UNIVERSITÀ CATTOLICA DEL SACRO CUORE

Sede di Piacenza

Dottorato di ricerca per il Sistema Agro-alimentare

Ph.D. in Agro-Food System

Cycle XXXV

S.S.D. AGR/16

Genomic insights for safety assessment of foodborne bacteria.

Coordinator:

Ch.mo Prof. Paolo Ajmone Marsan

Candidate:

Mireya Viviana Belloso Daza

Matriculation n: 4915184

Academic Year 2022/2023

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

Sede di Piacenza

Dottorato di ricerca per il Sistema Agro-alimentare

Ph.D. in Agro-Food System

Cycle XXXV

S.S.D. AGR/16

Genomic insights for safety assessment of foodborne bacteria.

Coordinator:

Ch.mo Prof. Paolo Ajmone Marsan

Tutor:

Prof. Pier Sandro Cocconcelli

Candidate:

Mireya Viviana Belloso Daza

Matriculation n: 4915184

Academic Year 2022/2023

Table of Contents

Preface

Safe food and the access to it is key to sustaining life and promoting good health. Unsafe food containing harmful microorganisms or chemical substances causes more than 200 diseases, ranging from diarrhoea to cancers that particularly affect infants, young children, elderly and immunocompromised individuals. The global burden of foodborne disease affects public health, society, and economy, therefore good collaboration between governments, producers and consumers is needed to help ensure food safety and stronger food systems. The most recent survey conducted by WHO (2015) showed an estimated 600 million ill individuals and 420 000 yearly deaths associated to unsafe food. The economic impact is mainly due to the lack of safe food in low and middle income causing a US$ 110 billion is lost each year in productivity and medical expenses. The main challenges to assure food safety remain tied to our food production and supply chain, where factors like environmental contamination, consumer preferences, timely detection and surveillance of outbreaks play a crucial role. Recently, DNA-based methodologies for microbial detection and investigation have sparked special interest, mainly linked to the development of sequencing technologies. Contrary to the traditional culture-dependent methods, DNA-based techniques such as Whole Genome Sequencing (WGS) that targets fast and sensitive results at a relative low price and short processing time. Moreover, WGS confers high discriminatory power that allows to determine important genomic characteristics linked to food safety like taxonomy, pathogenic potential, virulence and antimicrobial resistance and the genetic transfer thereof. The understanding of these characteristics is fundamental to design detection and mitigation strategies to apply along the entire food-chain following a 'One Health' perspective, leading to gain knowledge about the microbiota that affect humans, animals, and environment.

The aim of the thesis is to gain insight into the genomics of foodborne microbes for their characterization and to create or improve strategies for their detection and mitigation methods. Particularly, this thesis is focused on the assessment of the pathogenic potential based on genomic analyses including taxonomy, virulence, antibiotic resistance and mobilome studies. The second focus is to profit from the genomic insights to design rapid and time-effective detection devices and reliable mitigation methods to tackle foodborne pathogens. In more details the following topics will be handled:

The presence of multi-drug resistant strains in ready-to-eat fermented food represents a risk of public health for the spread of AMR determinants in the food chain and in the gut microbiota of consumers. Genomic analyses permitted to accurately assess the safety of *E. faecium* strain UC7251, with respect to its virulence and co-location of antibiotic and heavy metal resistance genes in mobile elements with conjugation capacity in different matrices. This work emphasizes the importance of a surveillance for the presence of AMR bacteria in food and to incite the development of innovative strategies for the mitigation of the risk related to antimicrobial resistance diffusion in food.

The accuracy of taxonomic identification drives the subsequent analysis and, for this reason, a suitable method to identify species is crucial. The species re-classification of *Enterococcus faecium* clade B was investigated,

using a combined approach of phylogenomics, multilocus sequence typing, average nucleotide identity and digital DNA–DNA hybridization. The goal is to show how the genome analysis is more effective and give more detailed results concerning the species definition, respect to the analysis of the 16S rRNA sequence. This led to the proposal to reclassify all the *E. faecium* clade B as *E. lactis,* recognizing the two groups are phylogenetically separate, where a specific safety assessment procedure can be designed, before their use in food or as probiotics, including the consideration for inclusion in the European QPS list.

From this taxonomic re-classification, we developed a PCR-based method for rapid detection and differentiation of these two species and to discuss main phenotypic and genotypic differences from a clinical perspective. To this aim, core-genome alignment base on pangenome analysis was used. Allelic difference between certain core genes allowed primer design and species identification through PCR with 100% specificity and no cross-reactivity. Moreover, clinical *E. lactis* genomes categorised as a potential risk due to the ability of enhanced bacterial translocation.

Antimicrobial agents that are alternative to antibiotics are one of the main areas of development and improvement in the current food chain. Metallic nanoparticles like Platinum nanoparticles (PtNPs), have awaken interest due to their potent catalytic activities similar to oxidases and peroxidases granting strong antimicrobial effects, have been proposed as potential candidates to overcome the drawbacks of antibiotics like drug resistance. The goal is to study the mode of action of PtNPs related to biofilm formation capacity, reactive oxygen species (ROS) coping mechanism and quorum sensing using foodborne bacteria.

CHAPTER 1

General introduction: Importance of WGS for surveillance, detection, and mitigation of foodborne bacteria.

Chapter 1

Importance of genomic surveillance for safety assessment, detection, and mitigation of foodborne bacteria.

1 Current risk assessment and implementation with WGS

The current food safety systems are facing challenges to improve the key components of the framework regarding regulatory schemes, surveillance, coordination mechanisms, emergency response, and food safety education and training. The way foods are being produced, delivered, and consumed are changing day-to-day and, together with the expanding globalization, ensuring food safety is a shared responsibility among many stakeholders. Food monitoring and surveillance systems are crucial for risk assessment and prevention of potential foodborne outbreaks and forecast of potential emerging threats. Recently, according to WHO/FAO in compliance with Codex Alimentarius, the Microbiological Risk Assessment Guidance for Food describes the risk assessment as an integral approach through the estimation of risk, uncertainty analysis and transparency. This approach is based on hazard identification and hazard characterization of microbial infectious agents or toxins and their adverse effects on human health, where the use of Whole Genome Sequencing (WGS) derived data is not adopted as a mandatory tool for outbreak prevention and epidemiological studies (1). Nevertheless, the European Food Safety Authority (EFSA)) developed the EFSA One Health WGS System that interoperates with the ECDC Molecular Typing to base hazard identification on a combination of laboratory-based studies and molecular typing methods based on WGS (2, 3). With that in mind, the overall benchmark for the current risk assessment framework to ensure microbiological safety in foods is based on a combination of microbial typing methods based on traditional molecular techniques that allow the identification of microbial pathogens. Yet, with the advancements in sequencing technologies, open sharing of genomic data and open source bioinformatic software tools are revolutionizing food safety science (4). Genomics and bioinformatics have been crucial in developing the standard molecular typing methodologies used for laboratory-based detection and investigation of foodborne disease outbreaks, the most used being Pulse-Field Gel Electrophoresis (PFGE), Multi-Locus Sequence Typing (MLST), and Multi-Locus Variable-Number of Tandem Repeats Analysis (MLVA) (5). However, these tests require limited bioinformatics and genomics compared to the potential contribution that this resource could provide. Together with the rapid development of different next-generation sequencing (NGS) and high-throughput technologies, WGS might soon replace routine molecular approaches for routine typing of microbial genomes. Cost-, labour- and time-effectiveness and high sensitivity and resolution are the strongest improvements to traditional techniques (6). Given that foodborne diseases represent a burden to public health linked especially to high medical costs and economic losses due to food recalls, the rapid identification of foodborne microbes and pathogenicity traits is crucial. Implementation of WGS has the potential to improve different aspects linked to foodborne bacteria, such as the management of infectious diseases, the prompt intervention during food outbreaks, the support of risk assessment, and the re-organization of taxonomy. Whitin this scope, the identification and characterization the pathogenicity potential of foodborne

microbes include the assessment of genes that contribute to virulence and antimicrobial resistance (AMR). In that respect, WGS is expected to improve the surveillance of emerging AMR and virulence factors (VF) by providing a greater understanding of the transmission of these specialty genes throughout the food chain, and therefore lead to a transformation of risk assessment in the food industry (7). The purpose of the application of WGS expands to the investigation of microorganisms and by-products thereof, intentionally added to food and feed as additives to support the risk assessment of these regulated products. Certainly, genome data can be exploited to study taxonomy and evolution relationships and presence of genes of concern carried by some strains usually applied in the food production (8).

2 Genomic data to support food safety assessment

2.1. Taxonomy and epidemiology

WGS is useful not only in the characterization of pathogens, but also for the identification of taxonomic relationships that could help to recognize the origin and cause at the beginning of an outbreak. At any point in time, a snapshot of pathogen DNA gathered from infected individuals can be analyzed to reconstruct the history of those transmission events. This evolutionary history and genome relatedness to other pathogens can provide information about the origin of disease outbreaks, including whether new strains are entering the population, and can help construct the epidemiology network (Table 1) (9).

The relationship between strains isolated in different steps of the food chain and from patients can be assessed with the investigation of Single Nucleotide Polymorphisms (SNPs) and allele-based approaches. A small number of different nucleotides detected in different isolates support the hypothesis that the analyzed genomes originate from the same strain. This means that a food product can be linked to a human infection. The genetic diversity should also consider the mutation rate, which characterizes different species; therefore, the exact number of SNPs to correlate different isolates should be determined case by case, in a species-based manner (10). Allele-based approaches, such as MLST and MLVA, where orthologs are identified using an automated approach against a curated database of possible alleles is used to confer a sequence type assigned to the isolate that can be used for downstream phylogenetic analyses (11).

Furthermore, taxonomy analyses based on 16S rRNA gene has been applied as the gold standard for sequence-based bacterial analysis for decades. However, it has been demonstrated several times that the targeting of this gene does not hold the high enough resolution capacity for species assignment (12). As WGS has become more widely accessible, tools such Average Nucleotide Identity (ANI) and Digital DNA-DNA Hybridization (dDDH) and pangenome analysis, provide the ultimate classification methods for microbial taxonomy, necessary to establish a valid Overall Genome Related Index (OGRI) (13). OGRI defines threshold values for 16S rRNA gene sequence similarity (98.65%), ANI (96%) and DDH (70%), which provide accurate results for species delineation. For example, using the OGRI approach, *Enterococcus faecium* clade B was demonstrated to belong to the *Enterococcus lactis* (14). Also, the reclassification and new genera and species delineation of the *Lactobacillaceae* family was performed using this approach (15). The implementation of these powerful tools

is wide applicable to all type of bacteria and has definitive implication on the evolution of the taxonomical assignment.

2.2. Antimicrobial resistance genes, virulence and mobile genetic elements

Together with taxonomical and epidemiological data, surveillance system of outbreaks involves the determination of genomic characteristics including traits like AMR profile, VFs and mobile genetic elements (MGE) (Table 1). Today, WGS has already become a crucial part for the safety assessment as recommended by European Food Safety Authority (EFSA) and the World Health Organization (WHO) encompassing the One Health approach to integrate human health, animal health and the environment (16). The use of WGS can refine the description of AMR genes, VF markers and MGE to unify the hazard identification and evaluate potential risks. Regarding genetic markers related to virulence, such as those conferring the capacity of attachment, adhesion, invasion or replication, genomic data facilitate their recognition and assessment. Several VF prediction tools have been developed, a few examples are VirulenceFinder (17), SPIFinder and VFDB (18). However, genes providing higher pathogenicity or are not always previously known. Whole Genome Association Studies (WGAS) is a useful tool that facilitates the identification genetic risk factors like genes, k-mers, mutations or SNPs associated with increased pathogenicity or virulence (19). Furthermore, as for AMR, the limit of resolution of phenotypic techniques is overcome by exhaustive list of AMR genes that can be predicted from molecular data, including the potential for occurrence of multidrug resistance (MDR) not only in pathogens but also foodborne bacteria such as probiotics (20). Bioinformaticians have developed multiple tools, mainly prediction tools, to detect the presence of AMR genes (acquired or spontaneous mutation) against a known reference database or with a gene annotation approach. Some of these tools are can be found on the Center for Genomic Epidemiology (CGE) like ResFinder (21), Comprehensive Antibiotic Resistance Database (CARD)(22), BLAST (Basic Local Alignment Search Tool) (23), ABRicate , Search Engine for Antimicrobial Resistance (SEAR) (24), ARG-ANNOT and Antimicrobial Resistance Identification By Assembly (ARIBA). Moreover, AMR can occur through various mechanisms, including mutations of chromosomal genes and the acquisition of antibiotic resistance genes from other strains in a process termed horizontal gene transfer (HGT). HGT of AMR genes between bacteria can occur via three main mechanisms, namely transformation (uptake of naked DNA), transduction (transfer by bacteriophages) and conjugation (transfer by plasmids and other MGE) (25, 26). Conjugation, in particular, seems to play an important role in the transmission and spread of foodborne AMR of public health importance, and it can occur within the gastrointestinal tract of animals and humans (27). Though WGS-derived genomic data is not able to provide information about the frequency of these events like in vivo or in vitro tests, it can be used to whether a plasmid has the molecular machinery for conjugation to occur and therefore predict the mobility profile of a plasmid (28). As an example, various studies investigated the occurrence of MDR strains in different foods using WGS, the most alarming being ready-to-eat foods such as, dairy and fermented meats (29–31). Particularly, the case of fermented sausage *E. faecium* isolate strain UC7251, where WGS and downstream analyses elucidated multiple antibiotic and heavy metal resistance cassettes on one mobilizable plasmid and a chromosomal transposon Tn916, giving insights into the AMR

dissemination in the swine production and consumption chain (31). Several other studies have been conducted using WGS as tool for AMR and virulence factors distribution, one of the mnay examples of WGS-based surveillance of AMR determinants include seven *E. coli* STEC strains found in raw milk cheese containing other AMR and virulence factors hazardous for human health (32), and the epidemiological reconstruction of *Salmonella* zoonosis with multiple serotypes, where the prevalence of AMR genes, plasmid replicons and virulence genes that were identical in different species highlighted exchange of serovars across different hosts (33).

2.3. Quorum Sensing and Biofilm Formation and Stress-response systems

Quorum Sensing (QS) is a communication mechanism among bacteria that enables the control of processes such as biofilm formation, virulence factor expression and stress adaptation mechanisms. The recent advances of high throughput sequencing and the increasing amount of genomic data have enabled to uncover the hidden secrets of microbial dynamics to survive and disseminate throughout the environment compromising so the One Health continuum (34). However, regarding QS further analyses must be performed in order to profit from the deep insights the genome sequence can provide in order to better understand this mechanism in different species.

One example is the study of QS mechanisms of *Pseudomonas aeruginosa* PAO1 through WGS analysis, where it was revealed that QS in *P. aeruginosa* is also regulated by transcriptional regulator *mexT* other than the normal *Rhl* system (35). Also, through similar genomic scrutiny it was revealed that low virulence *Staphylococcus aureus* develop mutations in quorum sensing system *agr,* highly involved in the control of virulence factors and biofilm formation (36). Moreover, concerning the latter, biofilms endure exogenous stressors through the production of EPS (Extracellular Polymeric Substance) enhancing adherence and anchorage of bacteria.

Two studies studied the biofilm formation capacity of *Listeria monocytogenes* and the corresponding distribution and composition of genetic operons coding for them. WGS revealed that the presence of genetic markers *inL*, *SSi* and *ermC* are strictly related to source of isolation, further help in the characterization of clonal complexes and so, epidemiology (37, 38).

WGS also unveiled that the enterococcal surface protein gene *esp* in *E. faecalis* and *E. faecium*, highly involved in biofilm production and surface attachment, in in fact inside a pathogenicity island (PAI).ecalis.. In Interestingly, genome derived data showed that in *E. faecium,* only isolates from clinical infections contain the PAI with *esp*, whereas in *E. faecalis* commensal (human, animal) and environmental strains may also contain it (39, 40). Indeed, apart from the quorum sensing machinery, bacteria are exposed to changing and challenging environmental conditions that are perceived as stresses, to which they need to adapt to ensure their survival. Pathogens in particular, have evolved intricate systems known as stress involving several molecular pathways at the levels of transcription, translation, and stability of transcripts and of proteins. A better understanding of these stress response mechanisms may be useful for developing new strategies to fight bacteria that, in some cases, represent an important life threat. A study by Liu and colleagues used NGS data to investigate the oxidative stress response of *Salmonella* ser. Enteritidis. Their findings suggested that there is a relation between

virulence and oxidative stress, as oxidatively stressed *S.* Enteritidis cells simultaneously repressed key motility encoding genes and induced a wide range of adhesin- and salmonellae-essential virulence-encoding genes, that are critical for the biofilm formation and intracellular survival (41).

Table 1. Schematic summary on the advantages and disadvantages of genomic surveillance of foodborne pathogens

	Taxonomy and Epidemiology	AMR, VF and MGEs	Detection and Mitigation
Advantages	- Local and global transmission reconstructed by WGS have shown that different clones have emerged following multiple independent events worldwide and have elucidated the role of this zoonotic pathogens in the spread of AMR. - WGS-derived dDDH, ANI and in silico 16S rRNA of a large set of available genomes facilitated the differentiation between clades and species. - Traditional serotyping using PCR is error prone, WGS has facilitated the identification of serotypes giving a more complete outlook of outbreak and non-outbreak isolates. - Nation-level systematic comparison of MLVA, core genome SNP (cgSNP), and core genome MLST (cgMLST) indicated that a combination of WGS and MLVA is a realistic approach to improve pathogen surveillance.	- Discovery of further putative virulence factors contributing to virulence other than genes defined by EFSA, derived from homology search in a large set of genomes. - Hybrid sequencing allows to predict transposon, bacteriophages and plasmids harbouring AMR genes with potential transfer capacities through conjugation and transduction in food and environmental models. - WGS can provide comprehensive resistance genotypes and is capable of accurately predicting resistance phenotypes, making it a valuable tool for surveillance.	- Genomic scrutiny allows for the detection of specialty genes that can be potentially used to design DNA specific probes for bacterial detection in different matrices. - Information gained from surveillance can then inform policy and risk mitigation strategies to combat increasing AMR and protect antimicrobials important to human health. - The rapid transformation from molecular epidemiology to genomics of infectious diseases is ushering in a new era of "precision public health" by uncovering the detailed dynamics of infection transmission and antimicrobial resistance to enable more effective and better targeted control interventions.
Disadvantages	- A key factor in the management of global outbreaks is the early and successful coordination of hospital practices and government agencies around the world. - Global outbreak management can succeed only if WGS data can be acquired, stored, and, most importantly, shared consistently in a variety of clinical settings. - The price and application range of most commercial software for bioinformatic analysis should harmonize affordability and user-friendliness.	- At the present time, WGS-based analyses cannot yield an inferred MIC or zone diameter. Hence the potential utility of WGS-based approaches must be considered at the level of detecting gene presence or absence. - WGS does not directly provide information on levels of gene expression. Although other technologies can do so, e.g. RNA sequencing, it seems unlikely that these will find a place in the clinical laboratory before WGS.	- To fully utilize the potential of WGS, open access databases need to be implemented for sharing the WGS data, while safeguarding legitimate data protection and ownership. - Further development, critical evaluation and harmonized application of WGS-based typing solutions for detection and mitigation of pathogens can only be delivered through engaging intersectoral and international collaborations. - Standardization of WGS outbreak analyses is key to ensure a precise level of flexibility to be tailored to the needs of specific situations and health care practices with various resources and capabilities.

3 WGS-based surveillance and mitigation of foodborne pathogens

Contamination in food and food producing environment led to a decreased in productivity and an increased health risk. Bacterial contamination frequently occurs as biofilms on foods or food contact materials. Biofilms allow bacteria to be more resistant to stress enhancers like sanitizers and antimicrobials. Sessile communities of bacteria have shown increased transmission of antimicrobial resistance, often associated with multi-drug resistance (MDR), and so have been recognized as a major AMR reservoir and transmission source. The development of innovative tools is urgently needed as often traditional detection and mitigation techniques are not as sensitive and effective to prevent bacterial growth. The use of WGS expands beyond hazard identification (AMR, VF, MGE) for outbreak prevention but reaches a further dimension where genomic information can be exploited to develop innovative techniques detection and mitigation of foodborne pathogens. A few characteristics that these methodologies can provide are rapid, on-site and sensitive tests that are also user-friendly and can be used along the entire food chain or so called Point-Of-Need (PON) approach (42). The genomic data provided by WGS can be used principally to design probes that are specific for determining hazards from a species level to very genomic features contributing to pathogenicity. Emerging methods commonly incorporate nucleic acids amplification, immunosensors, metabolic assays and nanomaterials. Such hybrid techniques aim to reduce the complexity of food screening processes and increase feasibility for in situ detection.

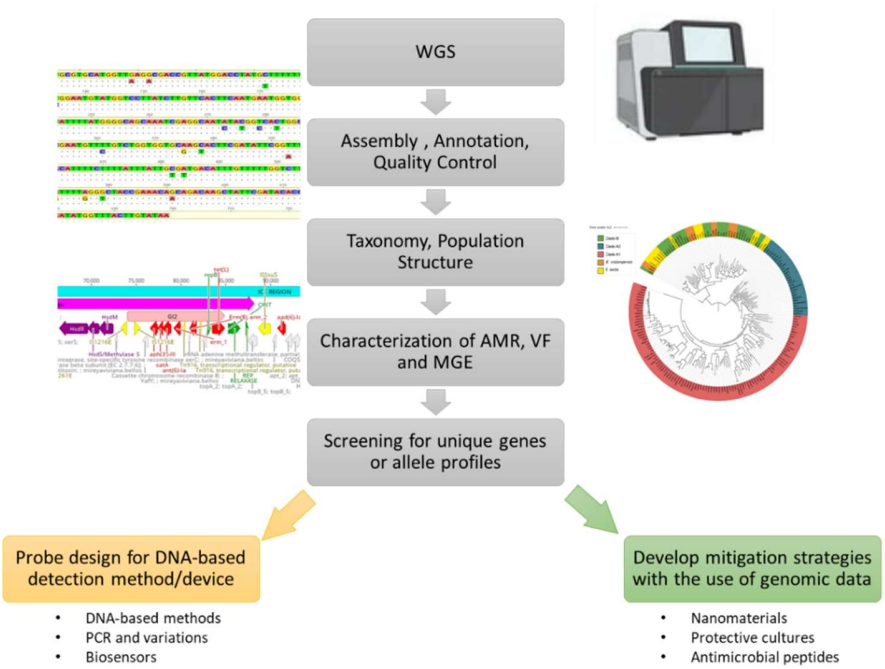


Figure 1. Surveillance and control pipeline of foodborne pathogens based on WGS-derived data.

3.1. Detection and biosensing

Amplification of DNA is essential for the detection of target food pathogen. These probes are generally amplified via polymerase chain reaction (PCR), the most commonly used nucleic acid amplification method for detecting pathogenic microorganisms. Since its discovery many different advances on the original PCR protocol have been describe and the viability of PCR tests for rapid detection of foodborne pathogens has been extensively explored (43). This method can detect a single copy of a target DNA sequence with respect to single pathogen in food. It is permissible because it amplifies the target organism sequence rather than the signal and by producing less false positives. So, PCR has become a very widely used detection method for the food-borne pathogens using nucleic acid as a target. Additionally, Real time PCR or quantitative PCR (qPCR) and Loop-mediated isothermal amplification (LAMP), have been successfully applied for diagnostics as well (43). qPCR is used to quantify DNA and cDNA, determining gene or transcript numbers present within a sample. Among its advantages there is a quick turn-around time, high sensitivity, and user-friendly technology . Furthermore, the LAMP method does not require a cyclic process with a specific temperature profile, but amplifies DNA with greater specificity, efficiency, and rapidity by maintaining a uniform temperature for an elongated period. Many amplification techniques should combine simple design, cost efficiency, and user-friendliness for development of Point-Of-Care Testing (POCT), which is one of the big challenges yet to achieve (44).

Current biosensor research is directed towards integration of Nucleic Acid Testing into microfluidic devices to further increase the biosensing capacity and develop diagnostic tools that can meet affordable, sensitive, specific, user-friendly, Rrpid and robust, equipment-free and deliverable (ASSURED) criteria recommended by the World Health Organization (WHO) (45). Use of nucleic acids for detection of pathogens involves the extraction of the dNA/RNA from the food matrix, an amplification stepas the complete test is minimized to several hours. In general, the procedure for rapid detection of DNA/RNA from pathogenic bacterial comprises five steps: pre-concentration, extraction, detection, signal transduction into a measurable signal and data analysis (46).

Furthermore, nanoscale agents, such as nanomaterials have awakened the interest in the field of biosensing and have presented great potential for pathogen detection. Nanomaterials are material units that are less than 100 nm in size, and biomarkers are biological molecules taken from an organism that can be reproduced for testing and analysis. Nanomaterials that have been utilized with biomarkers to detect foodborne and waterborne pathogens include quantum dots, gold, silver, magnetic materials, metal oxides, and carbon-based materials (47). Recent developments in nanotechnology have allowed for more rapid, accurate, and cost-effective biosensors for pathogen monitoring. Advances in the manipulation of these nanomaterials permit binding of different biomolecules such as bacteria, toxins, proteins, and nucleic acids. The nanosensors are operated at a scale similar to the biological processes to increase the specificity of biological response. The key challenge for developing effective biosensors

is their specificity, sensitivity and detection time to assess the presence of food-borne pathogens in normal and toxin-supplemented samples (48). Metallic nanoparticles such as gold and silver have been used in signal amplification of numerous bio-diagnostic devices. Gold nanoparticles (AuNPs) in particular have been used in a variable optical and electrical assay. The redox activity is an interesting characteristic of gold NPs, which enhances the sensitivity of electrochemical biosensors in the analysis of foodborne pathogens (48). In general, the use of gold NPs onto electrochemical biosensors in conjugation with ssDNA complementary to the microbial DNA under evaluation improves their binding with DNA-gold NPs on the transducer surface and enhances the sensitivity of the developed biosensor.

3.2. Mitigation and prevention

Information gathered from surveillance can contribute to the development of risk mitigation strategies to combat increasing widespread of harmful bacteria and their implications like AMR dissemination. Starting from the food processing steps, controlling the growth of microorganisms and limiting the transmission/expression of hazards by applying Good Agricultural Practices (GAP), and later in the food industry in the way of production, storage, processing, and distribution of foods of animal origin (49). Additionally, prudent antimicrobial use in animal husbandries and control procedures targeting all foods of animal origin throughout the processing are the main effective intervention strategies to prevent the transmission of resistant bacteria from foods to humans and vice versa. Moreover, the transmission of AMR can be tackled during production and between food industry workers carrying the resistant strains so that person-to-person spread of these pathogens in animal food sector can be reduced. So, efforts to prevent such a challenge should also be built on application of effective food safety management, including Good Manufacturing Practices (GMP) and Good Hygienic Practices (GHP) (50). Moreover, it is clear that AMR bacteria have outpaced any diagnosis or surveillance procedures, leaving behind untreatable infections in humans and animals, but, data generated from WGS, for example population structure genomics, may warn us from repeating patterns and emerging risks (51). The use of bioprotective cultures like lactic acid bacteria (LAB) is also proposed as a sustainable alternative to antimicrobials like antibiotics. The principal effect of LAB derives mainly from a decrease of pH values in foods as well as the antibacterial activity of organic acids or peptides (bacteriocins) and bacteriocin-like inhibitory substances (BLISs). However, special attention should be taken to minimize the potential role of LAB as a source of transferable AMR genes throughout the food chain. WGS plays a pivotal role in performing a complete risk assessment supported by a thorough study of their biochemical and genetic characteristics to determine the presence of AMR genes and their potential to transfer them to other bacteria. The strategies designed to combat this growing threat of MDR pathogens, face a particular challenge as there is a rapid dissemination of resistance genes between bacteria (52, 53). Other than natural occurring antimicrobial peptides, WGS has opened the door for the *in silico* design of novel antimicrobial peptide sequences from biosynthetic gene clusters using genome mining tools (54). With the use of machine-learning technologies and WGS, many databases with publicly available AMP sequences are being used to gather

data, extract potential peptide sequences and to predict their performance with Train-Test and K-fold cross-validation (55). Similar to detection strategies, WGS can aid in the design and development of mitigation techniques, or a combination of them, to tackle specific genetic traits or metabolic pathways. Novel technologies like nanoscale tools, can be combined with pathogen-specific DNA probes in combination with, for example, metallic or metal-oxide nanoparticles. A few of the mechanisms of include cell structure disruption, disintegration of cell membrane, protein dysfunction, generation of reactive oxygen species, impairing nutrient assimilation and others (56–58). Other than their intrinsic properties, nanoscale tools can be exploited as nanocarriers for antimicrobials. Efficient delivery of drugs at the right dosage has been demonstrated. Some examples are lipid-based, metallic-based and polymeric nanocarriers. However, recent research has shown that nucleic acid nanocarriers have improved physicochemical properties and antibacterial effect and possess excellent biocompatibility, biodegradability and targeting properties. For instance, they have shown promising results against bacterial biofilms, have excellent antimicrobial activities and reduce the effect of bacterial toxins (59).

4 Current Challenges of WGS

While this powerful tool brings many benefits for the scope of safety assessment, there are many different aspects that need to be improved for it to be applied routinely and become a gold standard. Some of these factors concern the standardization of wet lab protocols and bioinformatics pipelines, the ability to manage a large amount of data, the interpretation results and, the way of data sharing compliant with regulations (4). The effectiveness of the current standard techniques in microbiology and molecular biology is reduced by the extensive hands-on time protocols and associated to high costs. The workflow to obtain a genome sequence includes a part of wet lab (DNA extraction, library preparation and sequencing procedure) and a bioinformatic part, in which all the raw data are managed to check the quality (coverage and contaminations). Quality control procedures are required for all components of the WGS process including sample DNA quality and quantity, sequence quality scores including depth of sequence coverage, read length and sequence quality (60). As with other WGS components, the bioinformatics analysis process, once optimized, needs to be version controlled and any subsequent alterations will require some form of revalidation. With bioinformatic tools accessible to everyone, the challenge remains in data interpretation and quality of the analysis. Given that the data and results management should be handled by trained microbiologist, which rarely are trained as bioinformaticians, makes it a challenge. The use of command-based tools imposes a gap between non-connoisseurs and specialists, making the harmonization even harder to reach. Therefore, many web-based bioinformatic tools have been developed to obtain relevant information from genome analysis, both for expert and amateurs, with or without the availability of computational resources (61). Useful platforms have been developed for comprehensive genome analysis, from assembly to functional annotation of coding sequences, infer taxonomy and other functional integral analyses. A few examples of these

platforms are Galaxy (https://usegalaxy.org/) PATRIC/RAST (62) and Center of Genomic Epidemiology (http://genomicepidemiology.org/services/).

On Another note, the type of sequencing technology can also affect tremendously the quality and resolution of genomic data. NGS was firstly introduced with low throughput DNA fragment sequencing and has evolved to high throughput next generation and third generation sequencing techniques. Short-read sequencing is highly accurate and produces read lengths of 100-300 bp, which are then assembled into incomplete or draft genomes, this provides high read accuracy and low sequencing costs. The limitations of short-read sequencing is the lack of contiguity giving an incomplete image of the genome (63). On the other hand, single molecule sequencing technologies such as Pacific Biosciences (PacBio) and Oxford Nanopore Technologies generate long reads, which can resolve the majority of repeat elements in bacterial genomes and improve the contiguity of assemblies. However, long reads generated by these platforms could have high error rates, resulting in the introduction of single base substitutions and small indels into the final assembly. Alternatively, hybrid assembly using both short reads and long reads generated by single molecule sequencing technologies can facilitate assembly of complete bacterial genomes, but the high per-genome cost of long-read sequencing restricts the extensive use of this approach in bacterial genomics (64).

Another point of discussion concerns the management of the big amount of data generated from sequencing. The aim to use in a routinary way a pipeline for the genome analysis can be achieved only in the case that the data would be submitted in a standardized fashion. Some public repositories (e.g. SRA and ENA) make raw data available for everyone, with the risk of including low data quality. The harmonization should involve not only the management of data storage, but also the entire procedure of sequencing and the pipeline workflow; moreover, pipelines should be suitable for a precise application, including a well standardized protocol(65). The standardization also should involve the management of common, non-proprietary file formats and data storage and therefore the development and adoption of guidelines and standards for data collection, annotation, archiving, and reuse in an environment that supports user feedback and issue tracking.

With the current food safety assessment scheme, it remains unclear to which extent WGS data can be peacefully integrated into quantitative risk assessment models and how this incorporation could impact detection and mitigation measures. Gene prediction and annotation belonging to the hazard identification rely mostly on databases of known pathogenicity-related genomic features. Additionally, the study of the structure of a bacterial population and epidemiology studies is based on a reference database of isolates specific to a given point in time. In brief, the support information conferred by WGS deeply depends on the reference databases that are used to generate them (66). In this scenario, the success of outbreak investigations will also depend on how timely and accurate WGS data can be created and analyzed (67). Machine learning -based algorithms could further speed-up such investigations, especially as the number of complete microbial genomes in NCBI

(http://www.ncbi.nlm.nih.gov) is rapidly growing. Such algorithms could potentially increase the accuracy and speed of clinically and epidemiologically relevant predictions. Yet, to yield accurate predictions, these algorithms require large amounts of high-quality data and current microbial genome databases are mostly biased toward cultivable pathogenic bacteria. Future improvements are needed principally to achieve better data curation and collection such as comprehensive and standardized metadata collection from phenotypic profiling using traditional microbiology methods for isolate characterization together with high quality WGS data (60).

