri G, van Sinderen D, Turroni F, Ventura M. Comprehensive insights from composition to functional microbe-based biodiversity of the infant human gut microbiota. npj Biofilms Microbiomes. 2023;9(1):25. doi: 10.1038/s41522-023-00392-6.
25. Kessler C, Hou J, Neo O, Buckner MMC. In situ, in vivo, and in vitro approaches for studying AMR plasmid conjugation in the gut microbiome. FEMS Microbiol Rev. 2022;47(1). doi: 10.1093/femsre/fuac044.
26. Saeed MH, Hall LJ. Early-life antibiotic usage and impact on the gut microbiota, including emergence of antimicrobial resistant enterococcus. Microbiota And Host. 2023;1(1):e230002. doi: 10.1530/MAH-23-0002.
27. Zaghloul HAH, El Halfawy NM. Genomic insights into antibiotic-resistance and virulence genes of enterococcus faecium strains from the gut of Apis mellifera. Microbial Genom. 2022;8(11). doi: 10.1099/mgen.0.000896.
28. Zhang K, Jin M, Yang D, Shen Z, Liu W, Yin J, Yang Z, Wang H, Shi D, Yang J, et al. Antibiotic resistance genes in gut of breast-fed neonates born by caesarean section originate from breast milk and hospital ward air. BMC Microbiol. 2022;22(1):36. doi: 10.1186/s12866-022-02447-8.
29. Bargheet A, Klingenberg C, Esaiassen E, Hjerde E, Cavanagh JP, Bengtsson-Palme J, Pettersen VK. Development of early life gut resistome and mobilome across gestational ages and microbiota-modifying

treatments. EBioMedicine. 2023;92:104613. doi: 10.1016/j.ebiom.2023.104613.
30. Castledine M, Buckling A. Critically evaluating the relative importance of phage in shaping microbial community composition. Trends Microbiol. 2024;32(10):957–969. doi: 10.1016/j.tim.2024.02.014.
31. He W, Russel J, Klincke F, Nesme J, Sørensen SJ. The role of plasmids in the gut microbiome during the first year of life. 2023. bioRxiv. 2023.2004.2005.535656.
32. Kitamoto S, Nagao-Kitamoto H, Kuffa P, Kamada N. Regulation of virulence: the rise and fall of gastrointestinal pathogens. J Gastroenterol. 2016;51(3):195–205. doi: 10.1007/s00535-015-1141-5.
33. Xiong X, Tian S, Yang P, Lebreton F, Bao H, Sheng K, Yin L, Chen P, Zhang J, Qi W, et al. Emerging enterococcus pore-forming toxins with MHC/HLA-I as receptors. Cell. 2022;185(7):1157–1171.e22. doi: 10.1016/j.cell.2022.02.002.
34. Backes C, Martinez-Martinez D, Cabreiro F. C. elegans: a biosensor for host–microbe interactions. Lab Anim (NY). 2021;50(5):127–135. doi: 10.1038/s41684-021-00724-z.
35. Lengfelder I, Sava IG, Hansen JJ, Kleigrewe K, Herzog J, Neuhaus K, Hofmann T, Sartor RB, Haller D. Complex bacterial consortia reprogram the colitogenic activity of enterococcus faecalis in a gnotobiotic mouse model of chronic, immune-mediated colitis. Front Immunol. 2019;10:1420. doi: 10.3389/fimmu.2019.01420.
36. García-Mantrana I, Alcántara C, Selma-Royo M, Boix-Amorós A, Dzidic M, Gimeno-Alcañiz J, Úbeda-Sansano I, Sorribes-Monrabal I, Escuriet R, Gil-Raga F, et al. MAMI: a birth cohort focused on maternal-infant microbiota during early life. BMC Pediatr. 2019;19:140.
37. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol. 1991;173(2):697–703. doi: 10.1128/jb.173.2.697-703.1991.
38. Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, List M, Neuhaus K, Tringe SG. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. mSphere. 2021;6(1). doi: 10.1128/mSphere.01202-20.
39. Kiu R, Shaw AG, Sim K, Acuna-Gonzalez A, Price CA, Bedwell H, Dreger SA, Fowler WJ, Cornwell E, Pickard D, et al. Particular genomic and virulence traits associated with preterm infant-derived toxigenic clostridium perfringens strains. Nat Microbiol. 2023;8(6):1160–1175. doi: 10.1038/s41564-023-01385-z.
40. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. iMeta. 2023;2(2):e107. doi: 10.1002/imt2.107.
41. Wick RR, Judd LM, Gorrie CL, Holt KE, Phillippy AM. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol. 2017;13(6):e1005595. doi: 10.1371/journal.pcbi.1005595.
42. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043–1055. doi: 10.1101/gr.186072.114.
43. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol. 2017;35:725–731.
44. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH, Borgwardt K. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformat. 2022;38(23):5315–5316. doi: 10.1093/bioinformatics/btac672.
45. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformat. 2014;30(14):2068–2069. doi: 10.1093/bioinformatics/btu153.
46. Kiu R. Sequence-stats: generate sequence statistics from FASTA and FASTQ files. GitHub; 2020. https://github.com/raymondkiu/sequence-stats.
47. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016;44(W1):W16–21. doi: 10.1093/nar/gkw387.
48. Wishart DS, Han S, Saha S, Oler E, Peters H, Grant J, Stothard P, P. and Gautam V. PHASTEST: faster than PHASTER, better than PHAST. Nucleic Acids Res. 2023;51(W1):W443–W450. doi: 10.1093/nar/gkad382.
49. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. Nucleic Acids Res. 2011;39(suppl):W347–W352. doi: 10.1093/nar/gkr485.
50. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9(1):5114. doi: 10.1038/s41467-018-07641-9.
51. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016;44(14):6614–6624. doi: 10.1093/nar/gkw569.
52. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, et al. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol. 2012;50(4):1355–1361. doi: 10.1128/JCM.06094-11.
53. Katz LS, Griswold T, Morrison SS, Caravas JA, Zhang S, den Bakker HC, Deng X, Carleton HA. Mashtree: a rapid comparison of whole genome sequence files. J Open Source Softw. 2019;4(44):1762. doi: 10.21105/joss.01762.

54. Seeman T. ABRicate. Github. 2020. https://github.com/tseemann/abricate .
55. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AV, Cheng AA, Liu S. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 2020;48:D517–d525. doi: 10.1093/nar/gkz935.
56. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895–3903. doi: 10.1128/AAC.02412-14.
57. Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. Nucleic Acids Res. 2016;44(D1):D694–D697. doi: 10.1093/nar/gkv1239.
58. Blin K. Ncbi-genome-download (0.3.3). Zenodo. 2023. doi: 10.5281/zenodo.8192486.
59. O'Leary NA, Cox E, Holmes JB, Anderson WR, Hem V, Tsuchiya MTN, Schuler GD, Zhang X, Torcivia J, Ketter A, et al. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI datasets. Sci Data. [2024 Jul 5]. 11(1):732. doi: 10.1038/s41597-024-03571-y.
60. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, et al. Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol. 2021;6(1):3–6. doi: 10.1038/s41564-020-00834-3.
61. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformat. 2009;10(1):421. doi: 10.1186/1471-2105-10-421.
62. Delmont TO, Eren AM. Linking pangenomes and metagenomes: the prochlorococcus metapangenome. PeerJ. 2018;6:e4320. doi: 10.7717/peerj.4320.
63. Dongen SV. Graph clustering via a discrete uncoupling process. SIAM J Matrix Anal Appl. 2008;30:121–141. doi: 10.1137/040608635.
64. Buck M, Mehrshad M, Bertilsson S. mOtupan: a robust bayesian approach to leverage metagenome-assembled genomes for core-genome estimation. NAR Genomics Bioinf. 2022;4(3). doi: 10.1093/nargab/lqac060.
65. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2016;45(D1):D353–D361. doi: 10.1093/nar/gkw1092.
66. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2013;42(D1):D199–D205. doi: 10.1093/nar/gkt1076.
67. Veseli I, Chen YT, Schechter MS, Vanni C, Fogarty EC, Watson AR, Jabri B, Blekhman R, Willis AD, Yu MK, et al. Microbes with higher metabolic independence are enriched in human gut microbiomes under stress. eLife. d oi: 1 0.7554/eLife. 2025;12 12:RP89862. doi: 10.7554/eLife.89862.3.