5 References

1. WHO. 2022. WHO GLOBAL STRATEGY FOR FOOD SAFETY 2022-2030.

2. Costa G, Di Piazza G, Koevoets P, Iacono G, Liebana E, Pasinato L, Rizzi V, Rossi M. 2022. Guidelines for reporting Whole Genome Sequencing-based typing data through the EFSA One Health WGS System. EFSA Support Publ 19.

3. EFSA BIOHAZ Panel, Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Cocconcelli PS, Fernández Escámez PS, Maradona MP, Querol A, Suarez JE, Sundh I, Vlak J, Barizzone F, Correia S, Herman L. 2021. The list of QPS status recommended biological agents for safety risk assessments carried out by EFSA https://doi.org/10.5281/ZENODO.4428353.

4. Taboada EN, Graham MR, Carriço JA, Van Domselaar G. 2017. Food Safety in the Age of Next Generation Sequencing, Bioinformatics, and Open Data Access. Front Microbiol 8.

5. Wei X, Zhao X. 2021. Advances in typing and identification of foodborne pathogens. Curr Opin Food Sci 37:52–57.

6. Messens W, Hugas M, Afonso A, Aguilera J, Berendonk TU, Carattoli A, Dhollander S, Gerner-Smidt P, Kriz N, Liebana E, Medlock J, Robinson T, Stella P, Waltner-Toews D, Catchpole M. 2019. Advancing biological hazards risk assessment. EFSA J 17:170714.

7. Pennone V, Cobo-Díaz JF, Prieto-Maradona M, Álvarez-Ordóñez A. 2022. Integration of genomics in surveillance and risk assessment for outbreak investigation. EFSA J 20.

8. Food E, Authority S. 2021. EFSA statement on the requirements for whole genome sequence analysis of microorganisms intentionally used in the food chain. EFSA J 19.

9. Timme RE, Strain E, Baugher JD, Davis S, Gonzalez-Escalona N, Leon MS, Allard MW, Brown EW, Tallent S, Rand H. 2019. Phylogenomic pipeline validation for foodborne pathogen disease surveillance. J Clin Microbiol 57:1–11.

10. Singh N, Lapierre P, Quinlan TM, Halse TA, Wirth S, Dickinson MC, Lasek-Nesselquist E, Musser KA. 2019. Whole-Genome Single-Nucleotide Polymorphism (SNP) Analysis Applied Directly to Stool for Genotyping Shiga Toxin-Producing Escherichia coli: an Advanced Molecular Detection Method for Foodborne Disease Surveillance and Outbreak Tracking. J Clin Microbiol 57.

11. Ronholm J, Nasheri N, Petronella N, Pagotto F. 2016. Navigating Microbiological Food Safety in the Era of Whole-Genome Sequencing. Clin Microbiol Rev 29:837–857.

12. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, Sodergren E, Weinstock GM. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun 2019 101 10:1–11.

13. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 68:461–466.

14. Belloso Daza MV, Cortimiglia C, Bassi D, Cocconcelli PS. 2021. Genome-based studies indicate that the Enterococcus faecium Clade B strains belong to Enterococcus lactis species and lack of the hospital infection associated markers. Int J Syst Evol Microbiol 71.

15. Zheng J, Wittouck S, Salvetti E, Franz CMAP, Harris HMB, Mattarelli P, O'toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Gänzle MG, Lebeer S. 2020. A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. Int J Syst Evol Microbiol 70:2782–2858.

16. Koutsoumanis K, Allende A, Alvarez-Ordóñez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Jenkins C, Malorny B, Ribeiro Duarte AS, Torpdahl M, da Silva Felício MT, Guerra B, Rossi M, Herman L. 2019. Whole genome sequencing and metagenomics for outbreak investigation, source attribution and risk assessment of food-borne microorganisms. EFSA J 17.

17. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 52:1501–1510.

18. Liu B, Zheng D, Jin Q, Chen L, Yang J. 2018. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 47:687–692.

19. Lees JA, Bentley SD. 2016. Bacterial GWAS: not just gilding the lily. Nat Rev Microbiol 2016 147

14:406–406.

20. Wang Y, Liang Q, Lu B, Shen H, Liu S, Shi Y, Leptihn S, Li H, Wei J, Liu C, Xiao H, Zheng X, Liu C, Chen H. 2021. Whole-genome analysis of probiotic product isolates reveals the presence of genes related to antimicrobial resistance, virulence factors, and toxic metabolites, posing potential health risks. BMC Genomics 22.

21. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 75:3491–3500.

22. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG. 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

24. Rowe W, Baker KS, Verner-Jeffreys D, Baker-Austin C, Ryan JJ, Maskell D, Pearce G. 2015. Search Engine for Antimicrobial Resistance: A Cloud Compatible Pipeline and Web Interface for Rapidly Detecting Antimicrobial Resistance Genes Directly from Sequence Data. PLoS One 10:e0133492.

25. Tao S, Chen H, Li N, Wang T, Liang W. 2022. The Spread of Antibiotic Resistance Genes In Vivo Model. Can J Infect Dis Med Microbiol = J Can des Mal Infect la Microbiol Médicale 2022:1–11.

26. Kondo K, Kawano M, Sugai M. 2021. Distribution of Antimicrobial Resistance and Virulence Genes within the Prophage-Associated Regions in Nosocomial Pathogens. mSphere 6.

27. Sun D. 2018. Pull in and push out: Mechanisms of horizontal gene transfer in bacteria. Front Microbiol 9:2154.

28. Guédon G, Libante V, Coluzzi C, Payot S, Leblond-Bourget N. 2017. The Obscure World of Integrative and Mobilizable Elements, Highly Widespread Elements that Pirate Bacterial Conjugative Systems. Genes (Basel) 8.

29. Jans C, Sarno E, Collineau L, Meile L, Stärk KDCC, Stephan R, Rodriguez-Lazaro D, Stessl B, Jans C, Sarno E, Collineau L, Meile L, Stärk KDCC, Stephan R. 2018. Consumer exposure to antimicrobial resistant bacteria from food at Swiss retail level. Front Microbiol 9:362.

30. Chajęcka-Wierzchowska W, Zadernowska A, García-Solache M. 2020. Ready-to-eat dairy products as a source of multidrug-resistant Enterococcus strains: Phenotypic and genotypic characteristics. J Dairy Sci 103:4068–4077.

31. Belloso Daza MV, Milani G, Cortimiglia C, Pietta E, Bassi D, Cocconcelli PS. 2022. Genomic insights of Enterococcus faecium UC7251, a multi-drug resistant strain from ready-to-eat foods, highlight the risk of antimicrobial resistance in the food chain. Front Microbiol.

32. Cortimiglia C, Borney MF, Bassi D, Cocconcelli PS. 2021. Genomic Investigation of Virulence Potential in Shiga Toxin Escherichia coli (STEC) Strains From a Semi-Hard Raw Milk Cheese. Front Microbiol 11:3642.

33. Pornsukarom S, Van Vliet AHM, Thakur S. 2018. Whole genome sequencing analysis of multiple Salmonella serovars provides insights into phylogenetic relatedness, antimicrobial resistance, and virulence markers across humans, food animals and agriculture environmental sources. BMC Genomics 19.

34. Badul S, Abia ALK, Amoako DG, Perrett K, Bester LA, Essack SY. 2021. From the Farms to the Dining Table: The Distribution and Molecular Characteristics of Antibiotic-Resistant Enterococcus spp. in Intensive Pig Farming in South Africa. Microorganisms 9.

35. Cheng X, Lu M, Qiu H, Li Y, Huang L, Dai W. 2022. Spontaneous quorum-sensing hierarchy reprogramming in Pseudomonas aeruginosa laboratory strain PAO1. AMB Express 12:4–11.

36. Mannala GK, Koettnitz J, Mohamed W, Sommer U, Lips KS, Spröer C, Bunk B, Overmann J, Hain T, Heiss C, Domann E, Alt V. 2018. Whole-genome comparison of high and low virulent Staphylococcus aureus isolates inducing implant-associated bone infections. Int J Med Microbiol 308:505–513.

37. Di Ciccio P, Rubiola S, Panebianco F, Lomonaco S, Allard M, Bianchi DM, Civera T, Chiesa F. 2022. Biofilm formation and genomic features of Listeria monocytogenes strains isolated from meat and dairy industries located in Piedmont (Italy). Int J Food Microbiol 378.

38. Maggio F, Rossi C, Chiaverini A, Ruolo A, Orsini M, Centorame P, Acciari VA, Chaves López C, Salini R, Torresi M, Serio A, Pomilio F, Paparella A. 2021. Genetic relationships and biofilm formation of Listeria monocytogenes isolated from the smoked salmon industry. Int J Food Microbiol 356.

39. Freitas AR, Tedim AP, Novais C, Coque TM, Peixe L. 2018. Distribution of putative virulence markers in Enterococcus faecium: towards a safety profile review. J Antimicrob Chemother 73:306–319.

40. Zischka M, Künne CT, Blom J, Wobser D, Sakinç T, Schmidt-Hohagen K, Dabrowski PW, Nitsche A, Hübner J, Hain T, Chakraborty T, Linke B, Goesmann A, Voget S, Daniel R, Schomburg D, Hauck R, Hafez HM, Tielen P, Jahn D, Solheim M, Sadowy E, Larsen J, Jensen LB, Ruiz-Garbajosa P, Quiñones Pérez D, Mikalsen T, Bender J, Steglich M, Nübel U, Witte W, Werner G. 2015. Comprehensive molecular, genomic and phenotypic analysis of a major clone of Enterococcus faecalis MLST ST40. BMC Genomics 16:1–20.

41. Liu X, Omar M, Abrahante JE, Nagaraja K V., Vidovic S. 2020. Insights into the oxidative stress response of salmonella enterica serovar enteritidis revealed by the next generation sequencing approach. Antioxidants 9:1–20.

42. Campbell VR, Carson MS, Lao A, Maran K, Yang EJ, Kamei DT. 2021. Point-of-Need Diagnostics for Foodborne Pathogen Screening. SLAS Technol 26:55–79.

43. Umesha S, Manukumar HM. 2017. Advanced molecular diagnostic techniques for detection of food-borne pathogens: Current applications and future challenges. https://doi.org/101080/1040839820151126701 58:84–104.

44. Cecere P, Gatto F, Cortimiglia C, Bassi D, Lucchini F, Cocconcelli PS, Pompa PP. 2021. Colorimetric point-of-care detection of Clostridium tyrobutyricum spores in milk samples. Biosensors 11.

45. Land KJ, Boeras DI, Chen XS, Ramsay AR, Peeling RW. 2018. REASSURED diagnostics to inform disease control strategies, strengthen health systems and improve patient outcomes. Nat Microbiol 2018 41 4:46–54.

46. Vidic J, Vizzini P, Manzano M, Kavanaugh D, Ramarao N, Zivkovic M, Radonic V, Knezevic N, Giouroudi I, Gadjanski I. 2019. Point-of-Need DNA Testing for Detection of Foodborne Pathogenic Bacteria. Sensors (Basel) 19.

47. Mei L, Zhu S, Liu Y, Yin W, Gu Z, Zhao Y. 2021. An overview of the use of nanozymes in antibacterial applications. Chem Eng J 418:129431.

48. Kumar H, Kuča K, Bhatia SK, Saini K, Kaushal A, Verma R, Bhalla TC, Kumar D. 2020. Applications of Nanotechnology in Sensor-Based Detection of Foodborne Pathogens. Sensors (Basel) 20.

49. George A. 2019. Antimicrobial Resistance (AMR) in the Food Chain: Trade, One Health and Codex. Trop Med Infect Dis 2019, Vol 4, Page 54 4:54.

50. Bennani H, Mateus A, Mays N, Eastmure E, Stärk KDC, Häsler B. 2020. Overview of Evidence of Antimicrobial Use and Antimicrobial Resistance in the Food Chain. Antibiot 2020, Vol 9, Page 49 9:49.

51. Baker S, Thomson N, Weill FX, Holt KE. 2018. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. Science (80-) 360:733–738.

52. Bover-Cid S, Holzapfel WH. 1999. Improved screening procedure for biogenic amine production by lactic acid bacteria. Int J Food Microbiol 53:33–41.

53. Miranda C, Contente D, Igrejas G, Câmara SPA, Dapkevicius M de LE, Poeta P. 2021. Role of Exposure to Lactic Acid Bacteria from Foods of Animal Origin in Human Health. Foods 2021, Vol 10, Page 2092 10:2092.

54. de Oliveira IMF, Godoy-Santos F, Oyama LB, Moreira SM, Dias RG, Huws SA, Creevey CJ, Mantovani HC. 2022. Whole-Genome Sequencing and Comparative Genomic Analysis of Antimicrobial Producing Streptococcus lutetiensis from the Rumen. Microorganisms 10.

55. Ramazi S, Mohammadi N, Allahverdi A, Khalili E, Abdolmaleki P. 2022. A review on antimicrobial peptides databases and the computational tools. Database J Biol Databases Curation 2022:1–17.

56. Seong M, Lee DG. 2017. Silver Nanoparticles Against Salmonella enterica Serotype Typhimurium: Role of Inner Membrane Dysfunction. Curr Microbiol 74:661–670.

57. Chlumsky O, Purkrtova S, Michova H, Sykorova H, Slepicka P, Fajstavr D, Ulbrich P, Viktorova J, Demnerova K. 2021. Antimicrobial properties of palladium and platinum nanoparticles: A new tool for combating food-borne pathogens. Int J Mol Sci 22.

58. Wang L, Hu C, Shao L. 2017. The antimicrobial activity of nanoparticles: present situation and prospects for the future. Int J Nanomedicine 12:1227.

59. Obuobi S, Škalko-Basnet N. 2020. Nucleic Acid Hybrids as Advanced Antibacterial Nanocarriers. Pharmaceutics 12:1–25.

60. Vilne B, Meistere I, Grantiņa-Ieviņa L, Ķibilds J. 2019. Machine Learning Approaches for Epidemiological Investigations of Food-Borne Disease Outbreaks. Front Microbiol 10.

61. Kåhrström CT. 2014. Techniques & Applications: Bacterial WGS made easy. Nat Rev Microbiol 12:152–153.

62. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, Chlenski P, Conrad N, Dickerman A, Dietrich EM, Gabbard JL, Gerdes S, Guard A, Kenyon RW, MacHi D, Mao C, Murphy-Olson D,

Nguyen M, Nordberg EK, Olsen GJ, Olson RD, Overbeek JC, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomas C, Vanoeffelen M, Vonstein V, Warren AS, Xia F, Xie D, Yoo H, Stevens R. 2020. The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities. Nucleic Acids Res 48:D606–D612.

63. Jagadeesan B, Gerner-Smidt P, Allard MW, Leuillet S, Winkler A, Xiao Y, Chaffron S, Van Der Vossen J, Tang S, Katase M, McClure P, Kimura B, Ching Chai L, Chapman J, Grant K. 2019. The use of next generation sequencing for improving food safety: Translation into practice. Food Microbiol 79:96–115.

64. Derakhshani H, Bernier SP, Marko VA, Surette MG. 2020. Completion of draft bacterial genomes by long-read sequencing of synthetic genomic pools. BMC Genomics 2020 211 21:1–11.

65. Balloux F, Brønstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, Wang H, Eldholm V. 2018. From Theory to Practice: Translating Whole-Genome Sequencing (WGS) into the Clinic. Trends Microbiol 26:1035–1048.

66. Collineau L, Boerlin P, Carson CA, Chapman B, Fazil A, Hetman B, McEwen SA, Jane Parmley E, Reid-Smith RJ, Taboada EN, Smith BA. 2019. Integrating Whole-Genome Sequencing Data Into Quantitative Risk Assessment of Foodborne Antimicrobial Resistance: A Review of Opportunities and Challenges. Front Microbiol 10.

67. Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, Wertheim HFL. 2017. Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clin Microbiol Rev 30:1015–1063.

CHAPTER 2

Genomic insight of *Enterococcus faecium* UC7251, a multi-drug resistance strain from ready-to-eat foods, highlights the risk of antimicrobial resistance in the food chain.

Chapter 2

Genomic insight of *Enterococcus faecium* UC7251, a multi-drug resistance strain from ready-to-eat foods, highlights the risk of antimicrobial resistance in the food chain.

Mireya Viviana Belloso Daza[1], Giovanni Milani[1], Claudia Cortimiglia[1], Ester Pietta[1], Daniela Bassi[1], Pier Sandro Cocconcelli[1]*

[1] Dipartimento di Scienze e Tecnologie Alimentari per una Filiera Agro-Alimentare Sostenibile (DISTAS), Università Cattolica del Sacro Cuore, Piacenza, Italy

Corresponding author: Pier Sandro Cocconcelli: pier.cocconcelli@unicatt.it

Keywords: *Enterococcus faecium,* multi-drug resistant, ready-to-eat foods, genomic analysis, gene transfer

This article was published on Frontiers in Microbiology in June 2022 DOI: 10.3389/fmicb.2022.894241.

1 Abstract

The presence of multi-drug resistant (MDR) bacteria in ready-to-eat foods comprises a threat for the public health due to their ability to acquire and transfer antibiotic-resistant determinants that could settle in the microbiome of the human digestive tract. In this study, *Enterococcus faecium* UC7251 isolated from a fermented dry sausage was characterized phenotypically and genotypically to hold resistance to multiple antibiotics including aminoglycosides, macrolides, β-lactams and tetracyclines. We furtherly investigated this strain following a hybrid sequencing and assembly approach (short and long reads) and determined the presence of various mobile genetic elements (MGE) responsible of horizontal gene transfer (HGT). On the chromosome of UC7251, we found one Integrative Conjugative Element (ICE) and a conjugative transposon Tn*916* carrying tetracycline resistance. UC7251 carries two plasmids, one small plasmid harboring a rolling circle replication and one MDR megaplasmid. The latter was identified as mobilizable and containing a putative integrative conjugative element-like region, prophage sequences, insertion sequences, heavy-metal resistance genes and several antimicrobial resistance (AMR) genes confirming the phenotypic resistance characteristics. The transmissibility potential of AMR markers was observed through mating experiments where Tn*916*-carried tetracycline resistance was transferred at intra- and inter-species level. This work highlights the significance of constant monitoring of products of animal origin, especially RTE foodstuffs, to stimulate the development of novel strategies in the race for constraining the spread of antibiotic resistance.

2 Introduction

Enterococcus faecium is an ubiquitous species found in a large number of foods, mainly fermented products of animal origin like cheeses and fermented sausages (1). Some strains of this species have been also recognized as probiotics conferring benefits to their hosts (2). Nevertheless, in the past three decades, *E. faecium* emerged as an important nosocomial multi-drug resistant (MDR) pathogen responsible for hospital-acquired infections (3). The duality of this species has led the European Food Safety Authority (EFSA) to state a safety assessment scheme based on the absence of genetic markers generally present in the hospital-associated (HA) biotypes for those *E. faecium* strains that are intentionally introduced into the food chain (4). Although *E. faecium* is extensively used as a probiotic and as part of the fermentation processes, it does not actually hold the Qualified Presumption of Safety (QPS) status due to its potential pathogenicity (5).

Previous studies indicated that the population structure of *E. faecium* is divided into three distinct clades. Clade A1 bearing clinical isolates, clade A2 mainly represented by strains from animal and human commensals that might cause sporadic human infections, both carrying determinants for virulence and antimicrobial resistance (AMR). The third, clade B, is characterized by community associated (CA) isolates lacking HA traits (6). Recently, clade B isolates were proposed to be reclassified as *Enterococcus lactis* because of the closer genomic proximity

to this new species and lack of HA markers (7). The genetic transmission of HA markers among isolates, for instance between farm animals and humans in the agricultural setting, revealed consequently the contamination of products of animal origin that affect the entire production and supply chain (8). The rise of MDR enterococci in the food chain represents a major public health concern as they are easily disseminated through the environment (9). Livestock animals and the farm environment exemplify an important reservoir of AMR bacteria due to the widely use of antibiotics (10), particularly in swine for prophylactic reasons (11). Also, resistance to heavy metals is a matter of concern because of possible co-selection of antibiotic resistance. Specifically, resistance towards copper is common in swine derived isolates due to the use of copper sulfate as a growth promoter in feed for pigs (12). Enterococci harboring MDR genes have been frequently isolated from the swine samples (13) and their diffusion arises concerns about the potential transmission to meat-based ready-to eat (RTE) foods, which proposes a risk because of the lack of microbial inactivation prior consumption (14). Considering the emergence of MDR enterococci and HA isolates, the current criteria for safety assessment is represented by a MIC of ampicillin of ≤ 2mg/L and lack of IS*16*/*esp/hyl* genes, associated with plasticity, adhesion, and carbohydrate metabolism, respectively (4). Further information on epidemiology and population structure can be analyzed by applying the Multi Locus Sequence Typing (MLST) scheme. Following this, *E. faecium* can be classified in different sequence types (STs), where ST17 was identified as the ancestral clone of HA isolates, forming the Clonal Complex 17 (CC17) (15). Nonetheless, it is crucial to understand the distribution of other putative virulence markers (PVM), involved in colonization and resistance recognized in other studies (3, 16). Horizontal gene transfer (HGT) is one the mechanism at the base of AMR and virulence markers dissemination among bacteria that, facilitates their survival and adaptation in stressful conditions. HGT of AMR genes between *E. faecium* and other species has been investigated mostly in clinical settings; furthermore, gene exchange in food was also demonstrated (14). Additionally, the transfer of resistance towards linezolid (17), oxazolinodone (18), aminoglycosides (19), glycopeptides, erythromycin and tetracycline (20) has been demonstrated between food isolated strains. The detection of AMR has also reached the retail level with the presence of AMR dissemination in RTE foods such as dairy products (21), salads (22), seafood (23) and meat products (24), pork-origin included (25).

Whole Genome Sequencing (WGS) has facilitated the understanding of the mechanisms that support the dissemination of mobile genetic elements (MGEs) in bacteria. The aim of this study is to investigate the genomic characteristics of a Vancomycin-susceptible MDR (VSE-MDR) *E. faecium* strain isolated from ready-to-eat fermented sausage and to evaluate the potential transmissibility of AMR markers through MGEs.

3 Materials and methods

3.1 Bacterial strain, cultivation, and antibiotic susceptibility testing

The strain UC7251 was isolated from a dry fermented Italian salami on Slanetz & Bartley Medium (Oxoid) containing 4 µg/ml Ampicillin (Sigma). The strain was sub-cultivated in Brain Heart Infusion (Oxoid) overnight at 37°C and species-specific PCR using primers for the *ddl* gene (Table 1S) was performed to confirm its taxonomical classification. Susceptibility to different antibiotics was determined by broth microdilution method according to EUCAST (26). The antimicrobial agents used were ampicillin, vancomycin, gentamycin, kanamycin, streptomycin, erythromycin, clindamycin, tylosine, tetracycline and chloramphenicol. The antibiotics were obtained from Sigma (St. Louis, Missouri, USA). The minimum inhibitory concentrations (MICs) were compared to the breakpoints recommended by EUCAST (2003) (http://www.eucast.org/) and EFSA (4).

3.2 Heavy metal susceptibility testing

Susceptibility towards Copper (Cu), Zinc (Zn), Cadmium (Cd) and Mercury (Hg) was tested as previously described (27, 28). Briefly, overnight cultures were spotted onto Mueller-Hinton agar (Oxoid) supplemented with different concentrations (0.05 to 40 mM) of $ZnCl_2$ (Carlo Erba), $HgCl_2$ (Sigma Aldrich) and $CdSO_4$ (Sigma Aldrich) resuspended in distilled water, and CuSO4 (Merck Millipore) adjusted to pH 7.2 with 1 M NaOH. After 24 to 48 h of incubation at 37°C, the plates were visually inspected for bacterial growth on the spots.

3.3 Conjugal Transfer

In vitro conjugation experiments were performed as described before (29). UC7251 was used as donor strain and 29 bacterial strains as recipients (see Table 3S). Briefly, 1 ml of a culture (OD600= 0.8) of donor and recipient strains were passed through a 0.45µm filter (MF-Millipore Membrane Filters, Merck). Right after, the filter was placed onto a non-selective agar plates favoring the growth of recipient strains and incubated at 37ºC for 24h. Conjugation with *Bacillus, Enterococcus, Listeria, Pseudomonas,* and *Staphylococcus* as recipient strains, were carried out onto BHI (Oxoid), *Clostridium* on RCM (Oxoid), for lactobacilli, *Pedioccoccus* and *Weisella* onto MRS (Difco). After the respective incubation period, cells were resuspended from the filter using saline solution and were diluted in a 10-fold dilution series and enumerated by spread plating onto appropriate agar media. Transconjugant selection was performed using the selective conditions reported in Table 3S. Transconjugant colonies were randomly selected and analyzed to check the presence or absence of the antibiotic resistance genes, by extracting the DNA with microLYSIS kit (Microzone) and performing PCR with primers for tetracycline and erythromycin resistance genes (Table 1S). The passage of potential plasmid-borne antibiotic resistance genes coding for aminoglycosides (*aad6, aph3-IIIa, aadE, satA, ant(6)-Ia)* and lincosamides (*IsaE, LnuB*) resistance, was also tested by PCR using the primers listed on Table 1S.

3.4 Multilocus sequence typing (MLST) analysis

Allelic profiles and sequence types were derived by PubMLST (30). The obtained ST were analysed using Phyloviz and the goeBURST algorithm to compute a spanning forest graph to build the relatedness between isolates based on Single Locus Variants (SLV) to identify clonal complexes (31). Furthermore, given that the resolution of MLST is limited, cgMLST (core genome MLST) was also determined using the cgmlst.org website. This method uses an allele numbering system for a scheme of 1423 cgMLST target genes, which confers a higher level of discrimination (32).

3.5 Detection of markers relevant for the assessment of safety and antibiotic resistance determinants

The strain UC7251 was screened for the hospital-associated genetic markers IS*16*, *hylEfm* and *esp* by PCR, using primers previously listed in the related section. Strains U0317 and E980 were used respectively as positive and negative controls. The presence of the antibiotic resistance determinants coding for the phenotypical resistances observed in UC7251 was investigated by PCR using the primers reported in Table 1S. The complete *pbp*5 gene was amplified, sequenced, and analyzed as described before (33), while the amplification of *ermB, tetM, tetL, aph3-IIIa ,satA ,ant(6)-Ia* and *aadE* was performed as described elsewhere (34–39). Here, new primers aad6_F and aad6_R for *aad6* screening, Lnu-B_F and Lnu-B_R for *Lnu(B)* screening, IsaE_F and lsaE_R for *Isa(E)* screening were designed de novo using Primer3 (40), and run the amplification reaction with the following conditions: initial denaturation at 95°C for 2 min; 35 cycles at 94°C for 40 s, 53°C for 45 s, and 72°C for 50 s; and extension at 72°C for 5 min.

3.6 Genome sequencing and database submission

A hybrid sequencing approach (short and long read) was followed to complete the assembly of UC7251. Genomic DNA was extracted from the cultured bacterium with NucleoSpin Tissue (Macherey-Nagel, Germany). Short read resequencing was performed with Illumina Miseq, 250 paired-end after Nextera XT paired-end library preparation. Long read sequencing was performed with PacBio Sequel II SMRT sequencing. After trimming the sequences using trimgalore! (GitHub - FelixKrueger/TrimGalore), hybrid assembly was carried out using Unicycler (41). The finished genome was deposited on NCBI under assembly accession No. ASM41165v2.

3.7 Bioinformatic Analyses

A total of 74 *E. faecium* complete genomes, including reference strains were selected to carry out phylogenetic and taxonomic analyses in comparison with UC7251 (Table 2S). Assembled genomes were downloaded from NCBI in September 2021 and were subsequently annotated using Prokka (42). Annotation results were then submitted to pan- and core-genome analysis using Roary (43). The phylogenetic tree was constructed using

RAxML-NG, V1.0.0(44) and iTOL was used to visualize and organize the tree (45). The genomes were also submitted to digital DNA-DNA hybridization (dDDH) using the Genome-to-Genome Distance Calculator (GGDC) (46). Average Nucleotide Identity (ANI) analysis was performed using fastANI (47).

In silico investigation of UC7251 was performed using the bioinformatics software platform Geneious prime v. 10.1. The Basic Local Alignment Tool (BLAST) from NCBI was used to investigate the presence and identity of different genetic markers contributing AMR, VF and MGE. The genome was interrogated for the presence of AMR genes using the Comprehensive Antibiotic Resistance Database (CARD) (48) and ResFinder (49). Ampicillin resistance was studied by evaluating the allelic variation in the strain of interest, against the reference sequence for PBP5-S/R profiles. Virulence markers were investigated according to the latest guidelines of EFSA (4) using manual annotation, VirulenceFinder (50) and VFAnalyzer (51).

HGT determinants were analyzed through MobileElementFinder (52) and Island Viewer 4 (53). In addition, integrative and conjugative elements were predicted using ICEberg 2.0 (54), which detects the signature sequences of the integrative modules and conjugation modules based on the profile hidden Markov models (profile HMMs). The origin of transfer site (*oriT*) was determined with OriTFinder (55). Lastly, the genome was screened for the presence of sequences of phage origin with Prophage Hunter (56) and CRISPR-Cas sites using CRISPR-CasFinder (57).

4 Results and discussion

4.1 Isolation and characterization of MDR *E. faecium* UC7251 from RTE food

In the framework of risk assessment of MDR in ready to eat foods, UC7251 was isolated from a dry-fermented sausage at a count of 3 x 10^5 CFU g-1 and identified as *Enterococcus faecium* by species-specific amplification of the *ddl* gene. This strain was resistant to ampicillin, streptomycin, kanamycin, erythromycin, clindamycin, tylosine and tetracycline and presented a MIC higher than the cutoff values defined by EUCAST and EFSA (Table 1). PCR analyses, using a pool of primers pairs targeted to the most commonly AMR genes found in enterococci (Table 1S) identified the genetic determinants for these resistances. *E. faecium* UC7251 was identified as a MDR strain, and harbored genes coding for aminoglycoside modifying enzymes, three genes for macrolide resistance and two genes responsible for tetracycline resistance. Moreover, the sequence of the amplicon targeted to the penicillin binding protein 5 PBP5, involved in β-lactams resistance, demonstrated that this strain showed the *pbp5*-S1/R20 allelic profile, conferring resistance to ampicillin (58). *E. faecium* showed to be intrinsically resistant to low levels of ampicillin through cell wall synthesis protein complex PBP; *pbp5* is part of this operon and sequence variations allow to differentiate the two groups of *E. faecium* according to allelic profile and expression levels (33). Within the context of a study focusing on the detection of ampicillin resistant *E. faecium* in ready to eat

fermented foods, a strain that presented resistance towards ampicillin with a MIC value of 64 µg/ml and carried the hybrid allelic profile PBP5-S1/R20 is of concern for the consumers safety. It has been demonstrated that *pbp5* may spread through horizontal gene transfer and specifically that *pbp5* of resistant isolates was located on transferable chromosomal regions, which suggested its dissemination through the environment (59).

Table 1. Antimicrobial resistance genes and MIC values of strain UC7251, following the guidelines and cutoff values established by EFSA/EUCAST for the safety assessment of *E. faecium*.

Antibiotic Resistance	UC7251 (µg/ml)	EFSA Cut-off value (µg/ml)	EUCAST (µg/ml)	AMR gene
Ampicillin	64	2	4	*pbp5-S1/R20*
Vancomycin	1	4	4	-
Gentamycin	32	32	32	*aac(6')-Ii*
Kanamycin	>4096	1024	-	*aph(3')-III*
Streptomycin	>1024	128	128	*aad6, aadE*
Erythromycin	>512	4	4	*ermB, mrsC, sat4*
Clindamycin	>512	4	-	*ermB, InuB, IsaE*
Tylosine	>512	4	-	*ermB*
Tetracycline	128	4	4	*tetL, tetM*
Chloramphenicol	8	16	32	-

4.2 Whole Genome Sequence Analyses

UC7251 was submitted to genome sequencing following a hybrid approach using long and short read technology (GenBank assembly accession numbers for chromosome CP084886.1, plasmid pUC7251_1 CP084887.1, plasmid pUC7251_2 CP084888.1). The assembly of the genome of UC7251 built a total of 3 contigs, predicted as a 2,6 Mb chromosome and two plasmids, pUC7251_1 and UC7251_2 (192 kb and 1,9 kb, respectively). The presence of the two plasmids was also distinguished by total DNA extraction and Pulsed Field Gel Electrophoresis (PFGE) (data not shown). The annotation of UC7251 resulted in 2662 coding sequences (CDS), of which 27% are hypothetical proteins and 73% have known functional assignments. It also contained genes coding for 18 rRNAs (6 copies each of 23S rRNA, 16S rRNA, and 5s rRNA), 69 tRNAs and 1 tmRNA. Compositional analysis resulted in 17 genomic islands (GIs), three active prophage sequences and several VF and AMR genes distributed throughout the chromosome and plasmidome. Regarding mobile genetic elements, two mobile regions were predicted on the chromosome and one on pUC7251_1 (Table 4S).

pUC7251_1 is a mobilizable megaplasmid as predicted by Plascad. According to OriTfinder, the origin of replication is 39bp long and showed homology with *oriT*_pUB110.There are no predicted T4SS proteins and only

one T4CP protein on locustag UC7251_02595. The relaxase MobM is found on locustag UC7251_02679. Mobilizable plasmids carry their own *oriT* and relaxase gene but lack genes required for T4SS formation and can therefore be transferred to cells that carry elements encoding a compatible T4SS (60). This plasmid showed homology with plasmids pF88_1 (identity 83%), p17-318_1 (identity 83%) pE843-TC-299 (identity 82%) and pE843-171 (identity 80%). The first three are VSE-MDR plasmids carried by *E. faecium* strains of clade A2. These strains were isolated from environmental (pF88_1) and human samples (p17-318_1 and pE843-TC-299). The fourth plasmid pE843-171, is carried by *Enterococcus lactis* E843 and it is characterized as VSE-MDR (61). According to these results, pUC7251_1 holds unique traits, and although the prevalence of VSE-MDR is high, none of the results on BLAST showed VSE-MDR from food origin. UC7251_2 harbors a single open reading frame that codes for a rolling circle REP (rep14a). Small plasmid of such size was also found in other *E. faecium* isolates, making it a common genomic feature.