68. Shaiber A, Willis AD, Delmont TO, Roux S, Chen L-X, Schmid AC, Yousef M, Watson AR, Lolans K, Esen ÖC, et al. Functional and genetic markers of niche partitioning among enigmatic members of the human oral microbiome. Genome Biol. 2020;21(1):292. doi: 10.1186/s13059-020-02195-w.
69. Wiegand I, Hilpert K, Hancock RE. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. Nat Protoc. 2008;3(2):163–175. doi: 10.1038/nprot.2007.521.
70. Leclercq R, Cantón R, Brown DF, Giske CG, Heisig P, MacGowan AP, Mouton JW, Nordmann P, Rodloff AC, Rossolini GM, et al. EUCAST expert rules in antimicrobial susceptibility testing. Clin Microbiol and Infect : the Off Publ of the Eur Soc of Clin Microbiol and Infect Dis. 2013;19(2):141–160. doi: 10.1111/j.1469-0691.2011.03703.x.
71. Ritz C, Baty F, Streibig JC, Gerhard D. Dose-response analysis using R. PLoS One. 2015;10(12):e0146021. doi: 10.1371/journal.pone.0146021.
72. Lagkouvardos I, Fischer S, Kumar N, Clavel T. Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. PeerJ. 2017;5:e2836. doi: 10.7717/peerj.2836.
73. de Korte-de Boer D, Mommers M, Creemers HMH, Dompeling E, Feron FJM, Gielkens-Sijstermans CML, Jaminon M, Mujakovic S, van Schayck OCP, Thijs C, et al. LucKi birth cohort study: rationale and design. BMC Publ Health. 2015;15(1):934. doi: 10.1186/s12889-015-2255-7.
74. Moosavian M, Ghadri H, Samli Z. Molecular detection of vanA and vanB genes among vancomycin-resistant enterococci in ICU-hospitalized patients in Ahvaz in southwest of Iran. Infect Drug Resist. 2018;11:2269–2275. doi: 10.2147/IDR.S177886.
75. Sifri CD, Mylonakis E, Singh KV, Qin X, Garsin DA, Murray BE, Ausubel FM, Calderwood SB. Virulence effect of enterococcus faecalis protease genes and the quorum-sensing locus fsr in Caenorhabditis elegans and mice. Infect Immun. 2002;70(10):5647–5650. doi: 10.1128/IAI.70.10.5647-5650.2002.
76. Matle I, Atanda AC, Pierneef R, Magwedere K, Mafuna T. Resistome, mobilome, virulome analysis and phylogenomics of enterococcus faecalis isolated from raw muscle foods of beef origin in Gauteng, South Africa. Genomics. 2023;115(6):110742. doi: 10.1016/j.ygeno.2023.110742.
77. Leroux S, Zhao W, Bétrémieux P, Pladys P, Saliba E, Jacqz-Aigrain E. Therapeutic guidelines for prescribing antibiotics in neonates should be evidence-based: a French national survey. Arch Dis Child. 2015;100(4):394–398. doi: 10.1136/archdischild-2014-306873.
78. Prusakov P, Goff DA, Wozniak PS, Cassim A, Scipion CEA, Urzúa S, Ronchi A, Zeng L, Ladipo-

Ajayi O, Aviles-Otero N, et al. A global point prevalence survey of antimicrobial use in neonatal intensive care units: the no-more-antibiotics and resistance (NO-MAS-R) study. EClinicalMedicine. 2021;32:100727. doi: 10.1016/j.eclinm.2021.100727.
79. Rivera-Chaparro ND, Cohen-Wolkowiez M, Greenberg RG. Dosing antibiotics in neonates: review of the pharmacokinetic data. Future Microbiol. 2017;12 (11):1001–1016. doi: 10.2217/fmb-2017-0058.
80. WHO Team. WHO bacterial priority pathogens list, 2024: bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. Geneva: World Health Organization; 2024.
81. Werner G, Neumann B, Weber RE, Kresken M, Wendt C, Bender JK, Becker K, Borgmann S, Diefenbach A, Hamprecht A, et al. Thirty years of VRE in Germany – "expect the unexpected": the view from the National reference centre for staphylococci and enterococci. Drug Resist Updates. 2020;53:100732. doi: 10.1016/j.drup.2020.100732.