4.3 Phylogenomics and Population Structure show that foodborne UC7251 is neighboring HA isolates.

For phylogenomic evaluation, UC7251 was compared with the other selected 74 *E. faecium* genomes (fig. 1). The interrogation of the pangenome has been recently regarded as a useful tool for species delimitation based on identification of lineage- specific gene sets (62). Observing the distribution of core- and accessory- genomes of our analysis, isolates of clade A1 and to a smaller extent clade A2 have a high variability in their accessory genes. A highly variable accessory genome is conferred by the fact that *E. faecium* has an open pangenome and therefore a higher genomic diversity (6). The adaptation of *E. faecium* to specific environmental factors, such as antimicrobial pressure, have increased the genomic diversity through horizontal gene transfer, genome rearrangement and gene loss (63). Pan and core genome analysis uncovered an open pangenome, with a core-genome consisting of 9,5% and an accessory-genome of 90,5%. In this context, UC7251 contains 33 unique genes, mainly insertion sequences and hypothetical proteins located on the chromosome and on pUC7251_1. Transposases belonging to IS*3*, IS*30* and IS*256* families were detected as unique on both pUC7251_1 and chromosome. On the chromosome we found unique gene *arnB*, that catalyzes the conversion of UDP-4-keto-arabinose to UDP-4-amino-4-deoxy-L-arabinose. The modified arabinose is attached to lipid A and is required for resistance to polymyxin in Gram negative bacteria (64). Moreover, unique genes *epsM* and *epsL* coding for putative acetyltransferase and sugar transferase respectively, were detected. They are involved in the production of the exopolysaccharide (EPS) component of the extracellular matrix during biofilm formation (65). Gene *cbh_2,* cholollyglycine hydrolyse, catalyzes the de-conjugation of bile acids (66). In *Enterococcus*, bile salt hydrolase activity has a hypo-cholesterolemic effects on animal and human hosts, conferring probiotic properties (67).

Furthermore, all 75 genomes were subjected to dDDH and ANI for genomic distance calculations. Although dDDH and ANI have different computational methods and species threshold values (70% for dDDH and 96% for ANI), they showed consistent results, confirming the taxonomical identification of UC7251. Digital DDH showed that values among UC7251-Clade A1 strains varied from 82-91%, among UC7251-Clade A2 strains 87-100% and, among UC7251-Clade B/*E. lactis* strains 64-70%. Similarly, ANI computation showed that the UC7251 is closest to Clade A2 strains with values between 98-100%, whereas comparison with genomes from the remaining two clades was lower (UC7251- Clade A1: 98% and UC7251-Clade B/*E. lactis*: 94%) (Table 2S).

The population structure and location of UC7251 was also evaluated using MLST. The genome was submitted to PubMLST and it was assigned to ST673. The latter clusters together with clonal group of ST117, which is known to be a part of CC17 meroclone (fig.1) . Published data on PubMLST showed a unique isolate harboring ST673, that contains a strain from a non-hospitalized person collected in Spain in 2010. MLST global scheme shows that UC7251, as other isolates from animal origin, belonged to hospital associated clades (68). Thus, *E. faecium* from CC17 have been also previously recovered from swine, poultry and cow samples (69–71). The use of cgMLST, a clustering based on 1423 target genes of the core genome, indicated that UC7251 belonged to the unique cluster type CT745.

Subspeciation of *E. faecium* has been also studied considering the defense mechanisms against HGT, such as CRISPR-Cas systems and R-M systems (72). CRISPR-Cas systems constitute endogenous barriers to HGT and, as a consequence, the presence of increased MGEs is associated with the complete absence or partial sequences of CRISPR-Cas systems (73). This has been observed in UC7251, where no complete CRISPR-Cas systems were detected. Differently, UC7251 carries a type I R-M system with the allelic variations typical of clade A1 isolates, polymorphisms that are used for clade classification of *E. faecium* (74).

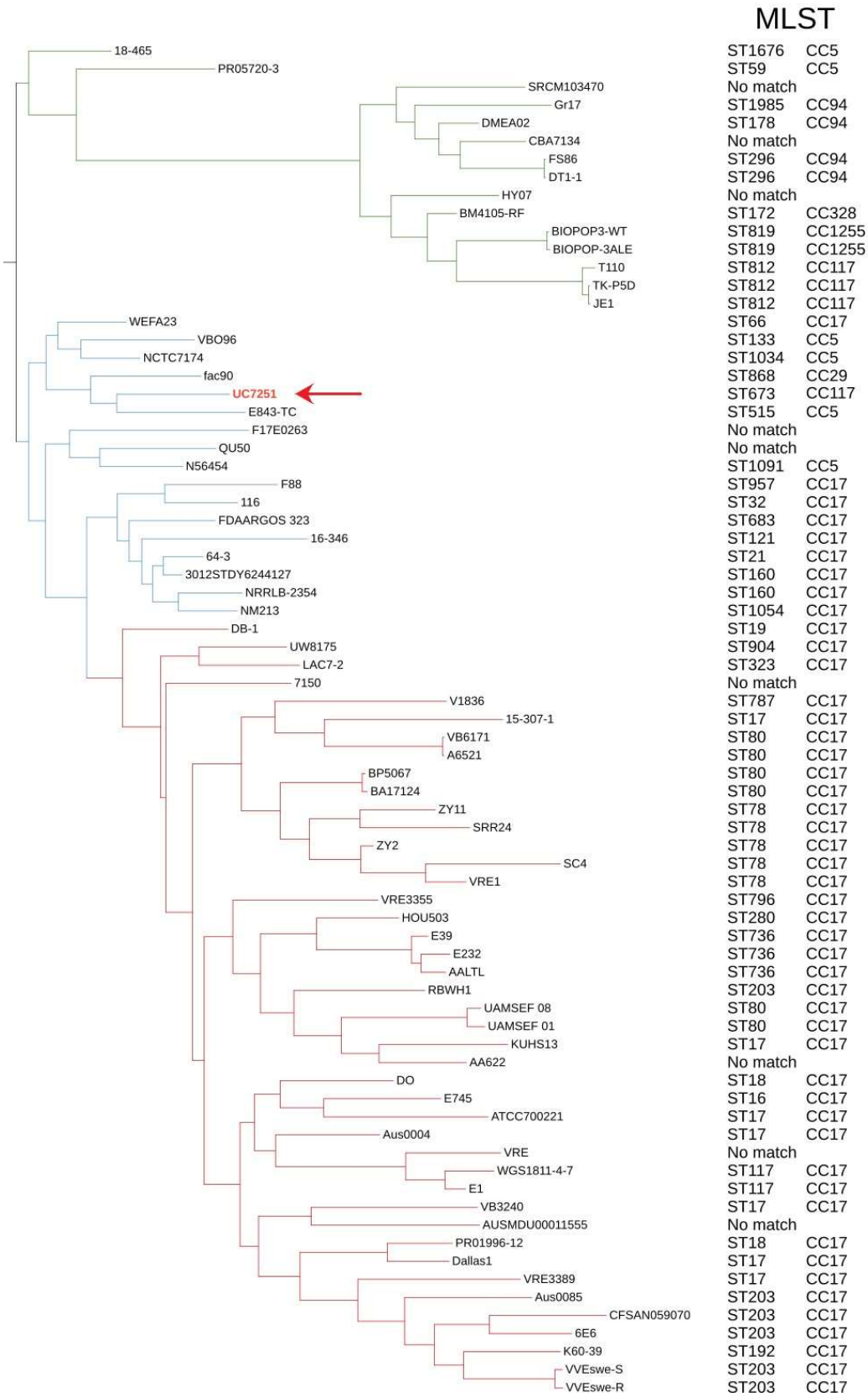

MLST
18-465 ST1676 CC5
PR05720-3 ST59 CC5
SRCM103470 No match
Gr17 ST1985 CC94
DMEA02 ST178 CC94
CBA7134 No match
FS86 ST296 CC94
DT1-1 ST296 CC94
HY07 No match
BM4105-RF ST172 CC328
BIOPOP3-WT ST819 CC1255
BIOPOP-3ALE ST819 CC1255
T110 ST812 CC117
TK-P5D ST812 CC117
JE1 ST812 CC117
WEFA23 ST66 CC17
VBO96 ST133 CC5
NCTC7174 ST1034 CC5
fac90 ST868 CC29
UC7251 ST673 CC117
E843-TC ST515 CC5
F17E0263 No match
QU50 No match
N56454 ST1091 CC5
F88 ST957 CC17
116 ST32 CC17
FDAARGOS 323 ST683 CC17
16-346 ST121 CC17
64-3 ST21 CC17
3012STDY6244127 ST160 CC17
NRRLB-2354 ST160 CC17
NM213 ST1054 CC17
DB-1 ST19 CC17
UW8175 ST904 CC17
LAC7-2 ST323 CC17
7150 No match
V1836 ST787 CC17
15-307-1 ST17 CC17
VB6171 ST80 CC17
A6521 ST80 CC17
BP5067 ST80 CC17
BA17124 ST80 CC17
ZY11 ST78 CC17
SRR24 ST78 CC17
ZY2 ST78 CC17
SC4 ST78 CC17
VRE1 ST78 CC17
VRE3355 ST796 CC17
HOU503 ST280 CC17
E39 ST736 CC17
E232 ST736 CC17
AALTL ST736 CC17
RBWH1 ST203 CC17
UAMSEF 08 ST80 CC17
UAMSEF 01 ST80 CC17
KUHS13 ST17 CC17
AA622 No match
DO ST18 CC17
E745 ST16 CC17
ATCC700221 ST17 CC17
Aus0004 ST17 CC17
VRE No match
WGS1811-4-7 ST117 CC17
E1 ST117 CC17
VB3240 ST17 CC17
AUSMDU00011555 No match
PR01996-12 ST18 CC17
Dallas1 ST17 CC17
VRE3389 ST17 CC17
Aus0085 ST203 CC17
CFSAN059070 ST203 CC17
6E6 ST203 CC17
K60-39 ST192 CC17
VVEswe-S ST203 CC17
VVEswe-R ST203 CC17

Figure 1. Maximum likelihood phylogenetic tree constructed using the core genome alignment of the selected 75 genomes and respective MLST and Clonal Complexes. *E. faecium* clade A1 strains are marked with red branches, clade A2 strains marked with blue and clade B strains with green. Strain UC7251 (pointing arrow and text in red) is grouped among clade A2 isolates and belongs to ST673 part of CC117.

4.4 Antimicrobial resistance profile and mobilome

Several AMR genes were detected on both the chromosome and pUC7251_1 (Table 2). The intrinsic determinant coding for aminoglycoside 6′-acetyltransferase enzyme (*aac(6')-II*), typical of *E. faecium* species (75), was found on the chromosome, together with the *liaFSR* operon, implicated in cell membrane-targeting lipopeptide antibiotic daptomycin (DAP) resistance. In previous studies, *E. faecium* isolates showed susceptibility and resistant allelic profiles of DAP (DAP-S and DAP-R, respectively) (76); UC7251 harbors the complete *liaFSR* system with the DAP-S allelic variation. Interestingly, occurrence of DAP resistance is inversely related to increased susceptibility to β-lactams, consistent with the ampicillin resistance in UC7251(77).

Genome sequencing and assembly following a hybrid approach, elucidated various details about UC7251 mobilome, crucial to understand the AMR mechanisms in this food isolated strain. A total of two mobile regions were predicted on the chromosome. Region 1 is classified as a putative integrative and mobilizable element (IME) with an insertion site and attachment sites and no detected origin of transfer (*oriT*). Proteins T2SSE, T4CP and VirB3, are also present within this region. T4CPs are phylogenetically and structurally associated to FtsK and SpoIIIE ATPases and the ability of translocating single-stranded DNA. Furthermore, type IV secretion protein VirB3 is an inner membrane protein and requires VirB4, VirB7, and VirB8 for stabilization (78). The IME contains several carbohydrate metabolism genes, suggesting acquired mechanisms for survival in environmental conditions. Region 2 is classified as an integrative conjugative element (ICE), including *oriT*, insertion and attachments sites. Additionally, it harbors T4SS machinery, integrase, relaxase and putative transposon Tn*916*. Tn*916* is a well described conjugative element that mediates tetracycline resistance (*tetM*) gene exchange principally among Gram positive bacteria (79). In the same molecule we found several inactive (score <0.50), 2 ambiguous (score 0.5-0.79) and one putatively active (score >0.80) prophage sequences, according to the scores attributed by Prophage Hunter software. The active prophage candidate showed the closest homology to *Halocynthia* phage JM-2012 (identity of 78%). This phage is classified as a "jumbo" bacteriophage from the *Myoviridae* family, initially identified within marine *Vibrio cyclitrophicus* (80). Limitations of the database of the phage prediction tool may interfere with estimation of the closest related phage. Phages from the *Myoviridae* family have been already identified in *Enterococcus* spp., making it a common feature within enterococci (81). Furthermore, pUC7251_1 harbors a large ICE-like (91 kb) with a total of 4 genomic islands (GIs) within (fig. 2). In detail, multiple AMR genes are found within GI2, GI3 and GI4 converting them in antibiotic resistance islands (ARI). ARI1 harbors against aminoglycoside coded by genes *aph(3')-IIIa*, *satA* and *ant(6')-Ia*, found from

UC7251_02667 to UC7251_02669. Interestingly, insertion sequence IS*1216E* is found flanking this region. IS*1216* is an enterococcal IS associated with resistance towards aminoglycosides, tetracyclines, macrolides and glycopeptides in Gram positive bacteria (82). IS*1216* has been identified as a vector for inter-plasmid recombination and dissemination of multi-drug elements in enterococci. Moreover, it has been found that IS*1216* is responsible for passing vancomycin-resistance with the help of transposon Tn*1546* (83) . Erythromycin resistance coded by *ermB* is found next to this region and adjacently the gene *tetL* is found next to relaxase MobM and the origin of transfer. It was reported that MobM has a dual role in autoregulation and initiation of transfer of plasmids or integrative mobilizable elements to other MGE members (84). Contiguous to this section, linezolid resistance genes *LnuB* and *IsaE* are found flanked by IS*Efa5*, which is typically found with high copy number in *E. faecium* strains. According to a previous study, it is suggested that IS*Efa5* may be contributing significantly to the genomic flexibility of the species with evidence of frequent integration and excision events (85). Additional studies have investigated that, plasmids harboring linezolid resistance genes acquired from other enterococcal plasmids through MGE, are associated with MDR phenotype (86, 87). Conjugation of these genes was also evidenced from *Enterococcus* to *Staphylococcus* elucidating their transmission potential (88). ARI3 carries two genes coding for aminoglycoside resistance *ant1, apt_3* and *ant(6)-Ia*, flanked by a putative recombinase and IS*4* family transposase IS*Dha5*. IS*4* family transposases are typically found among important clinical lineages in *E. faecium* (89). The presence of the complete operon for bacitracin resistance *bcrABDR* was found inside ARI4 flanked by IS*1485*. This is congruent with other studies suggesting the presence of this operon in swine isolates, as it is used for prophylaxis and therapy in food animals. The plasmid co-location of this locus and other resistance gene clusters might accelerate their dissemination (90).

Regarding prophage sequences, pUC7251_1 presented 2 sequence fragments with high homology (score > 0.8) to known prophage sequences, classified as active. These prophages show a high identity with *Staphylococcus* phage SPbeta-like and *Streptococcus* phage phiJH1301-2 prophage sequences, genetic elements that are common in clade A1 isolates (91). The annotated genes for both prophage code mainly for transposases and integrases as well as RelE/ParE toxin/antitoxin systems but. Additionally, 3 ambiguous and 4 inactive prophage sequences where detected. Interestingly, linezolid resistance genes *IsaE, LnuB* and bacitracin resistance operon *bcrABDR* were predicted to be within inactive prophage sequences. Recent studies have elucidated the role of phages in HGT of AMR genes as they are often carried within prophage sequences and stably inherited in the host genome carrying antibiotic resistant determinants (92). Inactive or defective phages, although categorized as non-functional, still may carry out important activities and functions like transposition and excision (93).

Copper resistance operon *tcrYAZB* was also found on an inactive prophage sequence. The swine industry is well-known for using copper as a feed additive. The presence of heavy metal resistance genes is a matter of concern also because of possible co-resistance with antibiotics (94, 95). The genetic system of this phenotypic resistance is coded by the *tcrYAZB* operon, which enhances bacterial survival and plasmid maintenance against high

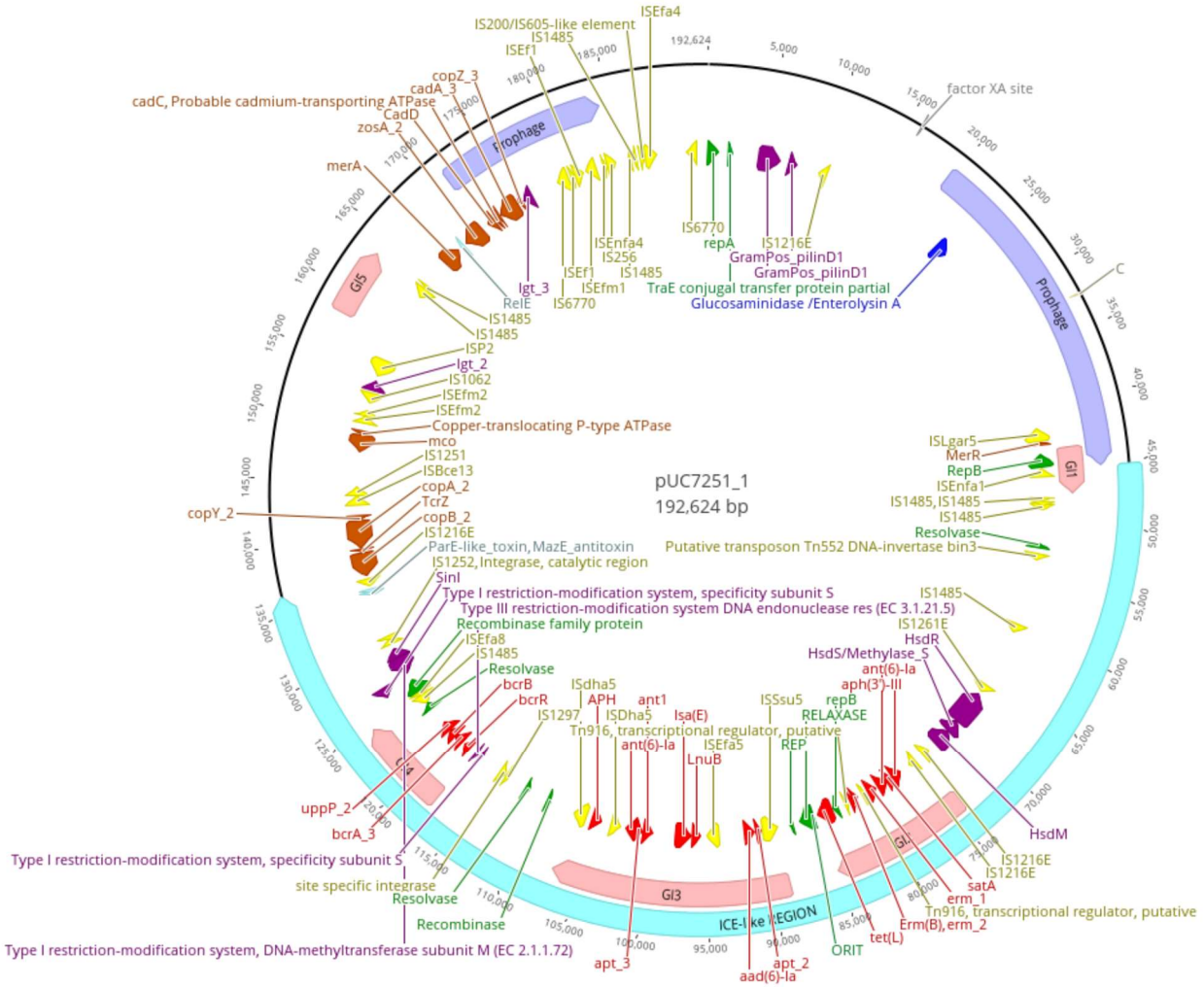
concentrations of this heavy metal (96). After genomic identification of copper resistance, we performed a susceptibility test to determine the MIC value, which resulted in 16mM, a level typical of high copper resistance. The mobility of this operon was evidenced by flanking IS*1216E* and IS*1251*, which are highly associated within *vanA*-type plasmids (97). A study by Silviera et al, (98) elucidated the presence of copper resistance genes with co-ocurrence of antibiotic resistance genes. Other heavy metal resistance genes were found in pUC7251_1. Zinc chloride, oxide or sulphate compounds are currently approved in the EU (up to 2500ppm) and used as additives in piglet feed (99). Similarly to copper, some regulatory genes and resistance mechanisms of Cu with known links to antibiotic resistance are also zinc-responsive (100). UC7251 has a MIC to Zn of 16mM and harbors gene *zosA*, observed also in *B. subtilis* for facilitating homeostasis to Zn (101). Moreover, resistance to Cd was also determined phenotypically at 2mM and genotypically by identifying genes *cadA*, *cadC* and *cadD*. The *cad* operons are typically found in *Staphylococci* carried by MGEs like plasmids and chromosomal cassettes and co-located with antibiotic resistance genes. Specifically, *cadAC* is known for conferring high resistance levels to Cd and it has been also reported to be present in other bacteria like *Lactococcus lactis, Streptococcus* spp. and *Listeria monocytogenes* (101). Furthermore, resistance to Hg was conferred by two genes *merA* and *merR*. In general, resistance Hg, is given by a set of genes clustered at the *mer* operon (*merRTPADE*) and are highly linked to AMR genes and MGEs (101). The lack of the complete operon is also confirmed by the low levels of resistance determined by MIC testing, that was of 50µM. Genes coding for resistance to Cu, Cd, Zn and Hg carried by plasmids, has been already observed in the plasmidome of other pig isolates (102).

Table 2. Distribution of virulence factors and AMR genes including antibiotic and heavy metal resistance genes in UC7251.

Molecule	Mechanism	Gene	Locus tag or position	Product
Chromosome	Antibiotic	*AAC(6')-Ia*	UC7251_02097	Aminoglycoside N(6')-acetyltransferase (EC 2.3.1.82)
		EfmM	UC7251_02049	rRNA) methyltransferase
		liaFSR	UC7251_01795-UC7251_01797	DAP
		pbp5	UC7251_01265	penicillin binding protein 5
		tet(M)	UC7251_02367	tetracycline resistance
	Heavy Metals	*cadA_1*	UC7251_00274	Cadmium-transporting ATPase
		cadA_2	UC7251_00904	Cadmium, zinc and cobalt-transporting ATPase
		copA_1	UC7251_00909	putative copper-importing P-type ATPase A
		copB_1	UC7251_00910	Copper-exporting P-type ATPase B
		copY_1	UC7251_00907	Transcriptional repressor CopY
		copZ_1	UC7251_00275	Copper chaperone CopZ
		copZ_2	UC7251_00908	Copper chaperone CopZ
		cutC	UC7251_02237	Copper homeostasis protein CutC
		czcD	UC7251_01786	Cadmium, cobalt and zinc/H( )-K( ) antiporter
		fief	UC7251_01380	Ferrous-iron efflux pump FieF
		ftsH	UC7251_02411	Cell division-associated, ATP-dependent zinc metalloprotease FtsH
		ziaA	UC7251_01739	Zinc-transporting ATPase
		znuA	UC7251_02450	High-affinity zinc uptake system binding-protein ZnuA
		znuB	UC7251_02448	High-affinity zinc uptake system membrane protein ZnuB
		znuC	UC7251_02449	High-affinity zinc uptake system ATP-binding protein ZnuC
		zosA	UC7251_01471	Zinc-transporting ATPase
		zupT	UC7251_00019	Zinc transporter ZupT
		zur	UC7251_00846	Zinc-specific metallo-regulatory protein
	Virulence	*swpB*	UC7251_00118	small WxL protein B
		swpC	UC7251_00593	small WxL protein C
		swpA	UC7251_00718	small WxL protein A
		acm	UC7251_02106	cell-wall-anchored collagen adhesin, MSCRAMM
		sagA	UC7251_02425	secreted antigen A
		scm	UC7251_02536	second collagen adhesin, MSCRAMM
		efaA	UC7251_00462	adhesion associated protein
		BopD	UC7251_00373	maltose operon transcriptional repressor
		cpsA/uppS	UC7251_01047	Undecaprenyl diphosphate synthase uppS
		cpsB/cdsA	UC7251_01048	Phosphatidate cytidylyltransferase cdsA
		fms3	UC7251_00358	Efm surface protein 3 orf371 (PGC-4)
		fms12	UC7251_00496	Efm surface protein 12 orf1996 (PGC-4)

		ebpA	UC7251_00550	PGC-3: endocarditis- and bio- film-associated pili A (MSCRAMM)
		epbB	UC7251_00551	PGC-3: endocarditis- and bio- film-associated pili B (MSCRAMM)
		ebpC	UC7251_00552	PGC-3: endocarditis- and bio- film-associated pili C (MSCRAMM)
		srtC	UC7251_00553	sortase C
		fms6	UC7251_00720	Efm surface protein 6 LPXTG family cell surface proteinPGC-4)
		fms7	UC7251_01220	Efm surface protein 7 orf2356 (PGC-4)
		fms22	UC7251_01278	Efm surface protein 22 orf884 (PGC-4)
		yidC	UC7251_00884	inner memebrane protein translocase and chaperone
pUC7251_1	Antibiotic	*ant(6)-Ia*	UC7251_02669	Aminoglycoside 6-adenylyltransferase
		ant1	UC7251_02694	Streptomycin 3"-adenylyltransferase
		ant(6)-Ia	UC7251_02696	Aminoglycoside 6-nucleotidyltransferase
		aph	UC7251_02698	aminoglycoside phosphotransferase family protein
		Lnu(B)	UC7251_02689	lincosamide nucleotidyltransferase
		lsa(E)	UC7251_02690	ABC-F type ribosomal protection protein Lsa(E)
		tet(L)	UC7251_02678	tetracycline efflux MFS transporter Tet(L)
		satA	UC7251_02668	Streptothricin acetyltransferase A
		erm_1	UC7251_02671	rRNA adenine N-6-methyltransferase
		erm_2	UC7251_02674	rRNA adenine N-6-methyltransferase
		aad(6)-Ia	UC7251_02684	Aminoglycoside 6-adenylyltransferase
	Heavy	*copZ_3*	UC7251_02781	Copper chaperone CopZ
	Metals	*cadA*	UC7251_02780	Cadmium, zinc and cobalt-transporting ATPase
		cadC	UC7251_02779	Cadmium, zinc and cobalt-transporting ATPase
		cadD	UC7251_02778	Cadmium, zinc and cobalt-transporting ATPase
		copA_2	UC7251_02740	Copper-exporting P-type ATPase
		copB_2	UC7251_02739	Copper-exporting P-type ATPase B
		copY_2	UC7251_02742	Transcriptional repressor CopY
		mco	UC7251_02750	Multicopper oxidase mco
		merA	UC7251_02772	mercuric reductase
		merR1	UC7251_02771	Mercuric resistance operon regulatory protein
		TcrZ	UC7251_02740	copper chaperone
		zosA	UC7251_02776	Zinc-transporting ATPase
	Virulence	*lgt*	UC7251_02756, UC7251_02782	surface protein anchor
		fms20	UC7251_02583-UC7251_02588	PGC-1: surface protein 20
		fms21 or pilA	UC7251_02583-UC7251_02588	PGC-1: surface protein 21

A


pUC7251_1
192,624 bp
Prophage
GI1
GI2
GI3
GI4
GI5
ICE-like REGION
ISEfa4
IS200/IS605-like element
IS1485
ISEf1
copZ_3
cadA_3
cadC, Probable cadmium-transporting ATPase
CadD
zosA_2
merA
factor XA site
IS6770
repA
IS1216E
GramPos_pilinD1
GramPos_pilinD1
TraE conjugal transfer protein partial
Glucosaminidase /Enterolysin A
ISEnfa4
IS256
ISEf1
IS1485
ISEfm1
Igt_3
IS6770
RelE
IS1485
IS1485
ISP2
Igt_2
IS1062
ISEfm2
ISEfm2
Copper-translocating P-type ATPase
mco
IS1251
ISBce13
copA_2
TcrZ
copB_2
IS1216E
copY_2
ParE-like_toxin, MazE_antitoxin
IS1252, Integrase, catalytic region
SinI
Type I restriction-modification system, specificity subunit S
Type III restriction-modification system DNA endonuclease res (EC 3.1.21.5)
Recombinase family protein
ISEfa8
IS1485
Resolvase
bcrB
bcrR
ISdha5
APH
ant1
IS1297
ISDha5
Isa(E)
Tn916, transcriptional regulator, putative
ant(6)-Ia
ISEfa5
LnuB
uppP_2
bcrA_3
Type I restriction-modification system, specificity subunit S
site specific integrase
Resolvase
Recombinase
Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)
apt_3
aad(6)-Ia
apt_2
ORIT
tet(L)
Erm(B),erm_2
Tn916, transcriptional regulator, putative
satA
erm_1
IS1216E
IS1216E
HsdM
HsdR
HsdS/Methylase_S
ant(6)-Ia
aph(3'')-III
ISSsu5
repB
RELAXASE
REP
IS1485
IS1261E
HsdR
HsdS/Methylase_S
ant(6)-Ia
aph(3'')-III
ISSsu5
repB
RELAXASE
REP
ISLgar5
MerR
RepB
ISEnfa1
IS1485, IS1485
IS1485
Resolvase
Putative transposon Tn552 DNA-invertase bin3

B


pUC7251_1 extraction
ICE-like REGION
Prophage
GI1
GI2
GI3
GI4
ISLgar5
MerR
RepB
ISEnfa1
IS1485
Resolvase
Putative transposon Tn552 DNA-invertase bin3
IS1261E
HsdR
HsdS/Methylase_S
HsdM
IS1216E
IS1216E
aph(...
ant(6)...
erm_1
tet(L)
RELAXASE
ISSsu5
aad(6...
ISEfa5
site specific integrase
apt_3
ant(6)-Ia
ant1
ISDha5
APH
ISdha5
Recombinase
Resolvase
IS1297
LnuB
Isa(E)
bcrR
bcrA_3
bcrB
uppP...
Type ...
Re...
IS1485
Recombin...
Type I restricti...
SinI
IS1...
copB_2
copA_2
ISBc...
IS1...
mco
ISE...

Figure 2. A. Map of plasmid pUC7251_1 harboring one large containing prophage sequences (mauve), an integrative conjugative-like element (cyan), five genomic islands (pink), insertion sequences (yellow), antibiotic resistance genes (red), metal resistance genes (orange), virulence factors (dark violet) and replication initiation systems (green), toxin-antitoxin systems (light blue). B. Details of the ICE-like region. The genetic elements are indicated with the color code above mentioned.

4.5 Conjugation experiments suggest AMR gene transfer by Insertion Sequences

The MDR profile of UC7251 endorsed the further evaluation of transmissibility of AMR genes and we focused on tetracycline resistance coded by two genes on Tn*916* and pUC7251_1 and the plasmid encoded erythromycin resistance. This was tested through conjugation experiments where gene exchange was demonstrated at inter- and intra-generic level (Table 3 and Table 3S). Filter mating experiments demonstrated that tetracycline resistance was transferred from UC7251 to *E. faecalis* OG1rf, *L. innocua* L7, *L. monocytogenes* DSM 15675, *S. aureus* UC7180, *L. rhamnosus* UC8647, with frequencies of transconjugants per donors varying from 6 x 10-3 to 5,7 x 10-6 CFU/ml. No gene transfer was observed toward Gram negative species. The transfer of the *tetM* gene was confirmed by PCR assays, whereas *tetL* was absent in all tetracycline positive transconjugants. The transfer of the *tetM* gene was found to be carried by chromosomal transposon Tn*916* from *E. faecalis*. This operon was predicted in chromosomal locus UC7251_02362-02376. The nucleotide identity between the 18,032 bp sequence of Tn*916* of UC7251 and *E. faecalis* (Genbank Accession No. U09422.1) sequences, was of 99.97%. It has been discovered that the presence of subinhibitory concentrations of specific classes of antibiotics can trigger the mobility of Tn*916*, as it has a broad inducibility of antibiotic resistance genes, implying that the dissemination of resistance genes is not necessarily linked to their selective pressure (103).

No gene transfer for the genes coding for erythromycin resistance was observed, consistently with the characteristics of pUC7251_1, a mobilizable but non conjugative plasmid lacking the complete conjugation apparatus.

Table 3. Conjugation of tetracycline resistance between *E. faecium* UC7251 and strains from other genera.

Donor	**Recipient Strain**	**Conjugation Frequency (T/D)**	**PCR Confirmation**	
			tetM	***tetL***
E. faecium UC7251	*E.s faecalis* OG1rf	6.01E-03	+	-
	L. innocua L7	5.68E-06	+	-
	L. monocytogenes DSM 15675	8.38E-04	+	-
	S. aureus UC7180	3.78E-02	+	-
	L. rhamnosus UC8647	6.84E-05	+	-

4.6 Virulence markers in the UC7251 genome show a collection of colonization facilitators

The complete assembly and annotation of UC7251 genome allowed investigation of the presence of putative virulence markers (Table 2). Adherence is an essential step in bacterial pathogenesis, required for colonization and attachment and it is therefore considered a type of virulence marker. When scrutinizing the genome of UC7251, several microbial surface components, recognizing adhesive matrix molecules (MSCRAMMs) including LPXTG family cell wall-anchored surface proteins as well as fimbriae proteins such as pili, were identified. It is important to denote the presence of genes *acm* (cell-wall-anchored collagen adhesin) and *scm* (second collagen adhesin). These proteins enhance initial adherence in vivo and interact with extracellular matrix components. Other genes associated with adhesion, *efaA* (*E. faecium* surface protein) and *sagA* (secreted antigen A) were detected. A novel class of cell surface proteins coded by WxL operon, found in clade A *E. faecium* isolates, with a functional role in virulence associated with endocarditis pathogenesis and bile salt resistance was previously investigated (104). The coding genes *swpA* (small WxL protein A), *swpB* (small WxL protein B) and *swpC* (small WxL protein C) were found in UC7251. Additionally, *malR*, a maltose-binding transcriptional regulator that increases biofilm production in the presence of this specific carbohydrate, was detected. Pili associated proteins, previously described as Pilin Gene Clusters (PGC-1, PGC-2, PGC-3, PGC-4), (16),were identified. PGC-1 is composed by the genes *fms20* and *fms21*; both are present along with a sortase A. This loci/operon is located between UC7251_02853 and UC7251_02588 in pUC7251_1. In addition, PGC-3 was found with 100% of nucleotide identity containing the endocarditis and biofilm associated pili genes *ebpA, ebpB*, *ebpC* accompanied by *srtC* (sortase) and flanked by IS*1216E*. This region is encompassed from UC7251_02583 to UC7251_02589 in the chromosome. PGC-4 cluster is incomplete lacking operon fms11-19-16 and PGC-2 associated genes fms14-17-13. UC7251 does not express the capsular polysaccharide, presenting the capsule operon polymorphism CPS type 1 (105) and does not harbor cytolysin (106) and BoNT/En toxin, a botulin type toxin found in a single strain of *E. faecium* (107). *E. faecium* UC7251 lacks the putative HA virulence markers as defined by EFSA (4) and does not harbor the complete operons coding pili-associated proteins, which is typical of clade A1 isolates.

5 Conclusion

The presence of multi-drug resistant strains in ready-to-eat fermented food represents a risk of public health for the spread of AMR determinants in the food chain and in the gut microbiota of consumers. In silico bioinformatic evaluations derived from genomic data permitted to accurately assess the safety of UC7251, a strain of *E. faecium* clade A2 which does not carry virulence factors typical of HA strains but presents the co-location of several antimicrobial resistance genes with heavy metal resistances on the mobilizable plasmid pUC7251_1 and the conjugative transposon Tn*916*. This work emphasizes the importance of a surveillance for the presence of AMR bacteria in food, with particular attention to fermented RTE foods. Moreover, the

presence of MDR strains carrying mobile AMR genetic elements incites the development of innovative strategies for the mitigation of the risk related to antimicrobial resistance diffusion in food.

6 Transparency Declaration

All authors contributed to the article and approved the submitted version. This work was partially funded by the PRIMA programme, under BioProMedFood project (Reference Number: 2019-SECTION2-4; CUP: J34I19004820005). The PRIMA programme was supported by the European Union H2020 programme and innovation programme

7 References

1. Ben Braïek O, Morandi S, Cremonesi P, Smaoui S, Hani K, Ghrairi T. 2018. Biotechnological potential, probiotic and safety properties of newly isolated enterocin-producing Enterococcus lactis strains. LWT - Food Sci Technol 92:361–370.

2. Ghattargi VC, Gaikwad MA, Meti BS, Nimonkar YS, Dixit K, Prakash O, Shouche YS, Pawar SP, Dhotre DP. 2018. Comparative genome analysis reveals key genetic factors associated with probiotic property in Enterococcus faecium strains. BMC Genomics 19:1–16.

3. Gao W, Howden BP, Stinear TP. 2018. Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen. Curr Opin Microbiol. Elsevier Ltd https://doi.org/10.1016/j.mib.2017.11.030.

4. EFSA, Rychen G, Aquilina G, Azimonti G, Bampidis V, Bastos M de L, Bories G, Chesson A, Cocconcelli PS, Flachowsky G, Gropp J, Kolar B, Kouba M, López-Alonso M, López Puente S, Mantovani A, Mayo B, Ramos F, Saarela M, Villa RE, Wallace RJ, Wester P, Glandorf B, Herman L, Kärenlampi S, Aguilera J, Anguita M, Brozzi R, Galobart J. 2018. Guidance on the characterisation of microorganisms used as feed additives or as production organisms. EFSA J 16.

5. EFSA BIOHAZ Panel, Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Cocconcelli PS, Fernández Escámez PS, Maradona MP, Querol A, Suarez JE, Sundh I, Vlak J, Barizzone F, Correia S, Herman L. 2021. The list of QPS status recommended biological agents for safety risk assessments carried out by EFSA https://doi.org/10.5281/ZENODO.4428353.

6. Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, Corander J, Cheng L, Saif S, Young S, Zeng Q, Wortman J, Birren B, Willems RJL, Earl AM, Gilmore MS. 2013. Emergence

of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. MBio 4.

7. Belloso Daza MV, Cortimiglia C, Bassi D, Cocconcelli PS. 2021. Genome-based studies indicate that the Enterococcus faecium Clade B strains belong to Enterococcus lactis species and lack of the hospital infection associated markers. Int J Syst Evol Microbiol 71.

8. Manyi-Loh C, Mamphweli S, Meyer E, Okoh A. 2018. Antibiotic use in agriculture and its consequential resistance in environmental sources: Potential public health implications. Molecules. MDPI AG https://doi.org/10.3390/molecules23040795.

9. Serwecińska L. 2020. Antimicrobials and Antibiotic-Resistant Bacteria: A Risk to the Environment and to Public Health. Water 12:3313.

10. Koutsoumanis K, Allende A, Álvarez-Ordóñez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, Ru G, Simmons M, Skandamis P, Suffredini E, Argüello H, Berendonk T, Cavaco LM, Gaze W, Schmitt H, Topp E, Guerra B, Liébana E, Stella P, Peixe L. 2021. Role played by the environment in the emergence and spread of antimicrobial resistance (AMR) through the food chain. EFSA J 19.

11. Pholwat S, Pongpan T, Chinli R, Rogawski McQuade ET, Thaipisuttikul I, Ratanakorn P, Liu J, Taniuchi M, Houpt ER, Foongladda S. 2020. Antimicrobial Resistance in Swine Fecal Specimens Across Different Farm Management Systems. Front Microbiol 11:1238.

12. Yu Z, Gunn L, Wall P, Fanning S. 2017. Antimicrobial resistance and its association with tolerance to heavy metals in agriculture production. Food Microbiol 64:23–32.

13. Tan SC, Chong CW, Teh CSJ, Ooi PT, Thong KL. 2018. Occurrence of virulent multidrug- resistant Enterococcus faecalis and Enterococcus faecium in the pigs, farmers and farm environments in Malaysia. PeerJ 2018.

14. Chajęcka-Wierzchowska W, Zadernowska A, Zarzecka U, Zakrzewski A, Gajewska J. 2019. Enterococci from ready-to-eat food - horizontal gene transfer of antibiotic resistance genes and genotypic characterization by PCR melting profile. J Sci Food Agric 99:1172–1179.

15. Lee T, Pang S, Abraham S, Coombs GW. 2019. Antimicrobial-resistant CC17 Enterococcus faecium: The past, the present and the future. J Glob Antimicrob Resist 16:36–47.

16. Freitas AR, Tedim AP, Novais C, Coque TM, Peixe L. 2018. Distribution of putative virulence markers in Enterococcus faecium: towards a safety profile review. J Antimicrob Chemother 73:306–319.

17. Tyson GH, Sabo JL, Hoffmann M, Hsu C-H, Mukherjee S, Hernandez J, Tillman G, Wasilenko JL, Haro J, Simmons M, Wilson Egbe W, White PL, Dessai U, Mcdermott PF. 2018. Novel linezolid resistance plasmids in Enterococcus from food animals in the USA. J Antimicrob Chemother 73:3254–

3258.

18. Kang Z-Z, Lei C-W, Kong L-H, Wang Y-L, Ye X-L, Ma B-H, Wang X-C, Li C, Zhang Y, Wang H-N. 2019. Detection of transferable oxazolidinone resistance determinants in Enterococcus faecalis and Enterococcus faecium of swine origin in Sichuan Province, China. J Glob Antimicrob Resist 19:333–337.

19. Kim Y Bin, Seo KW, Son SH, Noh EB, Lee YJ. 2019. Genetic characterization of high-level aminoglycoside-resistant Enterococcus faecalis and Enterococcus faecium isolated from retail chicken meat. Poult Sci 98:5981–5988.

20. Conwell M, Daniels V, Naughton PJ, Dooley JSGG. 2017. Interspecies transfer of vancomycin, erythromycin and tetracycline resistance among Enterococcus species recovered from agrarian sources. BMC Microbiol 2017 171 17:1–8.

21. Chajęcka-Wierzchowska W, Zadernowska A, García-Solache M. 2020. Ready-to-eat dairy products as a source of multidrug-resistant Enterococcus strains: Phenotypic and genotypic characteristics. J Dairy Sci 103:4068–4077.

22. Zhou SYD, Wei MY, Giles M, Neilson R, Zheng F, Zhang Q, Zhu YG, Yang XR. 2020. Prevalence of Antibiotic Resistome in Ready-to-Eat Salad. Front Public Heal 8:92.

23. Igbinosa EO, Beshiru A. 2019. Antimicrobial Resistance, Virulence Determinants, and Biofilm Formation of Enterococcus Species From Ready-to-Eat Seafood. Front Microbiol 10.

24. Chajęcka-Wierzchowska W, Zadernowska A, Łaniewska-Trokenheim Ł. 2016. Diversity of Antibiotic Resistance Genes in *Enterococcus* Strains Isolated from Ready-to-Eat Meat Products. J Food Sci 81:M2799–M2807.

25. Kim HJ, Koo M. 2020. Diversity of Enterococcus faecium in Processed Pork Meat Products in Korea. Foods 9:1–14.

26. The European Committee on Antimicrobial Susceptibility Testing. 2017. Breakpoint tables for interpretation of MICs and zone diameters, version 7.0, 2017. http://www.eucast.org. http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_5.0_Breakpoint_Table_01.pdf.

27. Capps KM, Amachawadi RG, Menegat MB, Woodworth JC, Perryman K, Tokach MD, Dritz SS, Derouchey JM, Goodband RD, Bai J, Apley MD, Lubbers B V, Nagaraja TG. 2020. Impact of added copper, alone or in combination with chlortetracycline, on growth performance and antimicrobial resistance of fecal enterococci of weaned piglets. J Anim Sci 98:1–11.

28. Sharifi Y, Abedzadeh A, Salighe A, Kalhor N, Motlagh MK, Javadi A. 2015. Antibiotics and heavy

metals resistance patterns of Enterococcus faecalis and faecium bacteria isolated from the human and the livestock sources. Environ Heal Eng Manag J 2:199–202.

29. Cocconcelli PS, Cattivelli D, Gazzola S. 2003. Gene transfer of vancomycin and tetracycline resistances among Enterococcus faecalis during cheese and sausage fermentations, p. 315–323. *In* International Journal of Food Microbiology. Elsevier.

30. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; referees: 2 approved]. Wellcome Open Res 3.

31. Francisco AP, Vaz C, Monteiro PT, Melo-Cristino J, Ramirez M, Carriço JA. 2012. PHYLOViZ: Phylogenetic inference and data visualization for sequence based typing methods. BMC Bioinformatics 13.

32. De Been M, Pinholt M, Top J, Bletz S, Mellmann A, Van Schaik W, Brouwer E, Rogers M, Kraat Y, Bonten M, Corander J, Westh H, Harmsen D, Willems RJL. 2015. Core genome multilocus sequence typing scheme for high-resolution typing of enterococcus faecium. J Clin Microbiol 53:3788–3797.

33. Pietta E, Montealegre MC, Roh JH, Cocconcelli PS, Murray BE. 2014. Enterococcus faecium PBP5-S/R, the Missing Link between PBP5-S and PBP5-R. Antimicrob Agents Chemother 58:6978–6981.

34. Jacob J, Evers S, Bischoff K, Carlier C, Courvalin P. 1994. Characterization of the sat4 gene encoding a streptothricin acetyltransferase in Campylobacter coli BE/G4. FEMS Microbiol Lett 120:13–17.

35. Swenson JM, Ferraro MJ, Sahm DF, Clark NC, Culver DH, Tenover FC, Charache P, Harrell LJ, Reller LB, Hardy D, Moellering RC, Wilson W, Hindler J. 1995. Multilaboratory evaluation of screening methods for detection of high-level aminoglycoside resistance in enterococci. National Committee for Clinical Laboratory Standards Study Group on Enterococci. J Clin Microbiol 33:3008–3018.

36. Olsvik B, Olsen I, Tenover FC. 1995. Detection of tet(M) and tet(Q) using the polymerase chain reaction in bacteria isolated from patients with periodontal disease. Oral Microbiol Immunol 10:87–92.

37. Sutcliffe J, Grebe T, Tait-Kamradt A, Wondrack L. 1996. Detection of erythromycin-resistant determinants by PCR. Antimicrob Agents Chemother 40:2562–2566.

38. Trzcinski K, Cooper BS, Hryniewicz W, Dowson CG. 2000. Expression of resistance to tetracyclines in strains of methicillin-resistant Staphylococcus aureus. J Antimicrob Chemother 45:763–770.

39. Ouoba LII, Lei V, Jensen LB. 2008. Resistance of potential probiotic lactic acid bacteria and bifidobacteria of African and European origin to antimicrobials: Determination and transferability of the resistance genes to other bacteria. Int J Food Microbiol 121:217–224.

40. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115.

41. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol 13:e1005595.

42. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

43. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

44. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455.

45. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. Nucleic Acids Res 47:W256–W259.

46. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60.

47. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:1–8.

48. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG. 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

49. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 75:3491–3500.

50. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol 52:1501–1510.

51. Liu B, Zheng D, Jin Q, Chen L, Yang J. 2018. VFDB 2019: a comparative pathogenomic platform with

an interactive web interface. Nucleic Acids Res 47:687–692.

52. Johansson MHKK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN. 2021. Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. J Antimicrob Chemother 76:101–109.

53. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, Brinkman FSL. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:W30.

54. Liu M, Li X, Xie Y, Bi D, Sun J, Li J, Tai C, Deng Z, Ou HY. 2019. ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. Nucleic Acids Res 47:D660–D665.

55. Li X, Xie Y, Liu M, Tai C, Sun J, Deng Z, Ou HY. 2018. OriTfinder: A web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. Nucleic Acids Res 46:W229–W234.

56. Song W, Sun HX, Zhang C, Cheng L, Peng Y, Deng Z, Wang D, Wang Y, Hu M, Liu W, Yang H, Shen Y, Li J, You L, Xiao M. 2019. Prophage Hunter: an integrative hunting tool for active prophages. Nucleic Acids Res 47:W74–W80.

57. Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, Rocha EPC, Vergnaud G, Gautheret D, Pourcel C. 2018. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res 46:W246–W251.

58. Galloway-Peña JR, Rice LB, Murray BE. 2011. Analysis of PBP5 of early U.S. isolates of Enterococcus faecium: Sequence variation alone does not explain increasing ampicillin resistance over time. Antimicrob Agents Chemother 55:3272–3277.

59. Morroni G, Brenciani A, Litta-Mulondo A, Vignaroli C, Mangiaterra G, Fioriti S, Citterio B, Cirioni O, Giovanetti E, Biavasco F. 2019. Characterization of a new transferable MDR plasmid carrying the pbp5 gene from a clade B commensal Enterococcus faecium. J Antimicrob Chemother 74:843–850.

60. Guédon G, Libante V, Coluzzi C, Payot S, Leblond-Bourget N. 2017. The Obscure World of Integrative and Mobilizable Elements, Highly Widespread Elements that Pirate Bacterial Conjugative Systems. Genes (Basel) 8.

61. Shan X, Yang M, Wang N, Schwarz S, Li D, Du X-D. 2022. Plasmid Fusion and Recombination Events That Occurred during Conjugation of poxtA-Carrying Plasmids in Enterococci. Microbiol Spectr 10.

62. Moldovan MA, Gelfand MS. 2018. Pangenomic definition of prokaryotic species and the phylogenetic structure of Prochlorococcus spp. Front Microbiol 9:428.

63. Bonacina J, Suárez N, Hormigo R, Fadda S, Lechner M, Saavedra L. 2017. A genomic view of food-related and probiotic Enterococcus strains. DNA Res An Int J Rapid Publ Reports Genes Genomes 24:11.

64. Lee M, Sousa MC. 2014. Structural Basis for Substrate Specificity in ArnB. A Key Enzyme in the Polymyxin Resistance Pathway of Gram-Negative Bacteria. Biochemistry 53:796–805.

65. Agius JE, Phalen DN, Rose K, Eden JS. 2021. Genomic Insights Into the Pathogenicity of a Novel Biofilm-Forming Enterococcus sp. Bacteria (Enterococcus lacertideformus) Identified in Reptiles. Front Microbiol 12:389.

66. Chand D, Panigrahi P, Varshney N, Ramasamy S, Suresh CG. 2018. Structure and function of a highly active Bile Salt Hydrolase (BSH) from Enterococcus faecalis and post-translational processing of BSH enzymes. Biochim Biophys Acta - Proteins Proteomics 1866:507–518.

67. Singhal N, Maurya AK, Mohanty S, Kumar M, Virdi JS. 2019. Evaluation of bile salt hydrolases, cholesterol-lowering capabilities, and probiotic potential of enterococcus faecium isolated from rhizosphere. Front Microbiol 10:1567.

68. Gouliouris T, Raven KE, Ludden C, Blane B, Corander J, Horner CS, Hernandez-Garcia J, Wood P, Hadjirin NF, Radakovic M, Holmes MA, de Goffau M, Brown NM, Parkhill J, Peacock SJ. 2018. Genomic surveillance of enterococcus faecium reveals limited sharing of strains and resistance genes between livestock and humans in the United Kingdom. MBio 9.

69. Freitas AR, Coque TM, Novais C, Hammerum AM, Lester CH, Zervos MJ, Donabedian S, Jensen LB, Francia MV, Baquero F, Peixe L. 2011. Human and Swine Hosts Share Vancomycin-Resistant Enterococcus faecium CC17 and CC5 and Enterococcus faecalis CC2 Clonal Clusters Harboring Tn1546 on Indistinguishable Plasmids †. J Clin Microbiol 49:925–931.

70. Werner G, Fleige C, Feßler AT, Timke M, Kostrzewa M, Zischka M, Peters T, Kaspar H, Schwarz S. 2012. Improved identification including MALDI-TOF mass spectrometry analysis of group D streptococci from bovine mastitis and subsequent molecular characterization of corresponding Enterococcus faecalis and Enterococcus faecium isolates. Vet Microbiol 160:162–169.

71. Getachew Y, Hassan L, Zakaria Z, Abdul Aziz S. 2013. Genetic variability of vancomycin-resistant enterococcus faecium and enterococcus faecalis isolates from humans, chickens, and pigs in malaysia. Appl Environ Microbiol 79:4528–4533.

72. Koonin E V., Makarova KS, Wolf YI. 2017. Evolutionary Genomics of Defense Systems in Archaea and Bacteria. Annu Rev Microbiol 71:233.

73. dos Santos BA, de Oliveira J da S, Parmanhani-da-Silva BM, Ribeiro RL, Teixeira LM, Neves FPG.

2020. CRISPR elements and their association with antimicrobial resistance and virulence genes among vancomycin-resistant and vancomycin-susceptible enterococci recovered from human and food sources. Infect Genet Evol 80:104183.

74. Huo W, Adams HM, Trejo C, Badia R, Palmer KL. 2019. A Type I Restriction-Modification System Associated with Enterococcus faecium Subspecies Separation. Appl Environ Microbiol 85.

75. Chow JW. 2000. Aminoglycoside resistance in enterococci. Clin Infect Dis 31:586–589.

76. Panesso D, Reyes J, Gaston EP, Deal M, Londoño A, Nigo M, Munita JM, Miller WR, Shamoo Y, Tran TT, Arias CA. 2015. Deletion of liaR reverses daptomycin resistance in Enterococcus faecium independent of the genetic background. Antimicrob Agents Chemother 59:7327–7334.

77. Diaz L, Tran TT, Munita JM, Miller WR, Rincon S, Carvajal LP, Wollam A, Reyes J, Panesso D, Rojas NL, Shamoo Y, Murray BE, Weinstock GM, Arias CA. 2014. Whole-Genome Analyses of Enterococcus faecium Isolates with Diverse Daptomycin MICs. Antimicrob Agents Chemother 58:4527.

78. Álvarez-Rodríguez I, Arana L, Ugarte-Uribe B, Gómez-Rubio E, Martín-Santamaría S, Garbisu C, Alkorta I. 2020. Type IV Coupling Proteins as Potential Targets to Control the Dissemination of Antibiotic Resistance. Front Mol Biosci 7:201.

79. Devirgiliis C, Coppola D, Barile S, Colonna B, Perozzi G. 2009. Characterization of the Tn916 conjugative transposon in a food-borne strain of Lactobacillus paracasei. Appl Environ Microbiol 75:3866–3871.

80. Lavysh D, Sokolova M, Minakhin L, Yakunina M, Artamonova T, Kozyavkin S, Makarova KS, Koonin E V., Severinov K. 2016. The genome of AR9, a giant transducing Bacillus phage encoding two multisubunit RNA polymerases. Virology 495:185–196.

81. Duerkop BA, Palmer KL, Horsburgh MJ. 2014. Enterococcal Bacteriophages and Genome DefenseEnterococci: From Commensals to Leading Causes of Drug Resistant Infection. Massachusetts Eye and Ear Infirmary. https://www.ncbi.nlm.nih.gov/books/NBK190419/. Retrieved 31 May 2022.

82. Partridge SR, Kwong SM, Firth N, Jensen SO. 2018. Mobile genetic elements associated with antimicrobial resistance. Clin Microbiol Rev 31.

83. Lin Y-T, Tseng S-P, Hung W-W, Chang C-C, Chen Y-H, Jao Y-T, Chen Y-H, Teng L-J, Hung W-C. 2020. A Possible Role of Insertion Sequence IS1216V in Dissemination of Multidrug-Resistant Elements MESPM1 and MES6272-2 between Enterococcus and ST59 Staphylococcus aureus. Microorganisms 8:1–12.

84. Lorenzo-Díaz F, Fernández-López C, Guillén-Guío B, Bravo A, Espinosa M. 2018. Relaxase MobM

induces a molecular switch at its cognate origin of transfer. Front Mol Biosci 5:17.

85. Bayjanov JR, Baan J, Rogers MRC, Troelstra A, Willems RJL, van Schaik W. 2019. Enterococcus faecium genome dynamics during long-term asymptomatic patient gut colonization. Microb Genomics 5.

86. Sadowy E. 2018. Linezolid resistance genes and genetic elements enhancing their dissemination in enterococci and streptococci. Plasmid 99:89–98.

87. Elghaieb H, Tedim AP, Abbassi MS, Novais C, Duarte B, Hassen A, Peixe L, Freitas AR. 2020. From farm to fork: identical clones and Tn6674-like elements in linezolid-resistant Enterococcus faecalis from food-producing animals and retail meat. J Antimicrob Chemother 75:30–35.

88. Yan XM, Wang J, Tao XX, Jia HB, Meng FL, Yang H, You YH, Zheng B, Hu Y, Bu XX, Zhang JZ. 2021. A Conjugative MDR pMG1-Like Plasmid Carrying the lsa(E) Gene of Enterococcus faecium With Potential Transmission to Staphylococcus aureus. Front Microbiol 12.

89. Mikalsen T, Pedersen T, Willems R, Coque TM, Werner G, Sadowy E, Van Schaik W, Jensen LB, Sundsfjord A, Hegstad K. 2015. Investigating the mobilome in clinically important lineages of enterococcus faecium and enterococcus faecalis. BMC Genomics 16:1–16.

90. Wang X-M, Li X-S, Wang Y-B, Wei F-S, Zhang S-M, Shang Y-H, Du X-D. 2015. Characterization of a multidrug resistance plasmid from Enterococcus faecium that harbours a mobilized bcrABDR locus. J Antimicrob Chemother 70:609–632.

91. Lisotto P, Raangs EC, Couto N, Rosema S, Lokate M, Zhou X, Friedrich AW, Rossen JWA, Harmsen HJM, Bathoorn E, Chlebowicz-Fliss MA. 2021. Long-read sequencing-based in silico phage typing of vancomycin-resistant Enterococcus faecium. BMC Genomics 22.

92. Kondo K, Kawano M, Sugai M. 2021. Distribution of Antimicrobial Resistance and Virulence Genes within the Prophage-Associated Regions in Nosocomial Pathogens. mSphere 6.

93. Mitchell S. 2014. ZOMBIES IN BACTERIAL GENOMES: IDENTIFICATION AND ANALYSIS OF PREVIOUSLY VIRULENT PHAGE. ZOMBIES Bact GENOMES Identif Anal PREVIOUSLY VIRULENT PHAGE.

94. Ahmed MO, Baptiste KE. 2018. Vancomycin-Resistant Enterococci: A Review of Antimicrobial Resistance Mechanisms and Perspectives of Human and Animal Health. Microb Drug Resist 24:590–606.

95. EFSA FEEDAP Panel. 2016. Revision of the currently authorised maximum copper content in complete feed. EFSA J 14.

96. Rebelo A, Mourão J, Freitas AR, Duarte B, Silveira E, Sanchez-Valenzuela A, Almeida A, Baquero F, Coque TM, Peixe L, Antunes P, Novais C. 2021. Diversity of metal and antibiotic resistance genes in Enterococcus spp. from the last century reflects multiple pollution and genetic exchange among phyla from overlapping ecosystems. Sci Total Environ 787:147548.

97. Wongnak K, Pattanachaiwit S, Rattanasirirat W, Limsrivanichakorn S, Kiratisin P, Assanasen S, Leelaporn A. 2021. First characterization of Tn1546-like structures of vancomycin-resistant Enterococcus faecium Thai isolates. J Infect Chemother 27:991–998.

98. Silveira E, Freitas AR, Antunes P, Barros M, Campos J, Coque TM, Peixe L, Novais C. 2014. Co-transfer of resistance to high concentrations of copper and first-line antibiotics among Enterococcus from different origins (humans, animals, the environment and foods) and clonal lineages. J Antimicrob Chemother 69:899–906.

99. Murphy D, Ricci A, Auce Z, Beechinor JG, Bergendahl H, Breathnach R, Bureš J, Duarte Da Silva JP, Hederová J, Hekman P, Ibrahim C, Kozhuharov E, Kulcsár G, Lander Persson E, Lenhardsson JM, Mačiulskis P, Malemis I, Markus-Cizelj L, Michaelidou-Patsia A, Nevalainen M, Pasquali P, Rouby J, Schefferlie J, Schlumbohm W, Schmit M, Spiteri S, Srčič S, Taban L, Tiirats T, Urbain B, Vestergaard E, Wachnik-Święcicka A, Weeks J, Zemann B, Allende A, Bolton D, Chemaly M, Fernandez Escamez PS, Girones R, Herman L, Koutsoumanis K, Lindqvist R, Nørrung B, Robertson L, Ru G, Sanaa M, Simmons M, Skandamis P, Snary E, Speybroeck N, Ter Kuile B, Wahlström H, Baptiste K, Catry B, Cocconcelli PS, Davies R, Ducrot C, Friis C, Jungersen G, More S, Muñoz Madero C, Sanders P, Bos M, Kunsagi Z, Torren Edo J, Brozzi R, Candiani D, Guerra B, Liebana E, Stella P, Threlfall J, Jukes H. 2017. EMA and EFSA Joint Scientific Opinion on measures to reduce the need to use antimicrobial agents in animal husbandry in the European Union, and the resulting impacts on food safety (RONAFA). EFSA J 15.

100. Poole K. 2017. At the Nexus of Antibiotics and Metals: The Impact of Cu and Zn on Antibiotic Activity and Resistance. Trends Microbiol 25:820–832.

101. Argudín MA, Hoefer A, Butaye P. 2019. Heavy metal resistance in bacteria from animals. Res Vet Sci 122:132–147.

102. Wist V, Morach M, Schneeberger M, Cernela N, Stevens MJA, Zurfluh K, Stephan R, Nüesch-Inderbinen M. 2020. Phenotypic and genotypic traits of vancomycin-resistant enterococci from healthy food-producing animals. Microorganisms 8.

103. Scornec H, Bellanger X, Guilloteau H, Groshenry G, Merlin C. 2017. Inducibility of Tn916 conjugative transfer in Enterococcus faecalis by subinhibitory concentrations of ribosome-targeting antibiotics. J Antimicrob Chemother 72:2722–2728.

104. Galloway-Peña JR, Liang X, Singh K V., Yadav P, Chang C, La Rosa SL, Shelburne S, Ton-That H, Höök M, Murray BE. 2015. The identification and functional characterization of WxL proteins from Enterococcus faecium reveal surface proteins involved in extracellular matrix interactions. J Bacteriol 197:882–892.

105. Hufnagel M, Hancock LE, Koch S, Theilacker C, Gilmore MS, Huebner J. 2004. Serological and Genetic Diversity of Capsular Polysaccharides in Enterococcus faecalis. J Clin Microbiol 42:2548.

106. Top J, Willems R, Bonten M. 2008. Emergence of CC17 *Enterococcus faecium* : from commensal to hospital-adapted pathogen. FEMS Immunol Med Microbiol 52:297–308.

107. Zhang S, Lebreton F, Mansfield MJ, Miyashita SI, Zhang J, Schwartzman JA, Tao L, Masuyer G, Martínez-Carranza M, Stenmark P, Gilmore MS, Doxey AC, Dong M. 2018. Identification of a botulinum neurotoxin-like toxin in a commensal strain of Enterococcus faecium. Cell Host Microbe 23:169.