82. Nasso C, Scarfone A, Pirrotta I, Rottura M, Giorgi DA, Pallio G, Irrera N, Squadrito V, Squadrito F, Irrera P, et al. Appropriateness of antibiotic prescribing in hospitalized children: A focus on the real-world scenario of the different paediatric subspecialties. Front Pharmacol. 2022;13:890398. doi: 10.3389/fphar.2022.890398.
83. Rodríguez-Noriega E, Hernández-Morfin N, Garza-Gonzalez E, Bocanegra-Ibarias P, Flores-Treviño S, Esparza-Ahumada S, González-Díaz E, Pérez-Gómez HR, Mendoza-Mujica C, León-Garnica G, et al. Risk factors and outcome associated with the acquisition of linezolid-resistant enterococcus faecalis. J Glob Antimicrob Resist. 2020;21:405–409. doi: 10.1016/j.jgar.2020.01.010.
84. Wada Y, Afolabi HA, Al-Mhanna SB, Bello KE, Irekeola AA, Wada M, Ahmed N, Harun A, Yean CY, Mohamad Nasir NS, et al. Global occurrence of linezolid-resistant enterococcus (LRE): the first systematic review and meta-analysis. The Microbe. 2024;2:100041. doi: 10.1016/j.microb.2024.100041.
85. Issi L, Rioux M, Rao R. The nematode Caenorhabditis elegans - a versatile in vivo model to study host-microbe interactions. JoVE. 2017; e56487. doi: 10.3791/56487.
86. Hilbi H, Weber SS, Ragaz C, Nyfeler Y, Urwyler S. Environmental predators as models for bacterial pathogenesis. Environ Microbiol. 2007;9(3):563–575. doi: 10.1111/j.1462-2920.2007.01238.x.
87. Dey P. Good girl goes bad: understanding how gut commensals cause disease. Microb Pathog. 2024;190:106617. doi: 10.1016/j.micpath.2024.106617.
88. Sethuvel DPM, Bakthavatchalam YD, Karthik M, Irulappan M, Shrivastava R, Periasamy H, Veeraraghavan B. β-Lactam resistance in ESKAPE pathogens mediated through modifications in penicillin-binding proteins: an overview. Infect Dis Ther. 2023;12(3):829–841. doi: 10.1007/s40121-023-00771-8.
89. de Jonge L, Bos HJ, van Langen IM, de Jong-van den Berg LTW, Bakker MK. Antibiotics prescribed before, during and after pregnancy in the Netherlands: a drug utilization study. Pharmacoepidemiology and Drug. 2014;23(1):60–68. doi: 10.1002/pds.3492.
90. Freitas AR, Werner G. Nosocomial pathogens and antimicrobial resistance: modern challenges and future opportunities. Microorganisms. 2023;11(7):1685. doi: 10.3390/microorganisms11071685.
91. Coelho GDP, Ayres LFA, Barreto DS, Henriques BD, Prado M, Passos CMD. Acquisition of microbiota according to the type of birth: an integrative review. Rev Lat Am Enfermagem. 2021;29:e3446. doi: 10.1590/1518.8345.4466.3446.
92. Song SJ, Wang J, Martino C, Jiang L, Thompson WK, Shenhav L, McDonald D, Marotz C, Harris PR, Hernandez CD, et al. Naturalization of the microbiota developmental trajectory of cesarean-born neonates after vaginal seeding. Med. 2021;2(8):951–964.e5. doi: 10.1016/j.medj.2021.05.003.
93. Garsin DA, Sifri CD, Mylonakis E, Qin X, Singh KV, Murray BE, Calderwood SB, Ausubel FM. A simple model host for identifying gram-positive virulence factors. Proc Natl Acad Sci, India, Sect B Biol Sci. 2001;98(19):10892–10897. doi: 10.1073/pnas.191378698.
94. Wang H, Chu W, Ye C, Gaeta B, Tao H, Wang M, Qiu Z. Chlorogenic acid attenuates virulence factors and pathogenicity of pseudomonas aeruginosa by regulating quorum sensing. Appl Microbiol Biotechnol. 2019;103(2):903–915. doi: 10.1007/s00253-018-9482-7.