8 Supplementary materials

Table 1S. Primers used in this work

Scope	Primer name	Primer sequence (5'-3')	Reference
Species identification	EM1A (ddl gene)	TTGAGGCAGACCAGATTGACG	Cheng et al. 1997
	EM1B (ddl gene)	TATGACAGCGACTCCGATTC	
Screening of HA markers	esp14F	AGATTTCATCTTTGATTCTTGG	Leavis et al., 2003
	esp12R	AATTGATTCTTTAGCATCTGG	
	hylEfm-F	GAGTAGAGGAATATCTTAGC	Rice et al., 2003
	hylEfm-R	AGGCTCCAATTCTGT	
	IS16-F	CATGTTCCACGAACCAGAG	Werner et al., 2011
	IS16-R	TCAAAAAGTGGGCTTGGC	
AMR genes	Efmpbp5-1outsideF	GGAATGACAAGCAAGAGAAGGAGG	Galloway-Pena et al., 2011
	Efmpbp5-1F	ATGAAAAGAAGTGACAAGCACGGC	
	Efmpbp5-1R	GCAAAGATGAATACCTCATTAGG	
	Efmpbp5-2F	CAAAGTAATCGGGTTGTACCCAGC	
	Efmpbp5-2R	GTCCCACGAAGATCCTTATCAAAAGCC	
	Efmpbp5-3F	GGCTTTTGATAAGGATCTTCGTGGGAC	
	Efmpbp5-3R	CCCATTTTCAACGTTTCTTGTGCC	
	Efmpbp5-4F	GGCACAAGAAACGTTGAAAATGGG	
	Efmpbp5-4R	TTATTGATAATTTTGGTTGAGGTATTG	
	Efmpbp5-4outsideR	CGCCACAGTCCTTTTACTGTAC	
	Rpbp5_1F	GCAAAGATGAATACCTCATTAGG	
	Rpbp5_1R	CAAAGTAATCGGGTTGTACCCAGC	
	Rpbp5_2F	CAGAACTTCCAGCTGGAGCTAC	
	Rpbp5_2R	GATCATAGCTTGGAGAGCTAGC	
	Rpbp5_3F	GCGACAGGTTATGCTCCTGG	
	Rpbp5_3R	GAATACATTGCTGCTTGCTGGATAGG	
	ermB1	GAAAAGGTACTCAACCAAATA	Sutcliffe et al., 1996
	ermB2	AGTAACGGTACTTAAATTGTTTAC	
	tetM1	GAACTCGAACAAGAGGAAAGC	Olsvik et al., 1995
	tetM2	ATGGAAGCCCAGAAAGGAT	
	tetL-up	ATAAATTGTTTCGGGTCGGTAAT	Trzcinski et al., 2000
	tetL-rev	AACCAGCCAACTAATGACAATGAT	
	aad6_F	TTCGAATTGTGACCCTTGAG	This study
	aad6_R	TGGTTCAGATGATCGATTGC	
	aph3-IIIa_F	GCCGATGTGGATTGCGAAAA	Ouoba et al., 2008
	aph3-IIIa_R	GCTTGATCCCCAGTAAGTCA	
	aadE_F	ATGGAATTATTCCCACCTGA	Ouoba et al., 2008
	aadE_R	TCAAAACCCCTATTAAAGCC	
	SatA_F	TCAAAGTTGGCGTATAA	Jacob et al., 1994
	SatA_R	TAAACCCAGCGAACCAT	
	Ant(6)-Ia_F	GCCTTTCCGCCACCTCACCG	Swenson et al., 1995
	Ant(6)-Ia_R	ACTGGCTTAATCAATTTGGG	
	Lnu-B_F	ATCGAGCAGTGGTCTTTGCA	This study
	Lnu-B_R	GGTTGTTTGACGTAGCTCCG	
	IsaE_F	TTGGCACGTTTCATCGCTTT	This study
	IsaE_R	ACGGACGCGGTAAAACTACT	

Table 2S. selected strains for taxonomic and phylogenetic analyses

Strain	Clade	Genbank Accesion	MLST	Clonal Complex according to MLST scheme and central ST	dDDH Average against UC7251	ANI against UC7251
7150	A1	GCA_019356355.1	No match		90.53	98.93
15-307-1	A1	GCA_002973755.2	17	CC17	82.73	98.57
6E6	A1	GCA_001518735.1	203	CC203	82.40	98.58
A6521	A1	GCA_012933195.2	80	CC80	84.63	98.60
AA622	A1	GCA_019977575.1	No match		86.83	98.78
AALTL	A1	GCA_002880635.1	736	CC17	87.27	98.58
ATCC700221	A1	GCA_001594345.1	17	CC17	86.80	98.89
Aus0004	A1	GCA_000250945.1	17	CC17	86.03	98.78
Aus0085	A1	GCA_000444405.1	203	CC203	82.27	98.67
AUSMDU00011555	A1	GCA_017301355.1	No match		82.73	98.54
BA17124	A1	GCA_012932975.2	80	CC80	85.50	99.04
BP5067	A1	GCA_012932985.2	80	CC80	85.60	99.00
CFSAN059070	A1	GCA_003071425.1	203	CC203	82.63	98.59
Dallas1	A1	GCA_015999605.1	17	CC17	83.73	98.51
DB-1	A1	GCA_006337045.1	19	CC64	87.20	98.92
DO	A1	GCA_000174395.2	18	CC18	87.33	98.80
E1	A1	GCA_001886635.1	117	CC117	83.27	98.62
E232	A1	GCA_002777275.1	736	CC17	86.43	98.58
E39	A1	GCA_001635875.1	736	CC17	86.50	98.60
E745	A1	GCA_001750885.1	16	CC17	84.17	98.84
HOU503	A1	GCA_005952885.1	280	CC280	87.90	98.65
K60-39	A1	GCA_002334625.1	192	CC192	85.60	98.71
KUHS13	A1	GCA_009938285.1	17	CC17	87.00	98.77
LAC7-2	A1	GCA_009036045.1	323	CC17	89.57	98.87
PR01996-12	A1	GCA_018219325.1	18	CC18	84.57	98.66
RBWH1	A1	GCA_003957785.1	203	CC203	85.17	98.66
SC4	A1	GCA_002848385.1	78	CC78	84.90	98.72
SRR24	A1	GCA_009734005.2	78	CC78	85.23	98.77
UAMSEF_01	A1	GCA_005886545.1	80	CC80	87.50	98.70
UAMSEF_08	A1	GCA_005886655.1	80	CC80	87.50	98.67
UW8175	A1	GCA_001587115.1	904	CC280	91.23	98.91
V1836	A1	GCA_008728455.1	787	CC80	85.43	98.75
VB3240	A1	GCA_005576735.1	17	CC17	85.73	98.55
VB6171	A1	GCA_017897965.1	80	CC80	84.63	98.60
VRE	A1	GCA_009697285.1	No match		82.70	98.69
VRE1	A1	GCA_006007925.1	78	CC78	83.27	98.63
VRE3355	A1	GCA_017584065.1	796	CC203	86.43	98.84
VRE3389	A1	GCA_015999405.1	17	CC17	85.43	98.71
VVEswe-R	A1	GCA_007917035.3	203	CC203	82.70	98.56
VVEswe-S	A1	GCA_007917315.3	203	CC203	82.87	98.58
WGS1811-4-7	A1	GCA_016864255.1	117	CC117	84.50	98.69
ZY11	A1	GCA_009938075.1	78	CC78	88.07	98.80
ZY2	A1	GCA_010120755.1	78	CC78	84.93	98.81

116	A2	GCA_018279145.1	32	CC32	88.03	98.90
16-346	A2	GCA_002761555.1	121	CC78	87.60	99.07
18-465	A2	GCA_018516845.1	1676	CC5	89.77	99.10
3012STDY6244127	A2	GCA_900683475.1	160	CC32	91.10	99.16
64/3	A2	GCA_001298485.1	21	CC32	89.43	98.99
E843-TC	A2	GCA_019774555.1	515	CC5	90.57	99.18
F17E0263	A2	GCA_006280355.1	No match		87.27	98.65
F88	A2	GCA_019175425.1	957	CC323/CC17	89.57	98.85
fac90	A2	GCA_016743855.1	868	CC29	92.47	98.91
FDAARGOS_323	A2	GCA_002983785.1	683	CC32	90.00	98.80
N56454	A2	GCA_006351845.1	1091	CC5	87.60	98.70
NCTC7174	A2	GCA_900637035.1	1034	CC5	91.87	99.21
NM213	A2	GCA_005166365.1	1054	CC32	89.10	98.90
NRRLB-2354	A2	GCA_001544255.1	160	CC32	89.60	99.05
QU 50	A2	GCA_006741355.1	No match		87.97	98.62
UC7251	A2	GCA_000411655.2	673	CC117	100.00	100.00
VBO96	A2	GCA_019456555.1	133	CC5	92.10	99.08
WEFA23	A2	GCA_002850515.1	66	CC66	91.87	99.11
BIOPOP-3 ALE	B	GCA_012045365.1	819	CC1255	69.07	94.37
BIOPOP3-WT	B	GCA_012045505.1	819	CC1255	69.07	94.34
BM4105-RF	B	GCA_003269465.1	172	CC328	64.63	94.25
CBA7134	B	GCA_004015145.1	No match		70.13	94.80
DMEA02	B	GCA_008330605.1	178	CC94	69.47	94.80
DT1-1	B	GCA_011745645.1	296	CC94	68.40	94.49
FS86	B	GCA_013201055.1	296	CC94	68.40	94.54
Gr17	B	GCA_003711605.1	1985	CC94	68.17	94.60
HY07	B	GCA_003574925.1	No match		68.47	94.48
JE1	B	GCA_003667965.1	812	CC117	67.70	94.66
SRCM103470	B	GCA_004103475.1	No match		69.30	94.83
T110	B	GCA_000737555.1	812	CC117	66.50	94.70
TK-P5D	B	GCA_015377765.1	812	CC117	67.70	94.65
PR05720-3	B	GCA_018219285.1	59	CC5	82.50	98.15

Table 3S. Strains used for mating experiments with tetracycline and erythromycin.

Conjugation			Tetracycline				Erythromycin		
			Conjugation frequency		PCR confirmation		Conjugation frequency		PCR confirmation
Donor	Recepient	Strain	(T/D)	(T/R)	*tetM*	*tetL*	(T/D)	(T/R)	*ermB*
E. faecium UC7251	*Enterococcus faecalis*	OG1rf	6.01E-03	2.80E-06	yes	no	/	/	no
	Listeria innocua	L7	5.68E-06	7.16E-05	yes	no	/	/	no
	Listeria monocytogenes	DSM 15675	8.38E-04	3.88E-03	yes	no	/	/	no
	Staphylococcus aureus	UC7180	3.78E-02	1.37E-01	yes	no	/	/	no
	Staphylococcus xylosus	UC8727	/	/	/	/	/	/	/
	Staphylococcus carnosus	UC8838	/	/	/	/	/	/	/
	Lactobacillus rhamnosus	UC8647	6.84E-05	1.30E-04	yes	no	/	/	/
	Lactobacillus paracasei	UC8477	/	/	/	/	/	/	/
	Lactobacillus casei	UC8477	/	/	/	/	/	/	/
	Lactobacillus fermentum	UC10045	/	/	/	/	/	/	/
	Lactobacillus plantarum	UC8479	/	/	/	/	/	/	/
	Lactobacillus reuteri	UC10043	/	/	/	/	/	/	/
	Bacillus cereus	UC4044	/	/	/	/	/	/	/
	Escherichia coli	BL21	/	/	/	/	/	/	/
	Pseudomonas koreensis	Psk	/	/	/	/	/	/	/
	Pseudomonas aeruginosa	Psa	/	/	/	/	/	/	/
	Pseudomonas chlororaphis	Psc	/	/	/	/	/	/	/
	Pseudomonas putida	Psp	/	/	/	/	/	/	/
	Pseudomonas fluorescens	Psf	/	/	/	/	/	/	/
	Weisella confusa	LMG 18478	/	/	/	/	/	/	/
	Weisella confusa	LMG17696	/	/	/	/	/	/	/
	Weisella confusa	LMG 17695	/	/	/	/	/	/	/
	Weisella confusa	BCC 2344	/	/	/	/	/	/	/
	Weisella confusa	BCC 3263	/	/	/	/	/	/	/
	Weisella confusa	BCC4255	/	/	/	/	/	/	/
	Weisella confusa	024F6	/	/	/	/	/	/	/
	Pediococcus pentosaceus	UC8487	/	/	/	/	/	/	/
	Pediococcus acidilactici	UC8715	/	/	/	/	/	/	/
	Clostridium tyrobutyricum	UC7086	/	/	/	/	/	/	/
	Clostridium sporogenes	UC9000	/	/	/	/	/	/	/

Table 4S. Annotation of ICE and IME in UC7251

	Chromosome				**pUC7251_1**	
	Region 1		**Region 2**		**Region 1**	
Location (nt)	1097029-1202552		2360857-2421852		45229..136214	
Length (bp)	105524		60996		90986	
GC content (%)	39.28		38.52		35.40	
oriT	-		2419860-2419992		85370..85407	
Insertion site	Aspartate racemase (1096652-1097383)		hypothetical protein (2421842-2422156)		-	
Direct Repeats	attL: 1097029-1097044 (aacagaaggaagtatg)/attR: 1202537-1202552 (aacagaaggaagtatg)		attL: 2360857-2360872 (tttctttattctttta)/attR: 2421837-2421852 (tttctttattctttta)		-	
Type	Putative IME		Putative ICE with T4SS		ICE-like region	
Annotations	Proteins	Position	Proteins	Position	Proteins	Position
	T2SSE	1119712-1120962	T2SSE	2384650-2386434	Integrase	4522-45909
	FtsK_SpoIIIE (T4CP)	1121045-1121632	TrbC	2393945-2394601	Pfam-B_706	47135-47335
	Pfam-B_6973 (Relaxase)	1123703-1124461	T2SSE	2396328-2398097	Integrase	48425-48928
	T2SSE	1140006-1141790	Integrase	2404634-2405851	Integrase	59358-60194
	VirB3	1181314-1181997	tet(M)	2408450-2410369	Integrase	61549-62634
			TrbL (Orf15_Tn, T4SS component)	2412673-2414850	Integrase	64313-64867
			AAA_10 (Orf16_Tn, T4SS component)	2414853-2417300	Integrase	65555-66235
			TrbC (Orf19_Tn, T4SS compont)	2418379-2418600	Integrase	73299-73979
			Rep_trans (Relaxase)	2418643-2419848	Integrase	74831-75511
			FtsK_SpoIIIE (T4CP)	2420026-2421411	Mob_Pre (Relaxase)	85476-86624
					Recombinase (Integrase)	107823-109391
					Recombinase (Integrase)	109393-109809
					Integrase	113755-114075
					Integrase	114194-114784
					Integrase	123820-124686
					Recombinase (Integrase)	125063-126289
					Integrase	135534-136214

CHAPTER 3

Genome-based studies indicate that the *Enterococcus faecium* Clade B strains belong to *Enterococcus lactis* species and lack of the hospital infection associated markers

Chapter 3

Genome-based studies indicate that the *Enterococcus faecium* Clade B strains belong to *Enterococcus lactis* species and lack of the hospital infection associated markers

Mireya Viviana Belloso Daza[1], Claudia Cortimiglia[1], Daniela Bassi[1]* and Pier Sandro Cocconcelli[1]

[1]Dipartimento di Scienze e Tecnologie Alimentari per una Filiera Agro-Alimentare Sostenibile (DISTAS), Università Cattolica del Sacro Cuore, 29122 Piacenza-Cremona, Italy.

Corresponding author: Pier Sandro Cocconcelli pier.cocconcelli@unicatt.it

Keywords: *Enterococcus faecium, Enterococcus lactis*, clade B, virulence markers, Whole-Genome based studies.

This article is published on the International Journal of Systematic and Evolutionary Microbiology in August 2021. DOI: 10.1099/ijsem.0.004948

1 Abstract

Enterococcus lactis and the heterotypic synonym *Enterococcus xinjiangensis* from dairy origin have recently been identified as a novel species based on 16S rRNA gene sequence analysis. *Enterococcus faecium* type strain NCTC 7171^T was used as the reference genome for determining *E. lactis, E. xinjiangensis* and *E. faecium* as separate species. However, this taxonomical classification did not consider the diverse lineages of *E. faecium*, and the double nature of hospital (clade A) and community-associated (clade B) isolates. Here, we investigate the taxonomical relationship among *E. faecium* isolates from different origin and *E. lactis*, using a genome-based approach. Additional to 16S rRNA gene sequence analysis, we estimated the relatedness among strains and species using phylogenomics based on the core pangenome, multilocus sequence typing, the average nucleotide identity and the digital DNA-DNA hybridization. Moreover, following the available safety assessment schemes, we evaluated the virulence profile and the ampicillin resistance of *E. lactis* and *E. faecium* clade B strains. Our results confirmed the genetic and evolutionary differences between clade A and the intertwined clade B and *E. lactis* group. We also confirmed the absence in these strains of virulence gene markers *IS16, hyl$_{Efm}$* and *esp* and the lack of PBP5 allelic profile associated to ampicillin resistance. Taken together, our findings support the reassignment of the strains of *E. faecium* clade B as *E. lactis*.

2 Introduction

Enterococci are gram-positive, catalase-negative, facultative anaerobic bacteria belonging to the lactic acid bacteria group. *Enterococcus faecium* is an important species of this genus, which is ubiquitous in nature. It is found in large numbers of foods, mainly fermented products of animal origin like cheeses and dry sausages (1), where they are part of the natural ripening processes or added as starter cultures. Some strains of this species have been recognized also as probiotics conferring benefits to their hosts (2,3). However, in the past three decades, *E. faecium* emerged as an important nosocomial multi-drug resistant pathogen responsible for hospital-acquired infections, such as endocarditis, urinary tract infections, and septicemia (4). The double nature of this species has led the European Food Safety Authority (EFSA) to define a safety assessment scheme for the *E. faecium* strains to be internationally introduced in the food chain, based on the absence of genetic markers generally present in the hospital-associated (HA) strains. Although *E. faecium* is widely used as a probiotic and as part of fermentation processes, it does not actually hold QPS status (5) due to its potential pathogenicity (6).

The population structure of *E. faecium* is divided into distinct clades, clade A containing the hospital-associated clade A and the community-associated (CA) clade B. A further split within clade A was identified and estimated to occur about 76 years ago (±30 years), namely clade A2, consisting mostly of animal-associated isolates, while clinical isolates are grouped in sub-clade A1. This separation most likely derived from the introduction of antibiotics both in clinical and agricultural settings (7). According to Multi Locus Sequence Typing (MLST), *E. faecium* can be divided into sequence types (STs) based on the allelic profile of six

housekeeping genes (8). Particularly, ST17 was identified as the ancestral clone of the hospital-associated clade A1 forming the Clonal Complex 17(CC17) and henceforth, most HA isolates have been identified as members of CC17 (9). Clade A isolates are frequently resistant to high levels of aminoglycosides, ampicillin, and vancomycin, they also hold an array of virulence factors and mobile genetic elements that set them apart from clade B strains (10). The putative virulence genes most thoroughly studied at the functional level include those encoding for an array of cell surface components involved in adhesion, biofilm formation and pili assembly (11). Whole Genome Sequencing (WGS) has facilitated the recognition of further genetic markers conferring genomic plasticity and other putative virulence markers (PVM) associated principally with HA strains (12). A further eventful fact, promoting the emergence of clade A in clinical facilities, was the acquisition of ampicillin resistance through a mutation in the penicillin-binding protein 5 (PBP5) (13). The evolution of PBP5 reflects directly on the phylogenetic diversification of this species (14,15). Clade A isolates carrying the phenotype for ampicillin resistance is described by PBP5-R sequences, while clade B genomes comprise susceptible strains portraying the PBP5-S profile (16)

Enterococcus lactis, as its species name suggests, was firstly isolated from milk samples (17) and was recognized as a new species closely related to *E. faecium* (18). However, differences in 16S rRNA gene sequences (19) and carbohydrate metabolism (20) have set these species apart. Morandi et al. (18) proposed *Enterococcus lactis* as a self-standing *Enterococcus* species, with strain BT159^T (=DSM 23655^T, =LMG 25958^T) as the species type strain. More recently, *Enterococcus xinjiangensis,* isolated from yogurt in China (21) was proposed as a later heterotypic synonym for *E. lactis* based on 16S rRNA gene sequence similarity, ANI and dDDH analyses (22). Although 16S rRNA gene analysis has been considered the classic standard for taxonomical analyses, the presence of high variability within the multiple copies of the 16S rRNA gene sequences impedes an accurate bacterial classification (23). As WGS has become more widely accessible, tools such as ANI, dDDH and pangenome analysis, provide the ultimate classification methods for microbial taxonomy, necessary to establish a valid Overall Genome Related Index (OGRI) (24)(25)(25)(25). OGRI defines threshold values for 16S rRNA gene sequence similarity (98.65%), ANI (96%) and DDH (70%), which provide accurate results for species delineation. Also, the taxonomic studies carried out for *E. lactis* and *E. xinjiangensis* referred to *E. faecium* type strain NCTC 7171^T (clade A2) for their analyses, and such classifications do not consider the different lineages of the species, especially clade B and the polyphyletics of clade A (26).

This study aims at reevaluating the taxonomic relationship among the HA and CA lineages of *Enterococcus faecium* respect to *Enterococcus lactis* and *Enterococcus xinjiangensis* by using a wide set of genomic analysis. Thus, the workflow includes the comparison among the three species through phylogenomics, pan- and core-genome analyses, MLST, 16S rRNA gene sequence similarity, ANI and dDDH. To complete the differentiation among these related species, virulence markers will be investigated according to the latest guidelines of EFSA (6), and the distribution scheme of PVM proposed by Freitas et al.(12). Lastly, ampicillin

resistance will be studied by evaluating the allelic identity of each strain of interest, against the reference sequence for PBP5-S/R profiles.

3 Methods and materials

3.1 Analyzed genomes

A total of 181 enterococcal strains were retrieved from the GenBank NCBI database in February 2021 (27). Specifically, a total of 14 strains of *E. lactis,* 10 strains of *E. xinjiangensis* and 157 *E. faecium* strains was based on relevance and completeness of WGS. For *E. faecium,* 102 strains belong to clade A1, 29 to clade A2, and 26 to clade B (Table 1 and supplementary Table S1). The type strains included in this analysis were NCTC7171^{T} (*E. faecium*), LMG25958^{T} (*E. lactis*), and JCM 30200^{T} (*E. xinjiangensis*). For this study, *E. xinjiangensis* is mentioned as *E. lactis* since it is a taxonomic synonym for this species.

3.2 Annotation, pangenome, phylogenesis and MLST

The listed 181 genomes were annotated using Prokka (28). The pan- and core-genome analysis was computed using Roary v3.11.2 (29) utilizing the annotation output of Prokka. The output given discriminates core genes as being present in 95% to 100% of the strains of interest. The accessory genes are constituted by shell (15% ≤ strains <95%) and cloud genes (0%≤ strains < 15%). The accessory genes are constituted by shell (15% ≤ strains <95%) and cloud genes (0%≤ strains < 15%). Roary creates a matrix based on the alignment of core genes among the strains which will be used to construct a maximum likelihood phylogenetic tree employing RAxML v1.0.0 (30). Results were viewed using iTOL, for organization purposes and mid-point rooting (31).

Multi Locus Sequence Typing analysis of the *E. faecium* species-specific housekeeping genes (*atpA*, *ddl*, *gdh*, *purK*, *gyd*, *pstS*, *adk*) was carried out to understand the population structure of clade B and *E. lactis* strains. Allelic profiles and sequence types (STs) were obtained from the PubMLST website (last visited: March 2021) (32).

3.3 Calculation of Genome Relatedness: 16S rRNA gene sequence similarity, ANI and dDDH

For gene comparison, annotation revision and sequence alignment, the selected 181 genomes were deposited onto Geneious prime V2021.1.1 (https://www.geneious.com). 16S rRNA gene sequence similarity was performed by extracting these sequences from the genomes and calculating distances with the MUSCLE alignment software (33). Same-species assignment was determined by a distance similarity of ≥98.65%.

DNA-DNA hybridization (dDDH) was made by using the Genome-to-Genome Distance Calculator (GGDC) (34). For this analysis, the distance values were computed using the total number of identities within high scoring pairs (HSPs) per total HSP length. In total, 134 genomes have a complete assembly level whereas the rest are at contig or scaffold level. Formula 2, highly recommended if the genomes submitted for analysis are

incomplete, was used. DDH values ≥70% indicate that the tested strain belongs to the same species as the type strain(s) (35).

Average Nucleotide Identity (ANI) analysis was performed using fastANI (36). For this analysis, *E. lactis* and *E. xinjiangensis* strains were compared against the complete repertoire of considered genomes and plotted using Graphpad Prism version 9.1.0 (Graphpad Prism Software, San Diego California, USA, www.graphpad.com). The resulting values were classified using 96% as the cut-off value for same-species determination.

Table 1. List of strains genomes of clade B *E. faecium, E. lactis and E. xinjiangensis* used in this study. These genomes were retrieved from NCBI, with respective accession numbers, assembly level and date of submission, MLST allelic profile and PBP5 profile. MLST allelic profiles not assigned to a specific ST are marked as "no match". PBP-5 sequence profiles for resistance or susceptibility. PBP5-R/-S were estimated according to similarity against the reference PBP5-S sequence from strain Com15.

Species	**Strain**	**Clade**	**Genbank Assembly Accession**	**Assembly level**	**MLST**	**PBP-5 profile**	**Date of submission**
E. faecium	1,141,733	B	GCA_000157575.1	Scaffold	ST327	S_{20}/R_1	16/06/2009
E. faecium	BIOPOP-3ALE	B	GCA_012045365.1	Complete genome	ST819	S_{21}/R_0	04/06/2020
E. faecium	BIOPOP-3WT	B	GCA_012045505.1	Complete genome	ST819	S_{21}/R_0	04/06/2020
E. faecium	BM4105_RF	B	GCA_003269465.1	Complete genome	ST172	S_{20}/R_1	06/28/2018
E. faecium	CBA7134	B	GCA_004015145.1	Complete genome	No match	S20/R1	14/01/2014
E. faecium	Com12	B	GCA_000157635.1	Scaffold	ST107	S_{20}/R_1	21/03/2013
E. faecium	Com15	B	GCA_004006255.1	Chromosome	ST583	S_{21}/R_0	09/01/2019
E. faecium	DMEA02	B	GCA_008330605.1	Complete genome	ST178 (CC94)	S_{21}/R_0	09/09/2019
E. faecium	DT1-1	B	GCA_011745645.1	Complete genome	ST296 (CC94)	S_{21}/R_0	25/03/2020
E. faecium	E1590	B	GCA_000321865.1	Scaffold	ST163	S_{20}/R_1	21/03/2013
E. faecium	E1604	B	GCA_000321885.1	Scaffold	ST75 (CC94)	/	21/03/2013
E. faecium	E1613	B	GCA_000321905.1	Scaffold	ST77	S_{19}/R_2	21/03/2013
E. faecium	E1861	B	GCA_000322085.1	Scaffold	ST289 (CC94)	S_{17}/R_4	21/03/2013
E. faecium	E1972	B	GCA_000322125.1	Scaffold	ST94 (CC94)	S_{17}/R_4	21/03/2013
E. faecium	E2620	B	GCA_000322225.1	Scaffold	ST1175	S_{19}/R_2	21/03/2013

E. faecium	E980	B	GCA_000172615.1	Contig	ST94 (CC94)	S_{17}/R_4	26/03/2010
E. faecium	FS86	B	GCA_013201055.1	Complete genome	ST296 (CC94)	S_{21}/R_0	31/05/2020
E. faecium	Gr17	B	GCA_003711605.1	Complete genome	No match	S_{17}/R_4	05/11/2018
E. faecium	HY07	B	GCA_003574925.1	Complete genome	No match	S_{21}/R_0	19/09/2018
E. faecium	JE1	B	GCA_003667965.1	Complete genome	ST812	S_{16}/R_5	21/10/2018
E. faecium	SRCM103341	B	GCA_004101385.1	Complete genome	No match	S_{17}/R_4	22/01/2019
E. faecium	SRCM103470	B	GCA_004103475.1	Complete genome	No match	S_{10}/R_{11}	23/01/2019
E. faecium	T110	B	GCA_000737555.1	Complete genome	ST812	S_{16}/R_5	04/08/2014
E. faecium	TK-P5D	B	GCA_015377765.1	Complete genome	ST812	S_{16}/R_5	12/11/2020
E. faecium	UC7256	B	GCA_000499925.1	Contig	ST74	S_{17}/R_4	26/11/2013
E. faecium	UC8668	B	GCA_000499905.1	Contig	No match	S_{17}/R_4	26/11/2013
E. lactis	AnGM4_AISHA		GCA_016863785.1	Scaffold	No match	S_{20}/R_1	02/10/2021
E. lactis	CCM8412		GCA_015751045.1	Contig	ST648	S_{21}/R_0	11/12/2020
E. lactis	CICC10840		GCA_009735445.1	Contig	ST1529	S_{18}/R_3	12/08/2019
E. lactis	CICC20089		GCA_009735435.1	Contig	No match	S_{21}/R_0	12/08/2019
E. lactis	CICC20680		GCA_009735475.1	Contig	ST76	/	12/08/2019
E. lactis	CICC24101		GCA_009735495.	Scaffold	ST648	S_{21}/R_0	12/08/2019
E. lactis	CICC6078		GCA_009735405.1	Scaffold	ST296 (CC94)	S_{21}/R_0	12/08/2019
E. lactis	G67-2		GCA_016767515.1	Contig	ST39	S_{16}/R_5	01/28/2021
E. lactis	KCTC21015		GCA_015767715.1	Complete genome	ST648	S_{21}/R_0	12/06/2020
E. lactis	L2672-1		GCA_015751085.1	Contig	ST812	S_{16}/R_5	11/12/2020
E. lactis	LMG25958^T		GCA_015904215.1	Scaffold	ST648	S_{21}/R_0	12/09/2020
E. lactis	MP10_1		GCA_017356435.1	Contig	ST296 (CC94)	S_{21}/R_0	1/28/2021
E. lactis	S10-4		GCA_016767545.1	Contig	ST39	S_{16}/R_5	7/28/2020
E. lactis	s-7		GCA_013867815.1	Scaffold	ST648	S_{21}/R_0	3/15/2021
E. xinjiangensis	HPCN38		GCA_016599235.1	Contig	ST94 (CC94)	S_{17}/R_4	1/10/2021
E. xinjiangensis	JCM30200^T		GCA_015751065.1	Contig	ST289	S_{17}/R_4	11/12/2020
E. xinjiangensis	NM29-3		GCA_016767675.1	Contig	ST94 (CC94)	S_{17}/R_4	1/28/2021
E. xinjiangensis	NM30-4		GCA_016767645.1	Contig	ST94 (CC94)	S_{17}/R_4	1/28/2021

E. xinjiangensis	NM31-5	GCA_016767635.1	Contig	No match	S_{17}/R_4	1/28/2021
E. xinjiangensis	Tb32-6	GCA_016767535.1	Contig	No match	S_{17}/R_4	1/28/2021
E. xinjiangensis	XJ28301	GCA_016767575.1	Contig	ST583	S_{21}/R_0	1/28/2021
E. xinjiangensis	XJ28304	GCA_016767495.1	Contig	ST583	S_{21}/R_0	1/28/2021
E. xinjiangensis	XZ35303	GCA_016767615.1	Contig	No match	S_{17}/R_4	1/28/2021
E. xinjiangensis	XZ37302	GCA_016767595.1	Contig	No match	S_{17}/R_4	1/28/2021

3.4 Hospital-associated markers

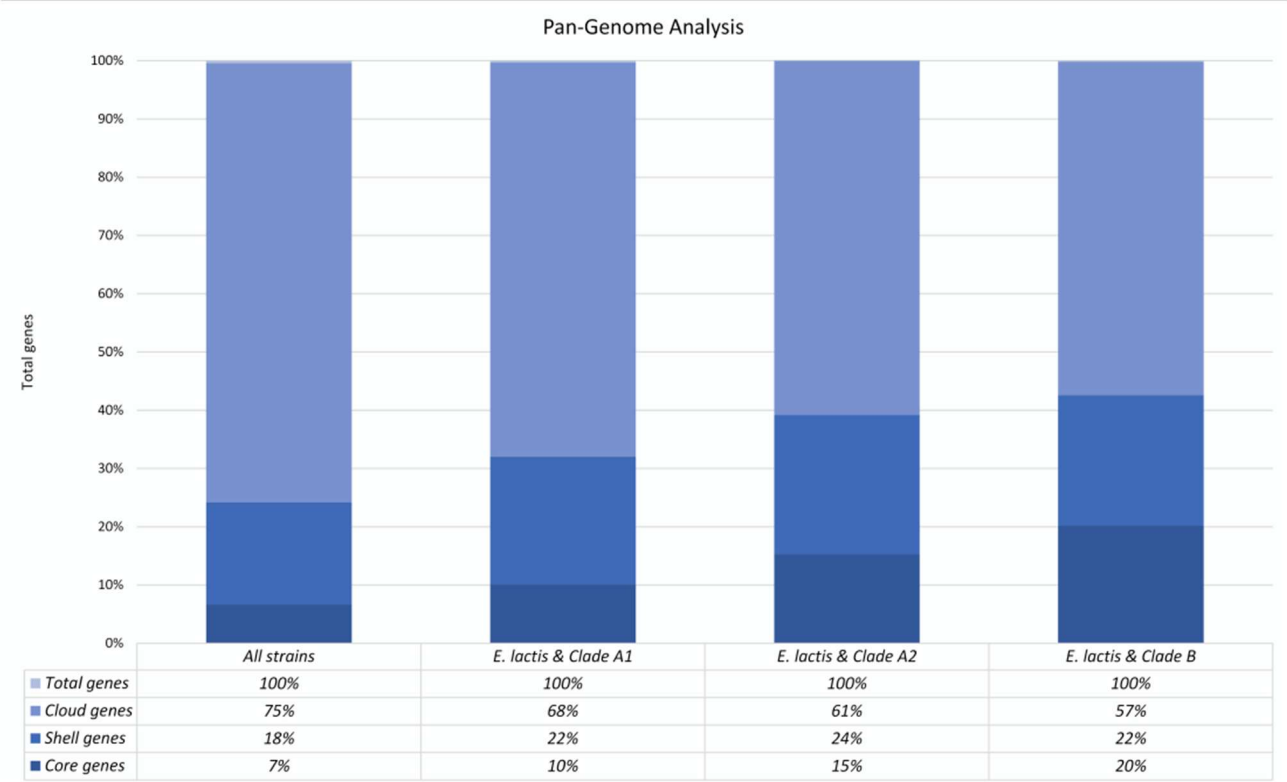
The identification of the major virulence factors was performed interrogating the considered genomes using the Basic Local Alignment Tool (BLAST) from NCBI for the presence of putative virulence factors identified by Freitas et al. (12) and EFSA (6). The considered genomes were interrogated for the presence of the insertion sequence IS16 (AF507977.1), hylfm(HMPREF0351_12988) and *esp*(EFAU004_02750) genes. The additional PVMs are genes coding for putative phosphotransferase *ptsD*(MBG7632288.1), sugar-binding protein encoded by a genomic island *orf1481*(EAN09962.1), collagen-binding protein *ecbA*(AFK59452.1) and nidogen-binding LPXTG surface adhesin *sgrA*(AFK59147.1), cell-wall-anchored collagen adhesin *acm*(WP_002310122.1). Pili gene clusters PGC-1, including *fms20*(VFA66863.1) and *fms21* (ACI49671.1). PGC-2 *fms14* (AFK58156.1), *fms13* (AFK58158.1) and *fms17* (AFK58157.1). PGC-3 containing *ebpA*(YP_006376000*), epbB*(WP_002286054.1) and *ebpC*(WP_002286053.1). Lastly, PGC-4 harboring *fms11*(VFA63747.1), f*ms19*(WP_080263466.1) and *fms16*(WP_002288981.1)

Ampicillin resistance will be analyzed *in silico* for the 50 genomes of *E. lactis, E. xinjiangensis* and *E. faecium*. The determination of the allelic type of penicillin-binding protein 5 sequences, PBP5-S/-R, was carried out as previously established (13). Briefly, PBP5 gene sequences were extracted and translated using Geneious prime V2021.1.1. Clade B strain Com15 was used as reference for PBP5-S amino acid sequence. The chosen clade B and *E. lactis* strains were then aligned against the reference sequence and herewith the allelic profile assigned.

4 Results

4.1 Pan- and core-genome analysis

The first step of this work was to study the genomic relatedness between *E. faecium* and *E. lactis* species, by identifying the pan-genome of 181 genomes of *E. faecium* and *E. lactis* strains. This is composed by 17,294 genes, of which 7% represents the core genome and 93% the accessory genome. In particular, *E. lactis* shares 10% of its core genes with clade A1, 15% with clade A2 and 20% with clade B (Fig. 1). The number of accessory genes in clade A2 and A1 increases gradually when compared with *E. lactis*, mainly because the large accessory genome of HA isolates is related to virulence markers and AMR genes (37). The number of conserved genes among all strains is higher between *E. lactis* and clade B, and it decreases against clade A2 and A1. In the same manner, the number of unique and new genes is lower among clade B and *E. lactis* strains, whereas it increases with clade A2 and A1 gradually.



	All strains	E. lactis & Clade A1	E. lactis & Clade A2	E. lactis & Clade B
Total genes	100%	100%	100%	100%
Cloud genes	75%	68%	61%	57%
Shell genes	18%	22%	24%	22%
Core genes	7%	10%	15%	20%

Figure 1. Pan- and core-genome analysis performed based on clustering of orthologous groups computed by Roary. The analysis was carried out for all strains of *E. faecium* and *E. lactis*, and the respective division between the different clades (A1, A2 and B) and *E. lactis*. Core genes (dark blue) are present in 95% -100% of the strains. Accessory genes, composed by shell genes (blue), shared by 15% ≤ to <95% of the strains and, cloud genes (light blue) shared in 0% - 15% of the strains. The evaluation among all strains depicts a large accessory genome (93%), characteristic of the open pangenome of *E. faecium*. Clade A1 and A2 share 10% and 15% of core genes and, 90% and 85% of accessory genes with *E. lactis*, respectively. Clade B and *E. lactis* shared the largest amount of core genes (20%) and the smallest amount of accessory genes (80%).

4.2 Phylogenesis and population structure

When the 181 considered genomes were used to construct a maximum-likelihood phylogenetic tree, based on the accessory gene sequence alignment, the previously described (7) clade separation in the *E. faecium* species (Fig. 2) was evident. Interestingly, all the *E. lactis* and E. xinjiangensis strains clustered together and were entirely intertwined among clade B strains, whereas clade A1 and A2 diverged into two separated branches, depicting their higher genomic distance.

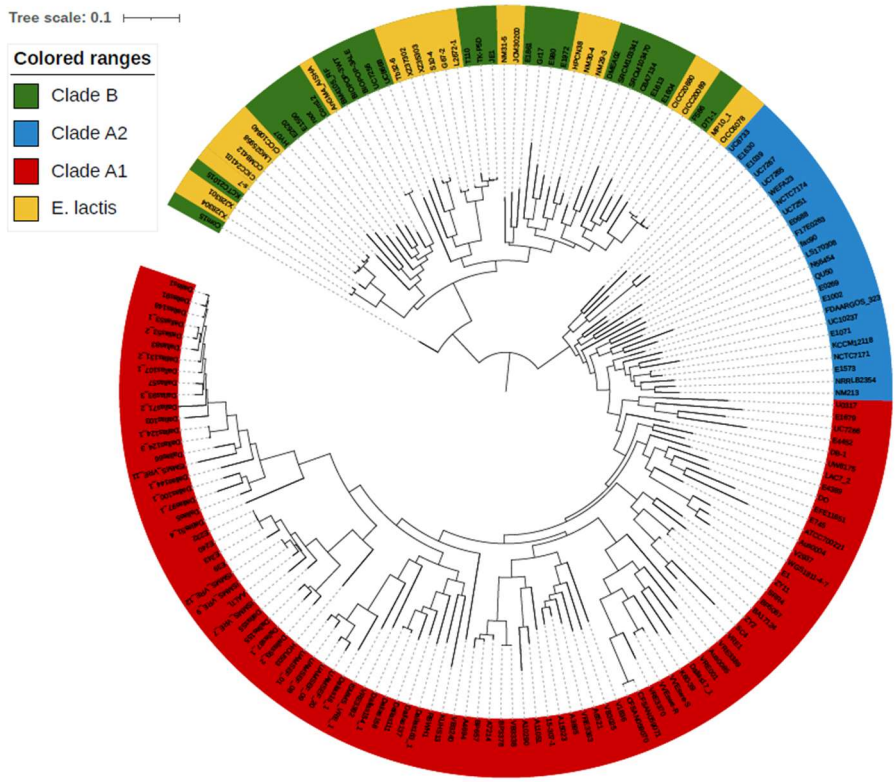


Figure 2. Maximum likelihood phylogenetic tree constructed using the core genome alignment of the selected 181 genomes, depicting *E. faecium* clade A1 strains marked in red, clade A2 strains marked in blue and clade B strains in green. The 14 strains belonging to *E. lactis* (yellow) and the 10 strains of *E. xinjiangensis* (orange) are completely integrated and grouped together within the branch of clade B strains

To better investigate the population structure of *E. lactis*, E. xinjiangensis and *E. faecium* clade B (50 genomes), the MLST analysis scheme (8) for *E. faecium* was applied. This approach indicated that none of the 26 clade B strains nor any of the 14 *E. lactis* or 10 E. xinjiangensis strains presented an allelic profile close to hospital-associated Clonal Complex 17 (CC17). As reported in Table 1, twelve strains were identified to belong to the non-pathogenic human isolated strains Clonal Complex 94 (CC94); six clade B *E. faecium* strains, two *E. lactis* strains and four E. xinjiangensis strains were not matched to an existing allelic profile; the remaining strains are spread in other fifteen sequence types.

The Overall Genome Related Index (38), a systematic workflow including 16S rRNA gene sequence analysis, ANI and dDDH, which allows a comprehensive taxonomical analysis, was applied to clarify the taxonomical relationships between *E. lactis* and *E. faecium*. As a first step, the calculation of the similarity values resulting from the 16S rRNA gene sequence alignment was evaluated against the established threshold of 98.65% for same-species analysis. The outcome of this analysis demonstrated that clade B *and E. lactis* strains present an identity higher than the defined threshold value (Fig. 3). On the contrary, strains from clade A1 and A2 presented a high variability in the similarity when compared to each other and the rest of the strains. Intra-strain variability is characterized by the presence of multiple *rrn* operons. For instance, the six copies of the *rrn* operon of reference strain DO (clade A1) showed heterogeneous results when compared against the *rrn* sequences of the type strains NCTC 7171^T (clade A2), LMG25989^T and JCM30200^T. Two 16S rRNA gene sequences remained below and three above the 98.65% species threshold, whereas the remaining single sequence showed a same-species value with *E. lactis* and *E. xinjiangensis* type strains. Vancomycin resistant strain Aus0004 (clade A1) showed no sequences with values above the cutoff with any of the type strains. The food isolated strain UC7251 (clade A2), presented values above 98.65% with all three type strains. Lastly, when compared to strain T110 (clade B), all six operons presented values above the cut-off with LMG25958^T and JCM30200^T exclusively. These data indicate a limitation in the use of 16S rRNA gene sequence analysis for assignation of strains to *E. faecium* and related species.

When ANI was applied to discriminate among species, using the threshold value of 96%, all the *E. lactis / E. xinjiangensis* strains presented an ANI value higher than 97% with strains of the same species and with all the considered *E. faecium* clade B strains (Fig. 4). Differently, the ANI values with Clade A1 and A2 strains were below the same species threshold value, emphasizing the distance between these two clades and the *E. lactis*/clade B group. This observation was confirmed by the analysis of Genome-to-Genome distance calculation, based on digital DNA-DNA hybridization with a species cut-off of 70%. The 181 considered strains were compared each other using the formula 2 that allows a constant error ratio independent from genome length and assembly(39). *E. lactis* (LMG25958^T) and *E. xinjiangensis* (JCM30200^T) type strains present DDH values higher than 80% against the entire selection of clade B, whereas clade A1 and A2 fall below 65% of identity (Fig. 5).

The overall analysis of the genomic data indicates that clade B strains belong to the same taxonomical unit of *E. lactis* and *E. xinjiangensis* and differ from *E. faecium* clade A1 and A2. Consequently, we propose to rename the *E. faecium* clade B strains as *E. lactis*.

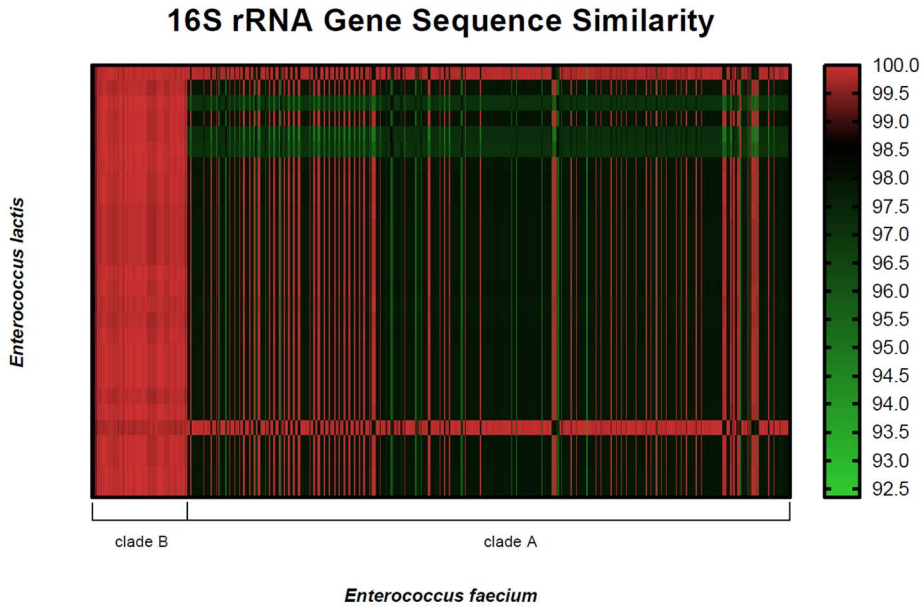


Figure 3. Heatmap representing similarity values of the 16S rRNA sequence alignment performed with MUSCLE. Each cell represents the similarity value of each *rrn* copy amongst the 181 selected strains of *E. lactis* (Y-axis) and *E. faecium* (X-axis). The color scale from green to red represents similarities from furthest to closest respectively. The green hue represents the lower values below the species threshold of 98.65%, whereas the red hue represents genomes with similarities above this species cut-off. *E. lactis* shows same-species values mainly with clade B strains and to a lesser extent to clade A1 and A2 strains.

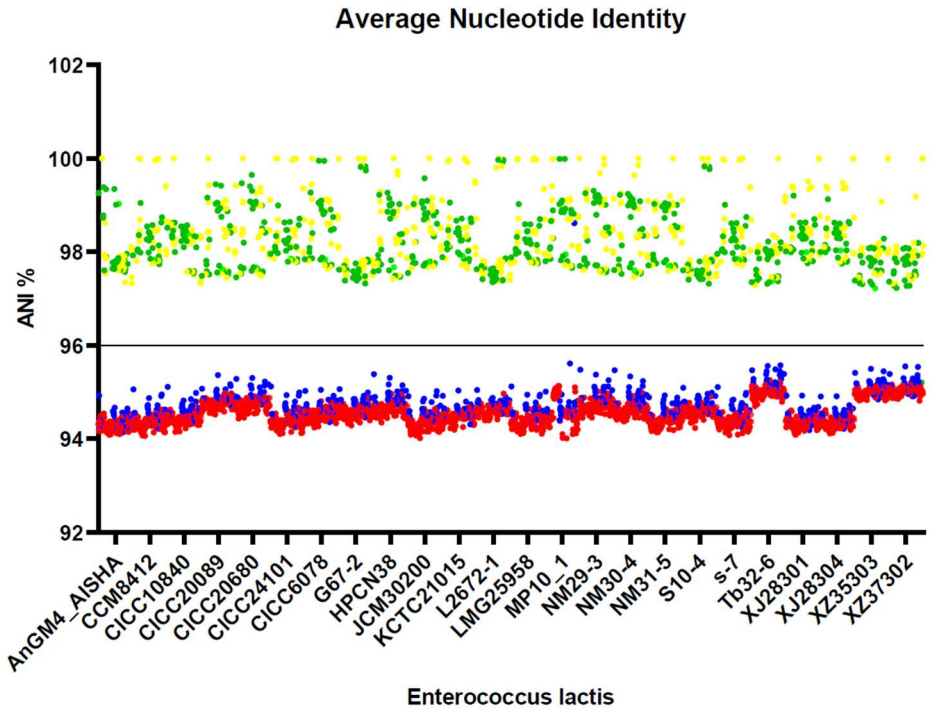


Figure 4. Average Nucleotide Identity (Y-axis) computed by fastANI for all selected *E. faecium* strains against *E. lactis* (X-axis). On the Y-axis the ANI in percentage is divided by the defined species-cutoff value of 96%.

The 102 strains of clade A1 (red) and A2 (blue) prevail below the threshold, while clade B strains (green) are grouped together with *E. lactis* strains (yellow) above the threshold with values higher than 97%.

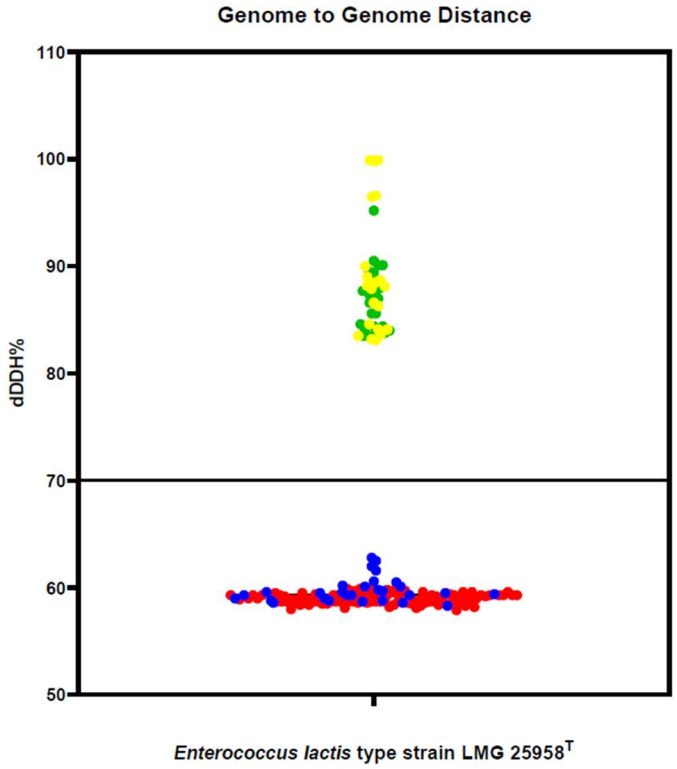


Figure 5. Digital DDH calculated for all selected strains against the type strains of *E. lactis* (LMG25958T) and *E. xinjiangensis* (JCM30200T) (X-axis).The computation formula for this value was selected accordingly to minimize error rates. The Y-axis depicts the DDH distance values, where 70% is defined as the species cut-off limit. The 131 genomes of *E. faecium* clade A1 (red) and clade A2 (blue) are clustered below 70% of DDH identity against both type strains of *E. lactis* and *E. xinjiangensis* (yellow). The 26 clade B strains (green) are grouped together with the 24 *E. lactis* and *E. xinjiangensis* (yellow), with values higher than 80% of DDH distance percentage.

4.3 Hospital-associated markers

The presence of hospital-associated markers in *E. lactis, E. xinjiangensis* and clade B strains was assessed by genome interrogation, using the criteria reported by EFSA (6). None of the analysed strains harbored the insertion sequence IS16, the hyaluronate lyase coding gene hyl_{fm} and the gene coding for the enterococcal surface protein *esp*. Additionally, we investigated on the allelic variations in the PBP5 protein that contribute to ampicillin resistance (13) by analysing the PBP5 amino acid sequences derived from the 50 selected genomes of *E. lactis, E. xinjiangensis* and clade B. Nine strains of *E. faecium* clade B, eight of *E. lactis* and two of *E. xinjiangensis* genomes, harbored the consensus profile of susceptible strains PBP5-S characterized by S_{21}/R_0.The remaining 31 strains presented "hybrid-like" PBP5 sequences as shown in Table 1. None of the analyzed strains contained the PBP5 profile associated to ampicillin resistance.

When the 50 considered genomes were analyzed for the presence of the genes identified by Freitas et al. (12) as putative virulence factors, none of the scrutinized strain contained the full set of genes typical of *E. faecium* strains responsible for human infections, as shown in Supplementary Table S2. Genes coding for adhesion, such as *sgrA* coding for nidogen-binding LPXTG surface adhesin implicated in biofilm formation was found in *E. lactis* strains Tb32-6, XZ35303 and XZ37302, as a cell wall anchor protein with 83% of identity cover. Among the clade B strains, *sgrA* was found in strain 1,141,733 with 84,78% of identity, as well as in strains E1613(87%), E1972(83%), E2620(87%), UC7256(88%) and UC8668(88%). Collagen binding protein *ecbA* was found in strain E1604 only with a 92% of identity. Collagen binding adhesin *acm* was found across all *E. lactis* and clade B strains. With regards to PGC-1, clade B strains BIOPOP-3WT, Com15, E1590, E1613, E1972 and *E. lactis* CICC10840 and CICC20089, and *E. xinjiangensis* HPCN38, NM30-4, XJ28301, XZ35303 and XZ37302 carry both genes coding for PGC-1(*fms20* and *fms21*). *E. xinjiangensis* strain NM30-4 and *E. lactis* CICC20089 carry PGC-4 with full identity for *fms13, fms14* and *fms17*. Concerning the pili gene clusters PGC-3 and PGC-4, none of the strains showed complete identity profiles for the entire array of genes forming the PGCs. The lack of genetic marker for virulence, and incomplete set of PGCs (1-4) together with the susceptible ampicillin genotype of both *E. lactis* and *E. xinjiangensis* together with Clade B, support the safety of this taxonomical unit and the clear separation from virulence and AM resistant isolates (Clade A2 and A1).

5 Discussion

Enterococcus lactis and heterotypic synonym *Enterococcus xinjiangensis*, were identified as novel species mainly by 16S rRNA gene sequence analysis against other enterococcal species. The analyses were carried out with *E. faecium* type strain NCTC 7171^{T}, and distance values set these species apart (18,22). While these studies compared strains of the new described species with the type strain of *E. faecium,* which belong to clade A2, in this study we compared the new species with the diverse genetic lineages of *E. faecium,* including the clade B, which contains isolates associated with commensal behavior and probiotic properties. As in previous reports (2,37,40,41), our study confirmed the phylogenetic divergence of *E. faecium* isolates and the clear separation of clades based on genomic differences. In addition, we have confirmed through 16S rRNA gene sequence analysis, that *E. lactis* and clade A1 and A2 are separate species. On the contrary, *E. lactis* and clade B strains exhibit distance values higher than 98.65%, placing them in the same phylogenetic group. However, taxonomic assignation based on 16S rRNA gene sequence presents some limitations. Thus, *E. faecium* generally carries six copies of the *rrn* operon and we have observed intra-species variability of *rrn* sequences, based on WGS analysis. As an example, in reference strain DO five of the six copies show homology with the *E. faecium* type strain, while the remaining operon is close to *E. lactis*. The diversity of the 16S rRNA gene sequence within a genome, a common phenomenon in bacteria (42), may result in incorrect phylogenetic assignation (43). Consequently, the use of 16S rRNA gene sequence remains limited for taxonomical identification of *E. faecium* and related species.

When WGS analyses were performed, a clearer differentiation between *E. lactis* and clades A1 and A2 of *E. faecium* was achieved. Thus, ANI results showed that strains of these clades cluster below 94-95%, when compared to *E. lactis* and clade B, while *E. lactis* and clade B were grouped above the 96% same-species threshold. This is consistent with other studies, that reported ANI values fluctuating from 94% to 99%, when comparing diverse *E. faecium* isolates from different sources (44,45). Different ANI thresholds have been established to discriminate species ranging from 94% to 96%; however, a stricter threshold (96%) provides a clearer species delimitation especially for species with diverse lineages such as *E. faecium* (24,46,47). Similarly, a different approach to genomic-based taxonomic identification, digital DNA-DNA Hybridization, confirmed the closeness between *E. lactis* with clade B as they were clustered together with values higher than 80%, above the cut-off value of 70% for same-species discrimination. Clade A1 and A2 are substantially distinct from *E. lactis* and clade B as they prevail below the species threshold of DDH. The coherence across the established thresholds for 16S rRNA gene identity, ANI and DDH (48), provide a valid overall genome relatedness index reliable for classifying Clade B strains of *E. faecium* as *E. lactis*.

The distinction between clades A and, *E. lactis and clade B* was also evident after the interrogation of the pangenome, which has been recently regarded as a useful tool for species delimitation based on identification of lineage-specific gene sets (49). Observing the distribution of core- and accessory-genome of our analysis, clade B and *E. lactis* share the highest number of core genes compared to clades A1 and A2. On the other hand, clade A1 and clade A2 show a larger accessory genome when analyzed against *E. lactis*. It has been repeatedly reported that, contrary to clade B, clade A1 and to a smaller extent clade A2 isolates, have a high variability in their accessory genes (9). A highly variable accessory genome is conferred by the fact that the *E. faecium* species has open pangenome and therefore a higher genomic diversity (7,26). The adaptation of *E. faecium* to specific environmental factors, such as clinical settings or antibiotic pressure, have altered the genomic diversity through HGT, genome rearrangement and gene loss (50,51). *E. faecium* isolates responsible for nosocomial infections are largely distinct than those from livestock settings or food origin, mainly because of the presence of AMR genes, virulence features and mobile genetic elements (52).

An additional factor that has been applied to differentiate HA strains of clade A1 and A2 from non-virulent clade B strains was the resistance to ampicillin (6,15). In fact, the phylogenomic diversification of *E. faecium* is highly influenced by the diversity of ampicillin resistance gene sequences (53). *E. faecium* is intrinsically resistant to low levels of ampicillin through cell wall synthesis protein complex PBP, *pbp5 is* part of this operon and sequence variation of PBP5 is sufficient to differentiate two groups of *E. faecium* (54). In addition to the different *pbp5* allelic types there are substantial differences in the expression levels of PBP5 which correspond with the ampicillin resistance phenotype (15). It has been stated that *pbp5* may spread through HGT and specifically that *pbp5* of resistant isolates is located on transferable chromosomal regions, which suggest its dissemination through the clinical environment (55). We confirmed that not only in *E. faecium* clade B the resistance phenotype of PBP5 is lacking but also in all the analyzed *E. lactis* genomes the PBP5-R, the allelic variations related to β-lactam resistance, was absent.

Enterococcal strains with higher ampicillin resistance contain a higher number of PVM, which represents greater risk in terms of pathogenicity (12). The analysis of PVM based on genomic data, indicates that the absence of major virulence factors like insertion sequence IS16, glycosyl hydrolase hyl_{Efm} and enterococcal surface protein *esp* in all considered *E. lactis* and clade B strains. This outcome indicates that the analyzed strains of *E. lactis* and clade B comply with the safety criteria defined by EFSA (6). With respect to the additional PVM analyzed, the isolates lack of putative phosphotransferase *pstD,* collagen binding protein *ecbA,* serine–glutamate containing protein *sgrA,* sugar-binding protein *orf148*. Distinct *E. faecium* genes encoding virulence factors, associated with biofilm formation and adherence are present mainly in clinical isolates (56), and to a shorter extent in CA isolates (12). In this study, none of the 50 clade B and *E. lactis* presented a complete PGC from cluster 1 to 4, while five clade B strains and five *E. xinjiangensis* strains contain a complete PGC-1 profile and two *E. lactis* strains harbor a complete PGC-2. It has been speculated that PGC-1 (*fms21-fms20*) was carried by ancestral and commensal isolates on large plasmids, which explains the presence of this feature in both in HA and CA isolates (57). Other studies revealed the presence of PGC-1 carried by non-clinical isolates (58). None of the 50 genomes of clade B and *E. lactis* hold complete PCG-3 and PGC-4 profiles. PGC-3 formed by the *empABC* operon is mainly present in clinical isolates and it is associated with urinary tract infections (59). Also, PGC-4 was proven to be related to high ampicillin resistance (12). The role of pilum in *E. faecium* and *E. lactis* stays unclear; even so, there are differential regulation and assembly mechanisms in the PGC of clinical isolates and commensal strains (60).

6 Conclusion

WGS based methodologies provide a good identification capacity for the *E. faecium* and *E. lactis* group and regardless of the species delineation method used, clade A is genetically and evolutionarily distinct from clade B and *E. lactis*. This led to the proposal to reclassify all the *E. faecium* clade B as *E. lactis,* recognizing the two groups are phylogenetically separate. Moreover, this re-classification may have an additional impact on the use of these bacteria in the food system. It is known that Clade B strain are used in fermented food industry or as animal probiotics, due to their phenotypic features (61–64). A clear distinction between the *E. faecium* HA strains and the *E. lactis* will also allow to design specific safety assessment procedure, before their use in food or as probiotics, including the consideration for inclusion in the European QPS list.

7 Transparency Declaration

The authors declare that there are no conflicts of interest. This work received no specific grant from any funding agency.

8 References

1. Bra\ek O Ben, Smaoui S. 2019. Enterococci: Between Emerging Pathogens and Potential Probiotics https://doi.org/10.1155/2019/5938210.

2. Ghattargi VC, Gaikwad MA, Meti BS, Nimonkar YS, Dixit K, Prakash O, Shouche YS, Pawar SP, Dhotre DP. 2018. Comparative genome analysis reveals key genetic factors associated with probiotic property in Enterococcus faecium strains. BMC Genomics 19:1–16.

3. Paula PLM de, Moraes ML de, Schueler J, Souza NAA de, Furlaneto MC, Furlaneto-Maia L, Katsuda MS. 2020. Enterococcus faecium in artisanal ripening cheese: technological and safety aspects. Res Soc Dev 9:e299119452.

4. Gao W, Howden BP, Stinear TP. 2018. Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen. Curr Opin Microbiol. Elsevier Ltd https://doi.org/10.1016/j.mib.2017.11.030.

5. EFSA BIOHAZ Panel, Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Cocconcelli PS, Fernández Escámez PS, Maradona MP, Querol A, Suarez JE, Sundh I, Vlak J, Barizzone F, Correia S, Herman L. 2021. The list of QPS status recommended biological agents for safety risk assessments carried out by EFSA https://doi.org/10.5281/ZENODO.4428353.

6. EFSA, Rychen G, Aquilina G, Azimonti G, Bampidis V, Bastos M de L, Bories G, Chesson A, Cocconcelli PS, Flachowsky G, Gropp J, Kolar B, Kouba M, López-Alonso M, López Puente S, Mantovani A, Mayo B, Ramos F, Saarela M, Villa RE, Wallace RJ, Wester P, Glandorf B, Herman L, Kärenlampi S, Aguilera J, Anguita M, Brozzi R, Galobart J. 2018. Guidance on the characterisation of microorganisms used as feed additives or as production organisms. EFSA J 16.

7. Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, Corander J, Cheng L, Saif S, Young S, Zeng Q, Wortman J, Birren B, Willems RJL, Earl AM, Gilmore MS. 2013. Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. MBio 4.

8. Homan WL, Tribe D, Poznanski S, Li M, Hogg G, Spalburg E, Van Embden JDA, Willems RJL. 2002. Multilocus sequence typing scheme for Enterococcus faecium. J Clin Microbiol 40:1963–1971.

9. Lee T, Pang S, Abraham S, Coombs GW. 2019. Antimicrobial-resistant CC17 Enterococcus faecium: The past, the present and the future. J Glob Antimicrob Resist 16:36–47.

10. Gorrie C, Higgs C, Carter G, Stinear TP, Howden B. 2019. Genomics of vancomycin-resistant Enterococcus faecium. Microb Genomics 5.

11. Freitas AR, Tedim AP, Novais C, Coque TM, Peixe L. 2018. Distribution of putative virulence markers in Enterococcus faecium: towards a safety profile review. J Antimicrob Chemother 73:306–319.

12. Pietta E, Montealegre MC, Roh JH, Cocconcelli PS, Murray BE. 2014. Enterococcus faecium PBP5-S/R, the Missing Link between PBP5-S and PBP5-R. Antimicrob Agents Chemother 58:6978–6981.

13. Freitas AR, Pereira AP, Novais C, Peixe L. 2021. Multidrug-resistant high-risk Enterococcus faecium clones: can we really define them? Int J Antimicrob Agents 57:106227.

14. Montealegre MC, Roh JH, Rae M, Davlieva MG, Singh K V., Shamoo Y, Murray BE. 2017. Differential Penicillin-Binding Protein 5 (PBP5) Levels in the Enterococcus faecium Clades with Different Levels of Ampicillin Resistance. Antimicrob Agents Chemother 61:1–10.

15. Guzman Prieto AM, van Schaik W, Rogers MRC, Coque TM, Baquero F, Corander J, Willems RJL. 2016. Global emergence and dissemination of enterococci as nosocomial pathogens: Attack of the clones? Front Microbiol 7:1–15.

16. Botina SG, Sukhodolets V V. 2006. Speciation in bacteria: Comparison of the 16S rRNA gene for closely related Enterococcus species. Russ J Genet 42:247–251.

17. Morandi S, Cremonesi P, Povolo M, Brasca M. 2012. Enterococcus lactis sp. nov., from Italian raw milk cheeses. Int J Syst Evol Microbiol 62:1992–1996.

18. Sukhodolets V V., Botina SG, Lysenko AM, Trenina MA. 2005. The lactic acid enterococci Enterococcus faecium and Enterococcus durans: nucleotide sequence diversity in 16S rRNA genes. Mikrobiologiia 74:810–815.

19. Morandi S, Brasca M, Lodi R. 2011. Technological, phenotypic and genotypic characterisation of wild lactic acid bacteria involved in the production of Bitto PDO Italian cheese. Dairy Sci Technol 91:341–359.

20. Ren X, Li M, Guo D. 2016. Enterococcus Xinjiangensis sp. nov., Isolated from Yogurt of Xinjiang, China https://doi.org/10.1007/s00284-016-1065-1.

21. Li YQ, Gu CT. 2021. Proposal of Enterococcus xinjiangensis Ren et al. 2020 as a later heterotypic synonym of Enterococcus lactis Morandi et al. 2012. Int J Syst Evol Microbiol 004716.

22. Ibal JC, Pham Q, Park CE, Shin J-H. 2019. Information about variations in multiple copies of bacterial 16S rRNA genes may aid in species identification https://doi.org/10.1371/journal.pone.0212090.

23. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 68:461–466.

24. Chung M, Munro JB, Tettelin H, Dunning Hotopp JC. 2018. Using Core Genome Alignments To Assign Bacterial Species. mSystems 3.

25. Bayjanov JR, Baan J, Rogers MRC, Troelstra A, Willems RJL, van Schaik W. 2019. Enterococcus faecium genome dynamics during long-term asymptomatic patient gut colonization. Microb Genomics 5.

26. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2013. GenBank. Nucleic Acids Res 41:D36.

27. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

28. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

29. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455.

30. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. Nucleic Acids Res 47:W256–W259.

31. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; referees: 2 approved]. Wellcome Open Res 3.

32. Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.

33. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60.

34. Auch AF, Klenk HP, Göker M. 2010. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci 2:142–148.

35. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:1–8.

36. Beukers AG, Zaheer R, Goji N, Amoako KK, Chaves A V., Ward MP, McAllister TA. 2017. Comparative genomics of Enterococcus spp. isolated from bovine feces. BMC Microbiol 17.

37. Chun J, Rainey FA. 2014. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. Int J Syst Evol Microbiol 64:316–324.

38. Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial

species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134.

39. Galloway-Peña J, Roh JH, Latorre M, Qin X, Murray BE. 2012. Genomic and SNP Analyses Demonstrate a Distant Separation of the Hospital and Community-Associated Clades of Enterococcus faecium. PLoS One 7:e30187.

40. Kim EB, Marco ML. 2014. Nonclinical and clinical enterococcus faecium strains, but not enterococcus faecalis strains, have distinct structural and functional genomic features. Appl Environ Microbiol 80:154–165.

41. Stolyar SM, Ravin N, Graf J, Espejo Romilio RT, Espejo RT, Plaza N. 2018. Multiple Ribosomal RNA Operons in Bacteria; Their Concerted Evolution and Potential Consequences on the Rate of Evolution of Their 16S rRNA https://doi.org/10.3389/fmicb.2018.01232.

42. Větrovský T, Baldrian P. 2013. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. PLoS One 8.

43. Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J, Feldgarden M, Haas B, Birren B, Gilmore MS. 2012. Comparative genomics of enterococci: Variation in Enterococcus faecalis, clade structure in *E. faecium*, and defining characteristics of E. gallinarum and E. casseliflavus. MBio 3:1–11.

44. Zhong Z, Kwok LY, Hou Q, Sun Y, Li W, Zhang H, Sun Z. 2019. Comparative genomic analysis revealed great plasticity and environmental adaptation of the genomes of Enterococcus faecium. BMC Neurosci 20.

45. Kanyó I, Molnár LV. 2016. Procaryotic species and subspecies delineation using average nucleotide identity and gene order conservation. Gene Reports 5:75–82.

46. Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol Microbiol 68:2386–2392.

47. Kim M, Oh HS, Park SC, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351.

48. Moldovan MA, Gelfand MS. 2018. Pangenomic definition of prokaryotic species and the phylogenetic structure of Prochlorococcus spp. Front Microbiol 9:428.

49. Bonacina J, Suárez N, Hormigo R, Fadda S, Lechner M, Saavedra L. 2017. A genomic view of food-related and probiotic Enterococcus strains. DNA Res An Int J Rapid Publ Reports Genes Genomes

24:11.

50. Mikalsen T, Pedersen T, Willems R, Coque TM, Werner G, Sadowy E, Van Schaik W, Jensen LB, Sundsfjord A, Hegstad K. 2015. Investigating the mobilome in clinically important lineages of enterococcus faecium and enterococcus faecalis. BMC Genomics 16:1–16.

51. Gouliouris T, Raven KE, Ludden C, Blane B, Corander J, Horner CS, Hernandez-Garcia J, Wood P, Hadjirin NF, Radakovic M, Holmes MA, de Goffau M, Brown NM, Parkhill J, Peacock SJ. 2018. Genomic surveillance of enterococcus faecium reveals limited sharing of strains and resistance genes between livestock and humans in the United Kingdom. MBio 9.

52. Novais C, Tedim AP, Lanza VF, Freitas AR, Silveira E, Escada R, Roberts AP, Al-Haroni M, Baquero F, Peixe L, Coque TM. 2016. Co-diversification of Enterococcus faecium core genomes and PBP5: Evidences of PBP5 horizontal transfer. Front Microbiol 7:1–17.