95. Nandu TG, Subramenium GA, Shiburaj S, Viszwapriya D, Iyer PM, Balamurugan K, Rameshkumar KB, Karutha Pandian S. Fukugiside, a biflavonoid from garcinia travancorica inhibits biofilm formation of streptococcus pyogenes and its associated virulence factors. J Med Microbiol. 2018;67 (9):1391–1401. doi: 10.1099/jmm.0.000799.
96. Aswathanarayan JB, Vittal RR. Inhibition of biofilm formation and quorum sensing mediated phenotypes by berberine in pseudomonas aeruginosa and salmonella typhimurium. RSC Adv. 2018;8(63):36133–36141. doi: 10.1039/C8RA06413J.
97. Pöntinen AK, Top J, Arredondo-Alonso S, Tonkin-Hill G, Freitas AR, Novais C, Gladstone RA, Pesonen M, Meneses R, Pesonen H, et al. Apparent nosocomial adaptation of enterococcus faecalis predates the modern hospital era. Nat Commun. 2021;12(1):1523. doi: 10.1038/s41467-021-21749-5.
98. Barreto Miranda I, Ignatius R, Pfüller R, Friedrich-Jänicke B, Steiner F, Paland M, Dieckmann S, Schaufler K, Wieler LH, Guenther S. High carriage rate of ESBL-producing Enterobacteriaceae at presentation and follow-up among travellers with

gastrointestinal complaints returning from India and Southeast Asia. J Travel Med. 2016;23(2):tav024. doi: 10.1093/jtm/tav024.

99. Bender JK, Hermes J, Zabel LT, Haller S, Mürter N, Blank HP, Werner G, Hüttner I, Eckmanns T. Controlling an unprecedented outbreak with vancomycin-resistant enterococcus faecium in Germany, October 2015 to November 2019. Microorganisms. 2022;10(8):1603. doi: 10.3390/microorganisms10081603.
100. Cimen C, Berends MS, Bathoorn E, Lokate M, Voss A, Friedrich AW, Glasner C, Hamprecht A. Vancomycin-resistant enterococci (VRE) in hospital settings across European borders: a scoping review comparing the epidemiology in the Netherlands and Germany. Antimicrob Resist Infect Control. 2023;12(1):78. doi: 10.1186/s13756-023-01278-0.
101. Piezzi V, Wassilew N, Atkinson A, D'Incau S, Kaspar T, Seth-Smith HM, Casanova C, Bittel P, Jent P, Sommerstein R, et al. Nosocomial outbreak of vancomycin-resistant enterococcus faecium (VRE) ST796, Switzerland, 2017 to 2020. Euro Surveill. 2022;27(48). doi: 10.2807/1560-7917.ES.2022.27.48.2200285.
102. Wu Y, Tang W, Ding Y, Zhao Y, Zhou C, Cheng Y, Dong R, Li Y, Zhao G, Xu A, et al. Enterococcus faecalis SI-FC-01 enhances the stress resistance and healthspan of C. elegans via AKT signaling pathway. Sci Rep. 2025;15(1):14454. doi: 10.1038/s41598-025-98440-y.
103. Tadesse BT, Svetlicic E, Zhao S, Berhane N, Jers C, Solem C, Mijakovic I. Bad to the bone? – Genomic analysis of enterococcus isolates from diverse environments reveals that most are safe and display potential as food fermentation microorganisms. Microbiological Res. 2024;283:127702. doi: 10.1016/j.micres.2024.127702.
104. Schiwon K, Arends K, Rogowski KM, Fürch S, Prescha K, Sakinc T, Van Houdt R, Werner G, Grohmann E. Comparison of antibiotic resistance, biofilm formation and conjugative transfer of staphylococcus and enterococcus isolates from International space station and antarctic research station concordia. Microb Ecol. 2013;65(3):638–651. doi: 10.1007/s00248-013-0193-4.
105. Coque TM, Cantón R, Pérez-Cobas AE, Fernández-de-Bobadilla MD, Baquero F. Antimicrobial resistance in the global health network: known unknowns and challenges for efficient responses in the 21st century. Microorganisms. 2023;11(4):1050. doi: 10.3390/microorganisms11041050.
106. Neuhaus K. Molecular ecology of food-borne bacteria in the environment and food chain: bacterial interpenetration of trophic levels [thesis]. München, Germany. 2016.