53. Miller WR, Munita JM, Arias CA. 2014. Mechanisms of antibiotic resistance in enterococci. Expert Rev Anti Infect Ther 12:1221–1236.

54. Morroni G, Brenciani A, Litta-Mulondo A, Vignaroli C, Mangiaterra G, Fioriti S, Citterio B, Cirioni O, Giovanetti E, Biavasco F. 2019. Characterization of a new transferable MDR plasmid carrying the pbp5 gene from a clade B commensal Enterococcus faecium. J Antimicrob Chemother 74:843–850.

55. Wagner T, Joshi B, Janice J, Askarian F, Škalko-Basnet N, Hagestad OC, Mekhlif A, Wai SN, Hegstad K, Johannessen M. 2018. Enterococcus faecium produces membrane vesicles containing virulence factors and antimicrobial resistance related proteins. J Proteomics 187:28–38.

56. Kim DS, Singh K V., Nallapareddy SR, Qin X, Panesso D, Arias CA, Murray BE. 2010. The fms21 (pilA)-fms20 locus encoding one of four distinct pili of Enterococcus faecium is harboured on a large transferable plasmid associated with gut colonization and virulence. J Med Microbiol. Microbiology Society https://doi.org/10.1099/jmm.0.016238-0.

57. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JEP, Schapendonk CME, Hendrickx APA, Nijman IJ, Bonten MJM, Tettelin H, Willems RJL. 2010. Pyrosequencing-based comparative genome analysis of the nosocomial pathogen Enterococcus faecium and identification of a large transferable pathogenicity island. BMC Genomics 11:1–18.

58. Sillanpää J, Prakash VP, Nallapareddy SR, Murray BE. 2009. Distribution of genes encoding MSCRAMMs and pili in clinical and natural populations of Enterococcus faecium. J Clin Microbiol 47:896–901.

59. García-Solache M, Rice LB. 2019. The enterococcus: A model of adaptability to its environment. Clin Microbiol Rev. American Society for Microbiology https://doi.org/10.1128/CMR.00058-18.

60. Hassanzadazar H, Ehsani A, Mardani K. 2014. Antibacterial activity of Enterococcus faecium derived from Koopeh cheese against Listeria monocytogenes in probiotic ultra-filtrated cheese. Vet Res forum an Int Q J 5:169–75.

61. do Nascimento M da S, Moreno I, Kuaye AY. 2010. Antimicrobial activity of Enterococcus Faecium Fair-E 198 against gram-positive pathogens. Braz J Microbiol 41:74–81.

62. Yerlikaya O, Akbulut N. 2020. In vitro characterisation of probiotic properties of Enterococcus faecium and Enterococcus durans strains isolated from raw milk and traditional dairy products. Int J Dairy Technol 73:98–107.

63. Pieniz S, Andreazza R, Okeke BC, Camargo FAO, Brandelli A. 2015. Antimicrobial and antioxidant activities of Enterococcus species isolated from meat and dairy products. Braz J Biol 75:923–931.

10 Supplementary materials

Table S1. List of the selected clade A1 and A2 Enterococcus faecium strains, publicly available.

Strain	Clade	Accession No.	Strain	Clade	Accession No.	Strain	Clade	Accession No.	Strain	Clade	Accession No.
6E6	A1	GCA_001518735.1	Dallas137	A1	GCA_016415445.1	E4389	A1	GCA_000322425.1	VRE3363	A1	GCA_015476295.1
1,230,933	A1	GCA_000157435.1	Dallas144_1	A1	GCA_016415405.1	E4452	A1	GCA_000239115.2	VRE3370	A1	GCA_015325925.1
15-307-1	A1	GCA_002973755.2	Dallas148	A1	GCA_016415425.1	E745	A1	GCA_001750885.1	VRE3382	A1	GCA_015999425.1
2014-VREF-268	A1	GCA_002025045.1	Dallas154_1	A1	GCA_016415545.1	EFE11651	A1	GCA_900044005.1	VRE3389	A1	GCA_015999405.1
2014-VREF-41	A1	GCA_002007625.1	Dallas155	A1	GCA_016415505.1	HOU503	A1	GCA_005952885.1	VVEswe-R	A1	GCA_007917035.3
2014-VREF-63	A1	GCA_002025065.1	Dallas158	A1	GCA_016415565.1	ISMMS_VRE_1	A1	GCA_001720945.1	VVEswe-S	A1	GCA_007917315.3
A10290	A1	GCA_012933345.2	Dallas16_1	A1	GCA_016406485.1	ISMMS_VRE_11	A1	GCA_001721905.1	WGS1811-4-7	A1	GCA_016864255.1
A11051	A1	GCA_012933285.2	Dallas163_1	A1	GCA_016415485.1	ISMMS_VRE_12	A1	GCA_001953235.1	ZY11	A1	GCA_009938075.1
A15023	A1	GCA_012933295.2	Dallas17_1	A1	GCA_016406445.1	ISMMS_VRE_7	A1	GCA_001721065.1	ZY2	A1	GCA_010120755.1
A3895	A1	GCA_012933265.2	Dallas5	A1	GCA_016126675.1	ISMMS_VRE_9	A1	GCA_001953255.1			
A4694	A1	GCA_012933245.2	Dallas51_4	A1	GCA_016406465.1	K60-39	A1	GCA_002334625.1			
A6521	A1	GCA_012933195.2	Dallas53_1	A1	GCA_016406365.1	KUHS13	A1	GCA_009938285.1			
A7214	A1	GCA_012933165.2	Dallas53_2	A1	GCA_016406545.1	LAC7_2	A1	GCA_009036045.1			
AALTL	A1	GCA_002880635.1	Dallas55	A1	GCA_016406385.1	RBWH1	A1	GCA_003957785.1			
ATCC700221	A1	GCA_001594345.1	Dallas57	A1	GCA_016406565.1	SC4	A1	GCA_002848385.1			
Aus0004	A1	GCA_000250945.1	Dallas66	A1	GCA_016406405.1	SRR24	A1	GCA_009734005.2			
Aus0085	A1	GCA_000444405.1	Dallas71_2	A1	GCA_016406525.1	U0317	A1	GCA_000172915.1			
BA17124	A1	GCA_012932975.2	Dallas83	A1	GCA_016406425.1	UAMSEF_01	A1	GCA_005886545.1			
BP3378	A1	GCA_012933055.2	Dallas87_1	A1	GCA_016406505.1	UAMSEF_08	A1	GCA_005886655.1			
BP5067	A1	GCA_012932985.2	Dallas91	A1	GCA_016406345.1	UAMSEF_09	A1	GCA_005886715.1			
BP657	A1	GCA_012933075.2	Dallas93_2	A1	GCA_016406345.1	UAMSEF_20	A1	GCA_005886735.1			
CFSAN059070	A1	GCA_003071425.1	Dallas93_3	A1	GCA_016406585.1	UC7266	A1	GCA_000764975.1			
CFSAN059071	A1	GCA_003071445.1	Dallas97_1	A1	GCA_016415025.1	UW8175	A1	GCA_001587115.1			
Dallas1	A1	GCA_015999605.1	DB-1	A1	GCA_006337045.1	V1836	A1	GCA_008728455.1			
Dallas100_1	A1	GCA_016415285.1	DO	A1	GCA_000174395.2	V2937	A1	GCA_008728475.1			
Dallas103	A1	GCA_016415105.1	E1	A1	GCA_001886635.1	VB3025	A1	GCA_005517315.1			
Dallas107_1	A1	GCA_016415325.1	E1679	A1	GCA_000172875.1	VB3240	A1	GCA_005576735.1			
Dallas111	A1	GCA_016415345.1	E232	A1	GCA_002777275.1	VB3338	A1	GCA_014874615.1			
Dallas124_1	A1	GCA_016415385.1	E240	A1	GCA_002761255.1	VRE	A1	GCA_009697285.1			
Dallas124_3	A1	GCA_016415365.1	E243	A1	GCA_002761275.1	VRE001	A1	GCA_001895905.1			
Dallas131_2	A1	GCA_016415465.1	E39	A1	GCA_001635875.1	VRE1	A1	GCA_006007925.1			

Strain	Clade	Accession No.	Strain	Clade	Accession No.	Strain	Clade	Accession No.	Strain	Clade	Accession No.

1,231,501	A2	GCA_000157555.1	E1071	A2	GCA_000172655.1	N56454	A2	GCA_006351845.1	UC7265	A2	GCA_000764985.1
16-346	A2	GCA_002761555.1	E1573	A2	GCA_000321765.1	NCTC7171	A2	GCA_900447735.1	UC7267	A2	GCA_000411635.1
3012STDY6244127	A2	GCA_900683475.1	E1630	A2	GCA_000322025.1	NCTC7174	A2	GCA_900637035.1	UC8733	A2	GCA_000499945.1
64-3	A2	GCA_001298485.1	F17E0263	A2	GCA_006280355.1	NM213	A2	GCA_005166365.1	UW7606x64/3 TC1	A2	GCA_001412695.1
E0269	A2	GCA_000321525.1	fac90	A2	GCA_016743855.1	NRRL B-2354	A2	GCA_000336405.1	WEFA23	A2	GCA_002850515.1
E0688	A2	GCA_000321605.1	FDAARGOS_323	A2	GCA_002983785.1	QU 50	A2	GCA_006741355.1			
E1002	A2	GCA_900066025.1	KCCM 12118	A2	GCA_015767695.1	UC10237	A2	GCA_000499965.1			
E1039	A2	GCA_000174935.1	LS170308	A2	GCA_002831505.1	UC7251	A2	GCA_000411655.1			

Table S2. Presence of major putative virulence factors interrogated for all *E. faecium* clade B and *E. lactis*. Percentage values represent the nucleotide identity between the query and subject sequence. The considered genomes were interrogated for the presence of the following genes: insertion sequence (*IS16*), hyalunorate lyase (*hylfm*) and enterococcal surface protein (*esp*), putative phosphotransferase (*ptsD*), sugar-binding protein encoded by a genomic island orf1481, collagen-binding protein (*ecbA*) and nidogen-binding LPXTG surface adhesin (*sgrA*), cell-wall-anchored collagen adhesin (*acm*). Pili Gene Cluster-1 (PGC-1), including pilus proteins genes *fms20* and *fms21*. Pili Gene Cluster-2 (PGC-2), including pilus proteins *fms14*, *fms13* and *fms17*. PGC-3 containing endocarditis- and biofilm-associated pili A(*ebpA*), endocarditis- and biofilm-associated pili B (*epbB*) and endocarditis- and biofilm-associated pili C (*ebpC*). PGC-4 harboring including pilus proteins *fms11*, *fms19* and *fms16.*

												PGC-1			PGC-2			PGC-3			PGC-4		
Species	**Strain**	**Accession No.**	***pstD***	***orf1481***	***sgrA***	***IS16***	***esp***	***hylEfm***	***ecbA***	***acm***	***fms20***	***fms21***	***fms14***	***fms13***	***fms17***	***ebpA***	***epbB***	***ebpC***	***fms11***	***fms19***	***fms16***
E. faecium	1,141,733	GCA_000157575.1	-	-	+ (84,78%)	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	BIOPOP-3ALE	GCA_012045365.1	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-
	BIOPOP-3WT	GCA_012045505.1	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-
	BM4105_RF	GCA_003269465.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-
	CBA7134	GCA_004015145.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+ (96%)
	Com12	GCA_000157635.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-
	Com15	GCA_004006255.1	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	+ (96%)
	DMEA02	GCA_008330605.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	DT1-1	GCA_011745645.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+ (96%)
	E1590	GCA_000321865.1	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-
	E1604	GCA_000321885.1	-	-	-	-	-	-	+ (92%)	+	-	-	-	-	-	-	-	-	-	-	-
	E1613	GCA_000321905.1	-	-	+ (87%)	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-
	E1861	GCA_000322085.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-
	E1972	GCA_000322125.1	-	-	+ (83%)	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-
	E2620	GCA_000322225.1	-	-	+ (87%)	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-

	E980	GCA_000172615.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+ (96%)
	FS86	GCA_013201055.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+ (96%)
	Gr17	GCA_003711605.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	HY07	GCA_003574925.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	JE1	GCA_003667965.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	SRCM103341	GCA_004101385.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+ (96%)
	SRCM103470	GCA_004103475.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	T110	GCA_000737555.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+ (96%)
	TK-P5D	GCA_015377765.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+ (96%)
	UC7256	GCA_000499925.1	-	-	+ (88%)	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	UC8668	GCA_000499905.1	-	-	+ (88%)	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+ (91%)	+ (96%)
E. lactis	AnGM4_AISHA	GCA_016863785.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	+ (96%)
	CCM8412	GCA_015751045.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	CICC10840	GCA_009735445.1	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+ (95%)
	CICC20089	GCA_009735435.1	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-
	CICC20680	GCA_009735475.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	CICC24101	GCA_009735495.	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	CICC6078	GCA_009735405.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+ (95%)
	G67-2	GCA_016767515.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	+ (95%)
	KCTC21015	GCA_015767715.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	L2672-1	GCA_015751085.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	+ (95%)
	LMG25958	GCA_015904215.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	MP10_1	GCA_017356435.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+ (95%)
	S10-4	GCA_016767545.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	+ (95%)
	s-7	GCA_013867815.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	HPCN38	GCA_016599235.1	-	-	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-
	JCM30200	GCA_015751065.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	NM29-3	GCA_016767675.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	NM30-4	GCA_016767645.1	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-
	NM31-5	GCA_016767635.1	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	+ (95%)

	Tb32-6	GCA_016767535.1	-	-	+ (83%)	-	-	-	-	+	+	-	-	-	-	-	-	-	+ (91%)	+ (95%)
	XJ28301	GCA_016767575.1	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	+ (95%)
	XJ28304	GCA_016767495.1	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+ (95%)
	XZ35303	GCA_016767615.1	-	-	+ (83%)	-	-	-	-	+	+	+	-	-	-	-	-	-	+ (91%)	+ (95%)
	XZ37302	GCA_016767595.1	-	-	+ (83%)	-	-	-	-	+	+	+	-	-	-	-	-	-	+ (91%)	+ (95%)

‘+’ indicates the presence of the interrogated gene

‘-’ indicates the absense of the interrogated gene

CHAPTER 4

Distinction between *Enterococcus faecium* and *Enterococcus lactis* by a *gluP*-PCR-based assay for accurate identification and diagnostics

Chapter 4

Distinction between *Enterococcus faecium* and *Enterococcus lactis* by a *gluP*-PCR-based assay for accurate identification and diagnostics

Mireya Viviana Belloso Daza[1#], Ana C. Almeida-Santos[2,3#], Carla Novais[2,3], Antónia Read[5], Valquíria Alves[5], Pier Sandro Cocconcelli[1], Ana R. Freitas[2,3,4*], Luísa Peixe[2,3*]

1) Università Cattolica del Sacro Cuore, Department for Sustainable Food Process (DISTAS), Università Cattolica del Sacro Cuore, Piacenza, Italy

2) UCIBIO/REQUIMTE, Applied Molecular Biosciences Unit, Department of Biological Sciences, Laboratory of Microbiology, Faculty of Pharmacy, University of Porto, Portugal

3) Associate Laboratory i4HB, Institute for Health and Bioeconomy, Faculty of Pharmacy, University of Porto, Portugal

4) TOXRUN – Toxicology Research Unit, University Institute of Health Sciences, CESPU, CRL, 4585-116 Gandra, Portugal.

5) Clinical Pathology Service- Microbiology, Pedro Hispano Hospital, Matosinhos, Portugal

#These authors contributed equally to this work.

***Corresponding authors:** Ana R. Freitas (afreitas@ff.up.pt) and Luísa Peixe (lpeixe@ff.up.pt)

Keywords: *Enterococcus faecium*, *Enterococcus lactis*, PCR-based differentiation, *gluP*, Public Health

This article was accepted for publication on Spectrum Microbiology: Manuscript #Spectrum03268-22R1-A

1 Abstract

It was recently proposed that *Enterococcus faecium* colonizing the human gut (previous clade B) actually corresponds to *Enterococcus lactis*. Our goal was to develop a PCR to rapidly differentiate these species and to discuss the main phenotypic and genotypic differences from a clinical perspective. The pangenome of 512 genomes from *E. faecium* and *E. lactis* was analysed to assess diversity in genes among the two species. Sequences were aligned to find the best candidate gene for designing species-specific primers and testing their accuracy in a collection of 382 enterococci. *E. lactis* isolates from clinical origin were further characterized by whole genome sequencing (Illumina). Pangenome analysis resulted in 12 gene variants, with gene *gluP* (rhomboid protease) selected as the candidate for species differentiation. The nucleotide sequence of *gluP* diverged by 90-92% between sets and allowed species identification through PCR with 100% specificity and no cross-reactivity. *E. lactis* were greatly pan-susceptible and non-host-specific. Clinical *E. lactis* isolates were susceptible to clinically-relevant antibiotics, lacked infection-associated virulence markers, and were associated with patients presenting risk factors for enhanced bacterial translocation. Here, we propose a PCR-based assay using *gluP* for an easy routine differentiation between *E. faecium* and *E. lactis* that could be implemented in different Public Health contexts. We further suggest that *E. lactis*, a dominant human gut species, can translocate the gut barrier in severely ill, immunodeficient, and surgical patients. Knowing that bacterial translocation may be a sepsis promoter, the relevance of infections caused by *E. lactis*, even if pan-susceptible, should be explored.

2 Introduction

Enterococcus faecium has emerged as a leading nosocomial multidrug-resistant (MDR) pathogen responsible for hospital-acquired infections worldwide (1).The population structure of *E. faecium* has been divided into distinct clades, clade A consisting of hospital- and animal-associated isolates and clade B containing community-associated isolates (2). In a previous study, we demonstrated that *Enterococcus lactis* (*E. lactis)* and *E. faecium* from clade B are genetically and evolutionarily distinct from clade-A *E. faecium* (3). In addition, other features distinguishing clade A from clade-B *E. faecium* isolates included the often resistance of the former to different antibiotics (*e.g.,* high levels of aminoglycosides, ampicillin, and/or vancomycin) and the enrichment in a variety of virulence factors and/or mobile genetic elements (3). Based on these data it was proposed to reclassify the clade B *E. faecium* as *E. lactis* (from now on always designated as *E. lactis*) as they are in fact the same species (3).

Although the extent of *E. lactis* causing human infections is much lower than that of clade-A *E. faecium*, they are currently being misidentified as *E. faecium* in hospitals worldwide. Large epidemiological studies previously showed *E. lactis* genomes in association with a significant number of bacteremia isolates and with vancomycin resistance as well (4). Also, many probiotics or feed formulas contain *E. faecium* which in fact could actually correspond to *E. lactis* (5). In this context, it is urgent to easily differentiate between *E. faecium*

and *E. lactis* not only for accurate patient diagnosis and infection prognosis but also for a correct taxonomic classification in different epidemiological and vigilance programs besides industry purposes. Therefore, we developed a PCR for rapid detection and differentiation between these species. Given, the lack of studies characterizing *E. lactis* from hospitalized patients, we also explored the genome and phenotypic features of the clinical *E. lactis* identified in this study.

3 Methods and materials

3.1 Pan-genome analysis and species-specific primer design

A total of 512 enterococcal genomes were retrieved from the GenBank NCBI database. The genomes were submitted to the genome-to-genome distance calculator online tool to compute digital DNA-DNA hybridization (dDDH) against *E. faecium* type strain ATCC 700221^T to discriminate between clade A (≥70%) and clade B/*E. lactis* (<70%) (Table 1S) (6). By the time of analysis (January 2022), the selection of strains was based on the level of genome completeness (only complete assemblies) as well as their relevance in the clinical or agri-food fields measured by their inclusion in different publications (Table S1). Genomes were annotated with Prokka (6) and submitted to pan-genome analysis using Roary v3.11.2 (7). The output discriminates core genes as being present in 95-100% of the strains of interest. The accessory genes are constituted by shell (15% ≤ - <95%) and cloud genes (0%≤ - < 15%). To differentiate between both species and create specific primers, we analysed unique genes in *E. lactis* that could be absent in *E. faecium* or gene variants differing among both species, by using the *query_pan_genome* command. We defined set1 of isolates containing *E. lactis* and set2 clade A *E. faecium*. Gene variants were then extracted from all genomes of set1 and were subsequently aligned to evaluate allelic differences among set2 isolates. We also submitted these genes to BLAST (8) to corroborate the allelic variance between *E. lactis* and *E. faecium* genomes. All extraction and alignment steps were performed with software Geneious Prime V2022.0.1. Good gene candidates were furtherly evaluated to test the accuracy of this method and the selected gene was used for primer design using Primer3 (9). Finally, the proposed primers were tested *in silico* with the genomes of *E. faecium* ATCC 700221^T and *E. lactis* LMG 25958^T type strains and *in vitro* in a collection of 137 well-characterized *E. faecium* (61 from human clinical origin, 42 from animals, 21 from healthy individuals, and 13 from miscellaneous sources) that have been previously classified as clade A (n=110) or clade B (n=27) in previous surveillance studies (10). They were also tested in other 245 enterococci isolates for which identification and/or clonality was not established (unknown clade or even species) to test eventual cross-reactions between *E. lactis, E. faecium* and non *E. faecium/E. lactis* isolates. PCR was performed in a BioRad iCycler equipment with PCR conditions and primer details being described in Table 1.

Table 1. Species-specific primers and PCR conditions for differentiation of *E. faecium* and *E. lactis* species by ampifying the *gluP* gene.

		Sequence (5'->3')	Length	Start	Stop	Tm	GC %	Product size	PCR conditions
E. faecium	Forward	GCGTGCATGGTT AAGACGAC	20	27	46	59.91	55	427	1 cycle 10'-94ºC; 30 cycles 30''-94ºC, 30''-61ºC, 30''-72ºC; 1 cycle 10'-72ºC
	Reverse	CTGCTGGATCGCT GGGTTAT	20	453	434	59.89	55		
E. lactis	Forward	TACGGTCACTGG CGGTTTTT	20	274	293	60.18	50	201	1 cycle 10'-94ºC; 30 cycles 30''-94ºC, 30''-58ºC, 30''-72ºC; 1 cycle 10'-72ºC
	Reverse	TGTCTGCTGTTTC GGTAGCC	20	597	578	60.32	55		

3.2 Antibiotic susceptibility and genomic profiling of clinical *E. lactis* isolates.

A total of five clinical isolates identified as *E. lactis* with the designed primers were further analysed and identified with the following strain names: HPH55b, HPH67, HPH133, HPH282 and HPH288. The five strains were subjected to antimicrobial susceptibility, which was studied by disk diffusion against 12 antibiotics (ampicillin, vancomycin, teicoplanin, ciprofloxacin, erythromycin, gentamicin, streptomycin, linezolid, tigecycline, tetracycline, chloramphenicol and quinupristin-dalfopristin). In general, we used EUCAST criteria (version 12.0) and in cases where EUCAST did not specify a clinical breakpoint, we referred to CLSI guidelines (11). Additionally, Minimum Inhibitory Concentrations (MICs) of ampicillin were determined by E-test (Liofilchem) (12).

Genomic DNA was extracted from 1 mL of overnight cultures in brain heart infusion broth using a Wizard Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA) according to the manufacturer's instructions and the concentration was determined with a Qubit 3.0 Fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Genome sequencing was accomplished by an Illumina NovaSeq 6000 platform (2× 300 bp pair-ended runs, ~6 Gb genome, coverage 100×) according to standard Illumina protocols performed at the Eurofins Scientific (Italy). Data were analysed using: FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to test the quality of the raw and pre-processed data; SPAdes (v.3.10.0) for *de novo* assembling the paired-end reads; and QUAST (http://quast.bioinf.spbau.ru) for evaluating the quality of genome assembly. After annotation with Prokka and pan and core-genome analysis with Roary, amaximum likelihood phylogenetic tree was constructed using the core genome alignment of the 269 *E. lactis* genomes (Table S1) and the five clinical *E. lactis* genomes (this study) by employing RaxML v1.0 (13) and results were edited using iTOL (14). The Whole Genome Shotgun project including the 5 clinical *E. lactis* isolates was deposited at DDBJ/ ENA/GenBank under BioProject accession number PRJNA851953 (Biosample accession numbers: SAMN29257244-SAMN29257248).

4 Results

The first part of this study was to compute the pan-genome of the 512 genomes including 269 *E. lactis* (with 183 deposited as *E. faecium* classified as clade B and 86 deposited and classified as *E. lactis*) and 243 *E. faecium* classified as clade A, as computed by dDDH. The pan-genome was composed by 32,380 genes, with 2% representing the soft core- and core-genome, which is defined as genes present in 95-100% of the genomes (Fig.1A and 1B). The remaining 98% represented the accessory genome which is defined by shell and cloud genes (>95% of the genomes). The pan-genome analysis of the 512 genomes resulted in 12 genes with high enough variance between *E. faecium* and *E. lactis* (Table 2), of which 7 did not have a functional annotation. The remaining five genes with functional annotations included *araR* (arabinose transcriptional repressor), *gluP* (rhomboid protease), *rlmA* (23S rRNA (G(745)-N(1))-methyltransferase), *ypjD* (inner membrane protein) and *yqgN* (inner membrane protein). The alignment of genes *araR_2*, *ypjD* and *yqgN* did not show promising results as they had high allelic variability among genomes from the same set. Concerning both genes *gluP* and *rlmA,* the alignment exhibited clear patterns of allelic differences between set1 and set2 of isolates. Nevertheless, the last gene was not further explored as a ribosomal subunit-based PCR, mainly due to the presence of multiple copies that may introduce high variability and inaccuracy to the assay. Consequently, *gluP* was chosen for primer design and further screening analysis.

The alignment of *gluP* from all *E. faecium* and *E. lactis* genomes showed two different nucleotide sequences: set1/set1 and set2/set2 identities ranged from 98-100% and set1/set2 between 90-92% (Fig. 1C, Table S2). Once the primers were designed for each species-specific sequence (sequences and PCR conditions in Table 1), we submitted them to BLAST to evaluate their *in silico* accuracy with all deposited genomes on NCBI. Indeed, primer pair for clade A showed 100% identity with the corresponding sequences of clade A isolates and <90% identity with *E. lactis gluP* sequences. Similar results were obtained while analysing *E. lactis* primers. *In silico* PCR of both primers resulted in 100% specificity with *E. faecium* ATCC 700221^{T} *and E. lactis* LMG 25958^{T} type strain.

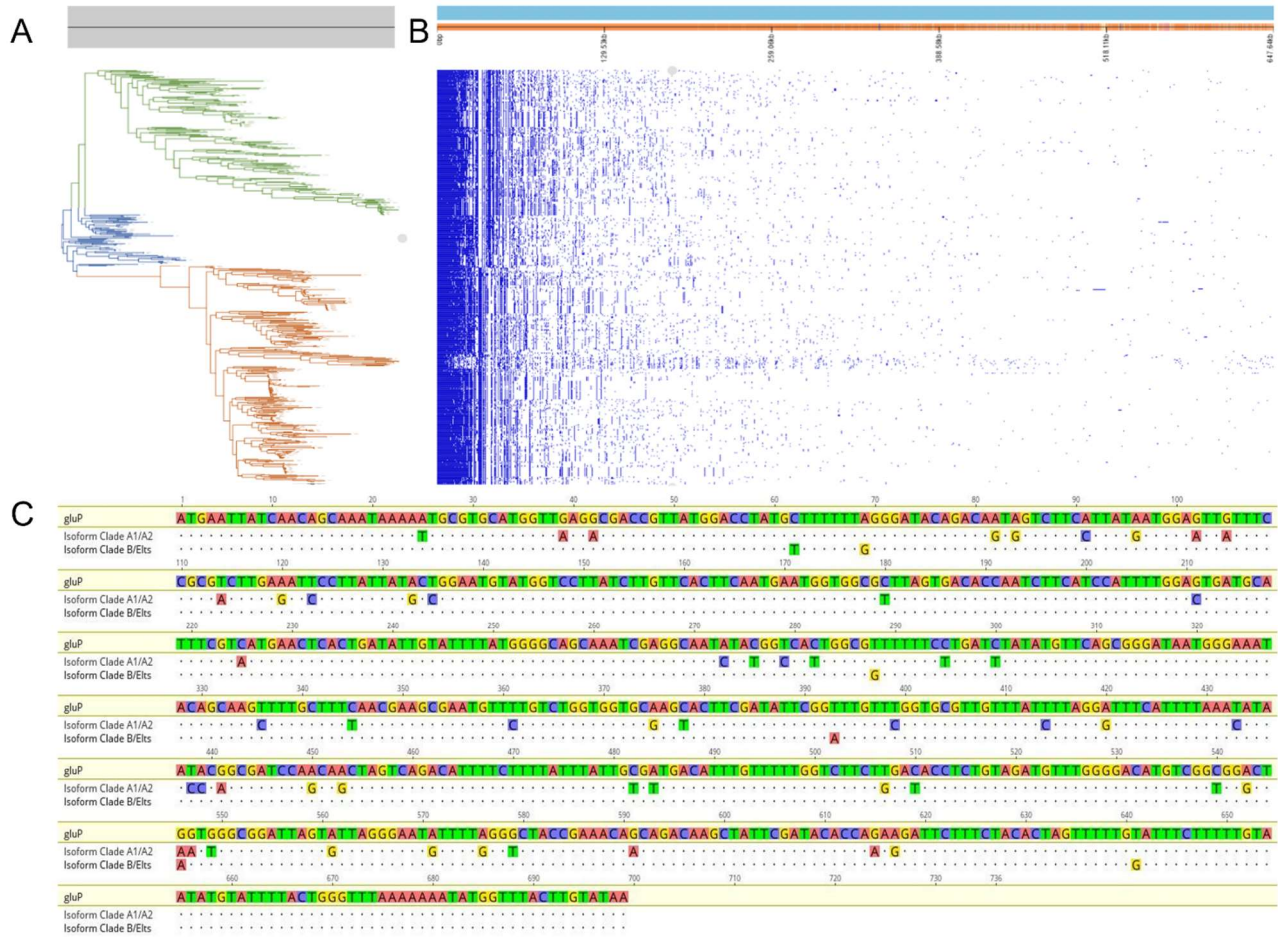


Figure 1. A) Maximum likelihood phylogenetic tree of *E. faecium* and *E. lactis* based on core genome alignment representing a clear clade separation, clade-A1 *E. faecium* (red), clade-A2 *E. faecium* (blue), and clade-B *E. faecium/E. lactis* (green). B) Presence and absence matrix of the core and accessory genomes respective to their phylogenetic position. C) Alignment of *gluP* gene showing the two main gene sequences from clade A/*E. faecium* (Genbank Accession: UDP42194.1) and *E. lactis* (Genbank Accession: WP_156271834.1) with different nucleotide patterns.

Table 2. Candidates of gene variants between *E. faecium* and *E. lactis* for primer design.

Gene	Product	Function
araR	Arabinose metabolism transcriptional repressor	Transcriptional repressor of the arabinose utilization genes.
comEA	hypothetical protein	
gluP	Rhomboid protease GluP	Rhomboid-type serine protease that catalyzes intramembrane proteolysis. Important for normal cell division and sporulation.
group_12706	hypothetical protein	
group_16273	hypothetical protein	

rimI	hypothetical protein	
group_21758	hypothetical protein	
group_21783	hypothetical protein	
group_21801	hypothetical protein	
rlmA	23S rRNA (guanine(745)-N(1))-methyltransferase	Methylation of 23S rRNA nucleotide G745
ypjD	putative protein YpjD	Inner membrane protein YpjD
yqgN	putative protein YqgN	Uncharacterized protein YpjD

We then tested and validated the PCR in enterococci collections recovered by the group in different surveillance studies over diverse time spans. PCR showed 100% accuracy when testing the 137 well-characterized *E. faecium*, where clade A isolates amplified with primer A1/A2 and *E. lactis* isolates (former clade B *E. faecium*) with primer B/*E. lactis*, exclusively (Table S3). The 27 *E. lactis* identified were obtained from human colonization (n=17), human clinical (n=5), animals (n=3) and the environment (n=2) (Table S3). At least in our dataset, *E. lactis* were generally more susceptible to antibiotics than *E. faecium* isolates (Table S3): only one *E. lactis* isolate was resistant to ampicillin (hospital sewage), two were resistant to ciprofloxacin and all remaining ones expressed resistance to erythromycin, tetracycline and/or aminoglycosides. We used the MLST *E. faecium* scheme to provide an overview of *E. lactis* clonal diversity and confirm their assignment as typical clade B *E. faecium* clones: they belonged to 18 different STs (4 novel), some previously associated with human clinical (ST74/ST108/ST123/ST329/ST361/ST798/ST994), hospital surveillance/environment (ST123/ST361/ST717), community (ST118/ST798) and animals (ST75/ST798) origins in different countries (pubmlst.org). Among the 245 isolates, amplifications were also highly specific: 98 amplified only with B/*E. lactis* primers (from human colonization), 101 (50 human clinical and 51 human colonization) only with clade A *E. faecium* primers, and the remaining (n=46) were negative for both primer pairs and then confirmed as *E. faecalis.* Most (102/180, 57%) *E. lactis* identified in all 382 enterococci (245 plus the 137) originate from human faecal colonization.

Details from the 5 clinical *E. lactis* identified in this study (2 from blood, 2 from bile and one abdominal pus) are included in Table 3. All patients but one presented co-morbidities and cholangitis/cholecystitis pathologies, for which gut bacterial translocation has been proposed as a possible cause. Indeed, these patients were co-infected with Gram-negative bacteria in 3/5 cases and underwent broad-spectrum antibiotic therapy, which is known to favour enterococci overgrowth in the lumen and possible gut translocation (15). Three out of the 5 isolates expressed resistance to erythromycin only, and the remaining were pan-susceptible. Although three of them carried some *pbp5* amino acid mutations and the Resfinder 4.1 predicted phenotype was of resistance on those cases, all presented a sensitive phenotype against ampicillin (MIC=0.05-0.75 mg/L). This may be explained by the absence of key mutations that are frequent among clinical ampicillin-resistant *E. faecium* (16). Regarding antibiotic resistance genes, only *aac(6')-Ii* and *msr(C)* genes were found, although both are

intrinsic for *E. faecium* and should be for *E. lactis* as well. Actually, we submitted to BLAST both genes against all *E. lactis* genomes and they were present in 100% of them. Even though the MLST scheme was designed for *E. faecium* and not for *E. lactis*, the 5 isolates were identified as ST118, ST329, ST361, ST994 and ST2215, and all but ST118 and the last one, that is novel, have been identified in hospitalized patients before (pubmlst.org) (4). These 5 *E. lactis* were further compared with available *E. lactis* genomes (n=269) and the resulting phylogenetic tree with the 274 *E. lactis* genomes clearly shows the intermixing of *E. lactis* from different sources with no obvious separation of isolates by source (Fig. 2). The 5 clinical *E. lactis* clustered with probiotic, dairy and animal samples. Additionally, they carried *acm*, *sgrA, ccpA, bepA, gls,* and pili genes involved in different cellular functions (Table S4), but most of them (59%) were either truncated (32%) or presented low similarity (27%) with reference strains.

Table 2. Epidemiological data and characterization of clinical *E. lactis* isolates from a Portuguese hospital in Porto area.

Isolate	ST[a]	Sex/Age	Date of isolation	Product	Pathology	Hospital Unit	Clinical case	Co-bacteria	Antibiotherapy	Co-morbilities
HPH55b	2215	M/63	11/10/19	bile	Cholangitis	Surgery	Hospitalization because of an episode of cholangitis. Antibiotic therapy with Pip+Taz. Percutaneous transhepatic cholecystectomy with bile aspiration. Bile culture: *E. coli* ESBL+ sensitive to gentamicin and ertapenem and *E. faecium*. Medicated with meropenem.	*E. coli* ESBL$^+$ sensitive to gentamicin and ertapenem	Pip+Taz Meropenem	Chronic pancreatitis of alcoholic aetiology with multiple episodes of cholangitis.
HPH67	329	M/70	28/10/19	abdominal pus	Necrotizing fasciitis	Medicine	Previous hospitalization by septic shock with abdomen necrotizing fasciitis after elective cholecystectomy. Surgery with pus collection.	*E. coli* Amoxi and cefuroxime resistant; P. *aeruginosa*	Pip+Taz and Vancomycin Addition of clindamycin	Ischemic heart disease since 1995; Chronic obstructive pulmonary disease; chronic kidney disease and gallstones.
HPH133	361	M/80	28/09/20	bile	Cholecystitis	Surgery	Laparoscopic cholecystectomy due to acute lithiasic cholecystitis. Initiates Pip+Taz. Collects bile during surgery - isolation of cefotaxime-sensitive and Pip+Taz-resistant *E. coli*; and ampicillin-sensitive *E. faecium*. Switch to cefotaxime + ampicillin. Discharge on 5/10/2020.	*E. coli* Pip/TazR	Pip+Taz	Nonrelevant
HPH282	994	M/84	25/02/22	blood	Cholangitis	Medicine	Multiple complications associated with left total hip prosthesis with prolonged hospital stay for periprosthetic infection with prosthesis extraction. Grade II acute cholangitis and prerenal AKI superimposed on CKD. Antibiotic therapy with Pip+Taz was started. Haemoculture: *E. faecium*. Excellent clinical evolution, with good response to antibiotic therapy, having completed 14 days of Pip+Taz.	No	Pip+Taz	Arterial hypertension, obesity, diabetes (type II), chronic kidney disease

HPH288	118	F/64	23/05/22	blood	Cholangitis	Surgery	During hospitalization maintained controlled pain. Progressive decrease in inflammatory parameters, without leucocytosis. Decrease in cholestasis parameters and progressive decrease in lipase and amylase, so currently no criteria for ERCP. At discharge, with innocent abdominal palpation, significant improvement in jaundice and in sustained apyrexia.	No	Ceftriaxone+ Metronidazol	Arterial hypertension, gastroesophageal reflux disease, dyslipidaemia, depressive/anxious disorder, lower limb venous insufficiency, colonic diverticulosis

AKI, acute kidney injury; Amoxi, amoxicillin; CKD, chronic kidney disease; ERCP, endoscopic retrograde cholangiopancreatography; ESBL, extended-spectrum beta-lactamase; F, female; M, male; Pip, piperacillin; R, resistant; S, susceptible; ST, sequence type; Taz, tazobactam.

[a]ST was defined according to the MLST scheme of *E. faecium* as there is any about *E. lactis*.

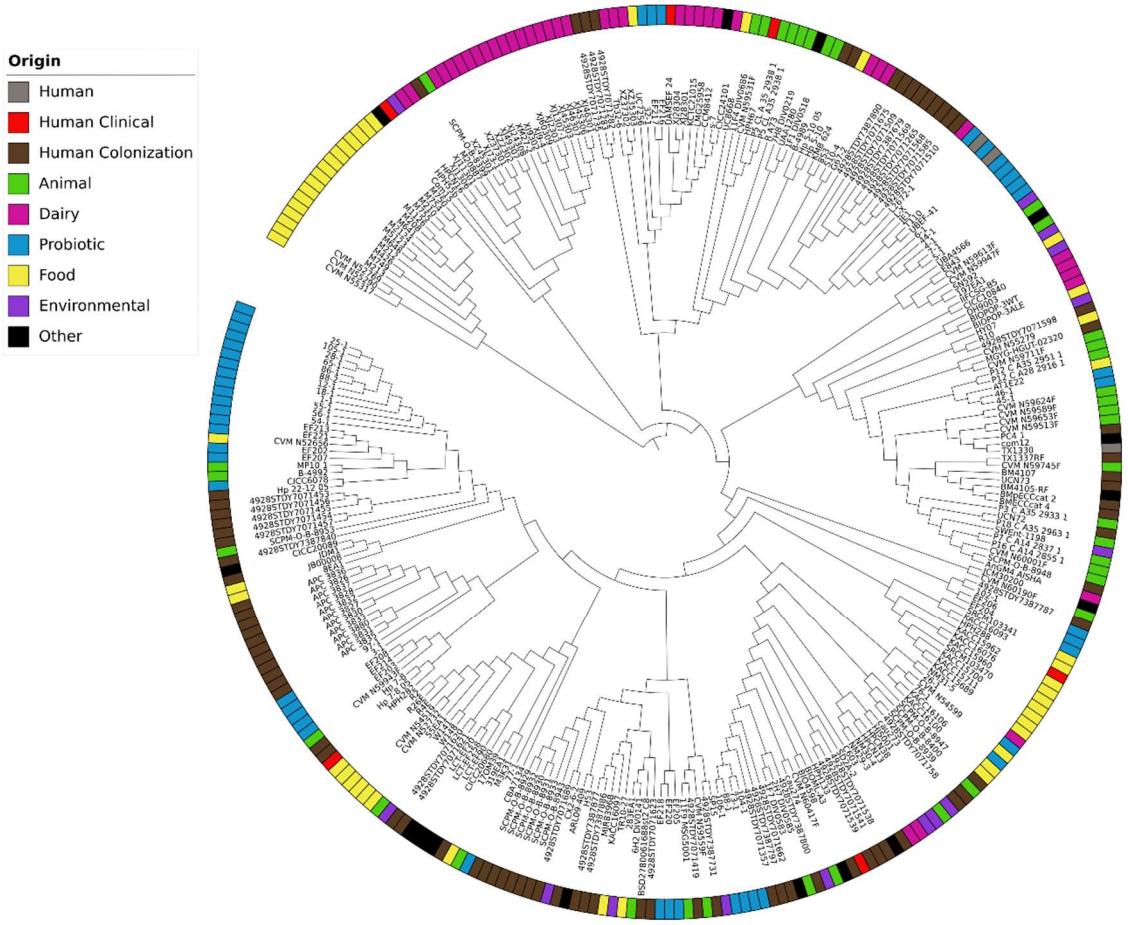


Figure 2. Maximum likelihood phylogenetic tree of 274 *E. lactis* genomes based on core genome alignment. Different isolation origins are classified as follows: human from clinical samples in red, human from colonization samples (stool, gastrointestinal, genitourinary and breast milk samples) in brown, non-determined origin of human samples in grey, animal isolates in green, food in yellow, dairy in magenta, probiotics in blue, environmental samples in purple and other origins (unknown) in black. No clear patterns between isolates from different sources is visible. Clinical *E. lactis* samples are clustered among probiotic/dairy and animal samples.

5 Discussion

The emergence of multidrug-resistant *E. faecium* in hospitals causing tenacious and hard-to-treat infections has been alarming over the last decades and has intensified the need to distinguish strains of public health concern. Currently, some of the enterococci causing hospital infections are being misidentified worldwide. In this work, we corroborate a recent proposal that a subset of *E. faecium* (clade B) are actually *E. lactis* (3) and further designed primers to correctly differentiate between these species for accurate identification. The primer design was based on the pan-genome alignment of both species aiming to find unique genes or gene variations that were sufficiently discriminatory to differentiate them through standard PCR. Gene *gluP*, coding for a rhomboid protease, showed two different sequences with enough nucleotide pattern differences to design species-specific primers. Rhomboid family proteases are a ubiquitous family of intramembrane serine proteases, with a unique evolutionary conservation level (17). Different studies have been conducted to investigate the structure and function of rhomboid proteases, especially AarA in Providencia stuartii (role in quorum sensing), GlpG in Escherichia coli and Haemophilus influenzae (role in antibiotic susceptibility) and GluP (also called YqgQ) in Bacillus subtilis (role in cell division and glucose uptake) (18). The function and structure description of GluP within the Bacillota phylum may suggest the potential function of GluP in enterococci, however this exceeds the scope of this study (19). More research will unveil the phenotypic impact of the allelic differences between species, however, for the scope of this study only its genotypic variation was considered.

E. lactis is genomically and evolutionarily distinct from *E. faecium*. Phenotypically, they are generally much more susceptible to antibiotics and lack key virulence markers known to be associated with outbreak/epidemic *E. faecium* (this study; 3,8). Even though *E. lactis* seem less prone to cause human infections, clade B *E. faecium* have been described as able to acquire the VanA (4) or VanN operon (20), and their proportion among human infections caused by enterococci may be undervalued since most surveillance studies focus on MDR *E. faecium* strains. According to the features of the clinical *E. lactis* detected in this and other (previously described as clade B *E. faecium*) studies, and their great association with human faecal colonization, we believe that *E. lactis*, as one dominant human gut species, can translocate the gut barrier in severely ill, immunodeficient and/or surgical patients. Indeed, all patients infected by *E. lactis* in this study present at least a risk factor for bacterial translocation (chronic diseases as pancreatitis, abdominal surgeries, broad-spectrum antibiotics). As the ability of different enterococcal species to translocate into host tissues seems evolutionarily-related (21), it makes sense that *E. lactis* is able to do it as well. Previous studies showed enterococci enriched in the faecal microbiome of patients with sclerosing cholangitis together with gram-negative bacteria (22) and that is one of the commonest genera in bile cultures (23), so more research is needed to unveil the amount of *E. lactis* versus other enterococcal species in these and other clinical cases. In common to previous studies describing *E. lactis* in association with bloodstream infections (4, 10, 24), here we describe two bacteraemia cases by *E. lactis* with clinical significance and systemic signals of infection. One limitation

of our study is the low sample, but future large-scale studies will unveil the real ability of *E. lactis* to cause bacteraemia and other infections as well as the best antibiotherapy to treat them.

6 Conclusion

To conclude, we designed and validated a PCR to discriminate between *E. faecium* and *E. lactis* species. To note that published primers widely used for years and designed to identify *E. faecium* (e.g., ddl gene) lack enough discriminatory power to distinguish these species. Very recent approaches to differentiate them by MALDI-TOF MS or qPCR showed promising results (25, 26), but until we have a robust collection of *E. lactis* mass spectra for hospital routine identification and other purposes, we have successfully designed a highly specific PCR that can be applied in a cost-effective and timely fashion. The development of a precise differentiation method has direct implications in both the clinical and food safety fields and could draw the line between *E. faecium* strains currently being used in probiotics and feed that actually correspond to *E. lactis* and/or associated with human infections when they are actually *E. lactis* having possible implications in infection management and overall, in different Public Health contexts.

7 Transparency Declaration

The authors do not report any conflict of interest in relation to this work. This work was funded by national funds from FCT—Fundação para a Ciência e a Tecnologia, I.P., in the scope of the project UIDP/04378/2020 and UIDB/04378/2020 of the Research Unit on Applied Molecular Biosciences—UCIBIO, the project LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy—i4HB, and the exploratory project EXPL/SAU-INF/0261/2021. ARF and A.C.A.-S. gratefully acknowledge the Junior Research Position (CEECIND/02268/2017 - Individual Call to Scientific Employment Stimulus 2017) and the UI/BD/151317/2021 fellowship, respectively, both granted by FCT/MCTES through national funds.

8 References

1. Gao W, Howden BP, Stinear TP. 2018. Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen. Curr Opin Microbiol. Elsevier Ltd https://doi.org/10.1016/j.mib.2017.11.030.
2. Freitas AR, Pereira AP, Novais C, Peixe L. 2021. Multidrug-resistant high-risk Enterococcus faecium clones: can we really define them? Int J Antimicrob Agents 57:106227.
3. Belloso Daza MV, Cortimiglia C, Bassi D, Cocconcelli PS. 2021. Genome-based studies indicate that the Enterococcus faecium Clade B strains belong to Enterococcus lactis species and lack of the hospital infection associated markers. Int J Syst Evol Microbiol 71.
4. Raven KE, Reuter S, Reynolds R, Brodrick HJ, Russell JE, Török ME, Parkhill J, Peacock SJ. 2016. A decade of genomic history for healthcare-associated Enterococcus faecium in the United Kingdom and Ireland. Genome Res 26:1388–1396.
5. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
6. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.
7. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115.
8. Freitas AR, Tedim AP, Novais C, Coque TM, Peixe L. 2018. Distribution of putative virulence markers in Enterococcus faecium: towards a safety profile review. J Antimicrob Chemother 73:306–319.
9. CLSI. 2020. Performance Standards for Antimicrobial Susceptibility Testing A CLSI supplement for global application. Clin Lab Stand Inst 1–296.
10. Freitas AR, Tedim AP, Almeida-Santos AC, Duarte B, Elghaieb H, Abbassi MS, Hassen A, Novais C, Peixe L. 2022. High-Resolution Genotyping Unveils Identical Ampicillin-Resistant Enterococcus faecium Strains in Different Sources and Countries: A One Health Approach. Microorganisms 10:1–13.
11. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455.
12. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. Nucleic Acids Res 47:W256–W259.
13. Archambaud C, Derré-Bobillot A, Lapaque N, Rigottier-Gois L, Serror P. 2019. Intestinal translocation of enterococci requires a threshold level of enterococcal overgrowth in the lumen. Sci Reports 2019 91 9:1–12.
14. Freitas AR, Tedim AP, Almeida-Santos AC, Duarte B, Elghaieb H, Abbassi MS, Hassen A, Novais C, Peixe L. 2022. High-Resolution Genotyping Unveils Identical Ampicillin-Resistant Enterococcus faecium Strains in Different Sources and Countries: A One Health Approach. Microorganisms 10.

15. Koonin E V, Makarova KS, Rogozin IB, Davidovic L, Letellier M-C, Pellegrini L. 2003. The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers. Genome Biol 4:R19.
16. Rather P. 2013. Role of rhomboid proteases in bacteria. Biochim Biophys Acta 1828:2849–54.
17. Began J, Cordier B, Březinová J, Delisle J, Hexnerová R, Srb P, Rampírová P, Kožíšek M, Baudet M, Couté Y, Galinier A, Veverka V, Doan T, Strisovsky K. 2020. Rhomboid intramembrane protease YqgP licenses bacterial membrane protein quality control as adaptor of FtsH AAA protease. EMBO J 39.
18. Lebreton F, Valentino MD, Schaufler K, Earl AM, Cattoir V, Gilmore MS. 2018. Transferable vancomycin resistance in clade B commensal-type Enterococcus faecium. J Antimicrob Chemother 73:1479–1486.
19. Fine RL, Manfredo Vieira S, Gilmore MS, Kriegel MA. 2020. Mechanisms and consequences of gut commensal translocation in chronic diseases. Gut Microbes 11:217–230.
20. Little R, Wine E, Kamath BM, Griffiths AM, Ricciuto A. 2020. Gut microbiome in primary sclerosing cholangitis: A review. World J Gastroenterol 26:2768.
21. Ozturk-Engin D, Agalar C, Cag Y, Can FK, Balkan II, Karabay O, Senbayrak S, Çetinkaya BM, Aydın MT, Tomas K, Disci E, Surmelioglu A, Alimoglu O, Ekinci O, Akın E, Köroglu M, Velidedeoglu M, Ankaralı H, Kocoglu E, Javadov M, Papilla-Kundaktepe B, Oguzoglu N, Ozmen E, Donmez R, Mega E, Aksaray S, Agalar F. 2022. Microorganisms isolated from the bile of the patients who have undergone cholecystectomy and their antibiotic resistance pattern: multicenter prospective study. Int Microbiol https://doi.org/10.1007/S10123-022-00251-Y.
22. Tedim AP, Ruíz-Garbajosa P, Rodríguez MC, Rodríguez-Baños M, Lanza VF, Derdoy L, Zurita GC, Loza E, Cantón R, Baquero F, Coque TM. 2017. Long-term clonal dynamics of Enterococcus faecium strains causing bloodstream infections (1995–2015) in Spain. J Antimicrob Chemother 72:48–55.
23. Kim E, Yang S-M, Kim H-J, Kim H-Y. 2022. Differentiating between Enterococcus faecium and Enterococcus lactis by Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry. Foods 11:1046.

9 Supplementary materials

Table S1. List of strains of *E. faecium* and *E. lactis* selected for in silico analyses and their respective Genbank Accession numbers.

Strain	Clade	Accession Number	dDDH against *E. faecium* ATCC 700221T (%)
DO	A1	GCA_000174395.2	86.6
Aus0004	A1	GCA_000250945.1	85.6
Aus0085	A1	GCA_000444405.1	81.6
6E6	A1	GCA_001518735.1	85.9
UW8175	A1	GCA_001587115.1	83.9
ATCC700221	A1	GCA_001594345.1	100
E39	A1	GCA_001635875.1	88.4
ISMMS_VRE_1	A1	GCA_001720945.1	85.6
ISMMS_VRE_7	A1	GCA_001721065.1	88
ISMMS_VRE_11	A1	GCA_001721905.1	90.7
E745	A1	GCA_001750885.1	92.8
E1	A1	GCA_001886635.1	89.2
VRE001	A1	GCA_001895905.1	84.5
ISMMS_VRE_12	A1	GCA_001953235.1	89.2
ISMMS_VRE_9	A1	GCA_001953255.1	89.1
2014-VREF-41	A1	GCA_002007625.1	81.1
2014-VREF-114	A1	GCA_002024245.1	84.4
2014-VREF-268	A1	GCA_002025045.1	80.4
2014-VREF-63	A1	GCA_002025065.1	81.6
K60-39	A1	GCA_002334625.1	86
A_020709_82	A1	GCA_002442955.1	92.3
E240	A1	GCA_002761255.1	89.6
E243	A1	GCA_002761275.1	89.5
E232	A1	GCA_002777275.1	89.6
SC4	A1	GCA_002848385.1	84.1
Efaecium_ER03933.3A	A1	GCA_002848625.1	88.4
Efaecium_ER04120.3A	A1	GCA_002848645.1	89.1
Efaecium_ER04484.3A	A1	GCA_002848665.1	89.1
Efaecium_ER04462.3A	A1	GCA_002848685.1	89.1
Efaecium_ER04619.3A	A1	GCA_002848705.1	88.3
Efaecium_ER04526.5A	A1	GCA_002848725.1	89.1
Efaecium_ER04526.3A	A1	GCA_002848745.1	88
AALTL	A1	GCA_002880635.1	89.5
15-307-1	A1	GCA_002973755.2	78.6
AUSMDU00004167	A1	GCA_003020685.1	82.7
AUSMDU00004055	A1	GCA_003020705.1	84.1
AUSMDU00004028	A1	GCA_003020725.1	88.3

AUSMDU00004024	A1	GCA_003020745.1	85
AUSMDU00004142	A1	GCA_003020765.1	89
CFSAN059070	A1	GCA_003071425.1	85.5
CFSAN059071	A1	GCA_003071445.1	85.5
RBWH1	A1	GCA_003957785.1	84.9
VB3025	A1	GCA_005517315.1	86.1
VB3240	A1	GCA_005576735.1	81.5
UAMSEF_01	A1	GCA_005886545.1	85.1
UAMSEF_09	A1	GCA_005886715.1	83.8
UAMSEF_20	A1	GCA_005886735.1	83.8
HOU503	A1	GCA_005952885.1	88.8
VRE1	A1	GCA_006007925.1	82.6
DB-1	A1	GCA_006337045.1	82.2
FB-1	A1	GCA_006351785.1	81.9
515	A1	GCA_006575625.1	88.3
VVEswe-R	A1	GCA_007917035.3	85.7
VVEswe-S	A1	GCA_007917315.3	85.9
V1836	A1	GCA_008728455.1	82.8
V2937	A1	GCA_008728475.1	86.4
LAC7.2	A1	GCA_009036045.1	80.6
VRE	A1	GCA_009697285.1	87.2
SRR24	A1	GCA_009734005.2	79.1
ZY11	A1	GCA_009938075.1	84.5
KUHS13	A1	GCA_009938285.1	81.4
ZY2	A1	GCA_010120755.1	81.6
BA17124	A1	GCA_012932975.2	84.4
BP5067	A1	GCA_012932985.2	84.6
BP3378	A1	GCA_012933055.2	83.6
BP657	A1	GCA_012933075.2	85.1
A7214	A1	GCA_012933165.2	85
A6521	A1	GCA_012933195.2	85.7
A4694	A1	GCA_012933245.2	83.5
A3895	A1	GCA_012933265.2	90.4
A11051	A1	GCA_012933285.2	81.3
A15023	A1	GCA_012933295.2	90.2
A10290	A1	GCA_012933345.2	85.9
AML0157	A1	GCA_014490015.1	79.4
VB3338	A1	GCA_014874615.1	83.6
VRE3370	A1	GCA_015325925.1	84.3
Cairo	A1	GCA_015356095.1	91.2
VRE3363	A1	GCA_015476295.1	88
VRE3389	A1	GCA_015999405.1	80.2

VRE3382	A1	GCA_015999425.1	79.1
Dallas1	A1	GCA_015999605.1	90
Dallas5	A1	GCA_016126675.1	89.9
Dallas91	A1	GCA_016406345.1	90
Dallas53_1	A1	GCA_016406365.1	90.1
Dallas55	A1	GCA_016406385.1	90.1
Dallas66	A1	GCA_016406405.1	92.4
Dallas83	A1	GCA_016406425.1	90.1
Dallas17_1	A1	GCA_016406445.1	87.4
Dallas51_4	A1	GCA_016406465.1	87.2
Dallas16_1	A1	GCA_016406485.1	81.6
Dallas87_1	A1	GCA_016406505.1	87.9
Dallas71_2	A1	GCA_016406525.1	89.2
Dallas53_2	A1	GCA_016406545.1	90.1
Dallas57	A1	GCA_016406565.1	89.9
Dallas93_3	A1	GCA_016406585.1	90
Dallas93_2	A1	GCA_016406605.1	89.4
Dallas97_1	A1	GCA_016415025.1	89.8
Dallas103	A1	GCA_016415105.1	88.8
Dallas100_1	A1	GCA_016415285.1	90.9
Dallas107_1	A1	GCA_016415325.1	89.7
Dallas111	A1	GCA_016415345.1	84.3
Dallas124_3	A1	GCA_016415365.1	89.9
Dallas124_1	A1	GCA_016415385.1	89.9
Dallas144_1	A1	GCA_016415405.1	86.1
Dallas148	A1	GCA_016415425.1	90
Dallas137	A1	GCA_016415445.1	87.6
Dallas131_2	A1	GCA_016415465.1	90
Dallas163_1	A1	GCA_016415485.1	81.4
Dallas155	A1	GCA_016415505.1	89.9
Dallas154_1	A1	GCA_016415545.1	83.8
Dallas158	A1	GCA_016415565.1	84
XJ11301	A1	GCA_016642695.1	79
WGS1811-4-7	A1	GCA_016864255.1	88.1
AUSMDU00011555	A1	GCA_017301355.1	82.8
VRE3355	A1	GCA_017584065.1	84
VB13828	A1	GCA_017603725.1	80.7
VB12993	A1	GCA_017815655.1	81.4
VB3895	A1	GCA_017815675.1	90.4
VB976	A1	GCA_017815695.1	80.1
VB6171	A1	GCA_017897965.1	85.7
VB6521	A1	GCA_017898005.1	85.7

VB3378	A1	GCA_017898025.1	83.6
PR01996-12	A1	GCA_018219325.1	92.8
18-276	A1	GCA_018516925.1	85.6
16-021	A1	GCA_018517105.1	84.6
18-204	A1	GCA_018517145.1	89.3
18-042	A1	GCA_018517185.1	82.4
M20887	A1	GCA_018531665.1	87.4
7150	A1	GCA_019356355.1	80.3
AA620	A1	GCA_019977495.1	79
AA622	A1	GCA_019977575.1	78.6
HJP554	A1	GCA_020091325.1	85.7
V1225	A1	GCA_020162155.1	82.1
V1164	A1	GCA_020162175.1	81.9
4995-20	A1	GCA_020221735.1	82.6
SC1762	A1	GCA_020736585.1	79.4
SC1762-D	A1	GCA_020736625.1	79.4
V13-21-E11-012-001	A1	GCA_021172105.1	76.2
NMVRE-001	A1	GCA_021228615.1	87.8
EFE11651	A1	GCA_900044005.1	92.9
EFE10021	A1	GCA_900066025.1	72.8
Ef_aus00233	A1	GCA_900092475.1	82.9
Ef_DMG1500501	A1	GCA_900094185.1	82.3
E6043	A1	GCA_900635415.1	79.7
E1774	A1	GCA_900638785.1	85.2
E4402	A1	GCA_900638805.1	85.9
E7067	A1	GCA_900639335.1	85.2
E6055	A1	GCA_900639345.1	83.4
E7025	A1	GCA_900639355.1	80
E7171	A1	GCA_900639365.1	81.6
E7040	A1	GCA_900639385.1	84.6
E6975	A1	GCA_900639395.1	86.5
E4457	A1	GCA_900639405.1	82.9
E6988	A1	GCA_900639415.1	79.6
E7098	A1	GCA_900639425.1	82.2
E7199	A1	GCA_900639445.1	87.4
E7237	A1	GCA_900639455.1	81.6
E7429	A1	GCA_900639465.1	89.8
E7240	A1	GCA_900639485.1	86.5
E7471	A1	GCA_900639495.1	84.4
E7356	A1	GCA_900639505.1	87.4
E7654	A1	GCA_900639515.1	89.5
E7663	A1	GCA_900639525.1	89.4

E8202	A1	GCA_900639535.1	89.3
E7933	A1	GCA_900639545.1	87
E8172	A1	GCA_900639555.1	85.1
E8195	A1	GCA_900639565.1	88.9
E8014	A1	GCA 900639575.1	84.8
E7948	A1	GCA_900639585.1	75.9
E8377	A1	GCA_900639595.1	86.7
E8290	A1	GCA 900639605.1	85.7
E8328	A1	GCA_900639615.1	87.5
E8284	A1	GCA_900639625.1	86.8
E8927	A1	GCA_900639635.1	80.2
E8414	A1	GCA_900639715.1	84.2
E8423	A1	GCA_900639725.1	84.2
E4456	A1	GCA_900639745.1	87.5
AUS2001	A1	GCA_907163255.1	86.9
AUS2002	A1	GCA_907163315.1	83
USZ_VRE32_P32	A1	GCA_907165365.1	84.3
USZ_VRE5_P5	A1	GCA_907176135.1	84.4
USZ_VRE67_P60	A1	GCA_907176815.1	84.3
USZ VRE53 P46	A1	GCA 907177285.1	81.8
NRRLB-2354	A2	GCA_000336405.1	74.6
UC7251	A2	GCA_000411655.2	81
64-3	A2	GCA 001298485.1	76
LS170308	A2	GCA_002831505.1	71.4
WEFA23	A2	GCA_002850515.1	74.6
FDAARGOS_323	A2	GCA_002983785.1	80.7
HPCN16	A2	GCA_003173545.1	91.9
FSIS1608820	A2	GCA_004332055.1	76
NM213	A2	GCA_005166365.1	77.6
UAMSEF_08	A2	GCA_005886655.1	85.1
F17E0263	A2	GCA_006280355.1	73.4
N56454	A2	GCA_006351845.1	71.4
QU 50	A2	GCA_006741355.1	70.9
SCPM-O-B-8399	A2	GCA 012935525.	76.8
XM27-2	A2	GCA_016642205.1	72.3
XJ1307-1	A2	GCA_016642495.1	78.2
XJ60309	A2	GCA 016642565.1	77.9
XJ46307	A2	GCA_016642585.1	71.3
XJ49307	A2	GCA_016642595.1	70.7
XJ46301	A2	GCA 016642605.1	71.4
XJ45306	A2	GCA_016642615.1	71.4
XJ45303	A2	GCA_016642665.1	71.3

XJ9302	A2	GCA_016642685.1	79.7
XJ1306	A2	GCA_016642705.1	74.4
XJ24308	A2	GCA_016642715.1	70.7
XJ2303	A2	GCA_016642765.1	78.3
XJ73-1	A2	GCA_016642785.1	75.8
XJ9-4	A2	GCA_016642825.1	70.5
XZ37301	A2	GCA_016643025.1	75.9
XZ45301	A2	GCA_016643045.1	75.8
XZ2302	A2	GCA_016643075.1	75.8
fac90	A2	GCA_016743855.1	75.9
PR05720-3	A2	GCA_018219285.1	78.3
116	A2	GCA_018279145.1	75.1
18-465	A2	GCA_018516845.1	71.3
18-133	A2	GCA_018517025.1	80.7
17-508	A2	GCA_018517045.1	80.7
17-318	A2	GCA_018517065.1	76
16-164	A2	GCA_018517085.1	74.8
18-201	A2	GCA_018517165.1	79.6
F88	A2	GCA_019175425.1	74.7
F39	A2	GCA_019175445.1	72
F179	A2	GCA_019175465.1	75.3
AVS0243	A2	GCA_019175525.1	73.4
VBO96	A2	GCA_019456555.1	73.6
VBO39	A2	GCA_019456575.1	73.7
VBR48	A2	GCA_019456595.1	73.6
E843xGE-1-TC1	A2	GCA_019774555.1	79.1
E0139	A2	GCA_900634805.1	74.2
E1334	A2	GCA_900635365.1	78.7
NCTC7174	A2	GCA_900637035.1	74.1
E2079	A2	GCA_900638765.1	72.3
E4227	A2	GCA_900638775.1	70.9
E4413	A2	GCA_900638795.1	76.4
E0595	A2	GCA_900638815.1	73.9
E4438	A2	GCA_900638825.1	77.8
E8691	A2	GCA_900639655.1	78.3
E9101	A2	GCA_900639705.1	72
E4239	A2	GCA_900640265.1	70.9
3012STDY6244127	A2	GCA_900683475.1	78.7
com12	B	GCA_000157635.1	63.7
TX1330	B	GCA_000159675.1	62.6
PC4_1	B	GCA_000178235.1	61.6
LCT-EF90	B	GCA_000258325.1	62.9

TX1337RF	B	GCA_000294345.2	59.6
LCT-EF20	B	GCA_000313155.1	63
LCT-EF258	B	GCA_000313195.1	63
UAA1280	B	GCA_000393735.1	58.8
SD2A-2	B	GCA 000415285.2	69.1
UC8668	B	GCA_000499905.1	65.6
UC7256	B	GCA_000499925.1	62.8
T110	B	GCA 000737555.1	64.4
L-X	B	GCA_000787065.1	64.8
KACC15960	B	GCA_001025245.1	65.1
KACC15700	B	GCA_001025255.1	64.8
KACC15689	B	GCA_001025265.1	64.4
KACC15962	B	GCA_001025315.1	63.6
KACC16076	B	GCA_001025325.1	66.6
KACC16093	B	GCA_001025335.1	62.5
KACC16097	B	GCA_001025375.1	62
KACC16100	B	GCA_001025385.1	62.5
KACC16106	B	GCA_001025405.1	63.3
KACC15711	B	GCA_001025435.1	63.7
M3K31	B	GCA 001039515.1	64.4
MJR8396B	B	GCA_001546375.1	62.5
17OM39	B	GCA_001652715.1	64.6
UCN73	B	GCA 001696275.1	59.8
BM4107	B	GCA_001696285.1	60
UCN72	B	GCA_001696305.1	55.7
11F4_DIV0686	B	GCA_002140385.1	61.1
7H8_DIV0219	B	GCA_002140865.1	53.5
11F9_MSG5001	B	GCA_002141115.1	61
1F1_DIV0518	B	GCA_002141175.1	57.2
1F7_DIV0583	B	GCA_002141255.1	64.9
6H2_DIV0141	B	GCA_002141355.1	63.9
2H7_DIV0585	B	GCA_002174445.1	65
BMpECCcat_2	B	GCA_002263125.1	64.5
BMECCcat 4	B	GCA 002263205.1	64.5
UBA4566	B	GCA_002387065.1	64.5
JB00008	B	GCA_002591965.2	64.9
Hp 7-8	B	GCA 002631225.1	67.1
Hp_5-10	B	GCA_002631295.1	66.1
CVM_N59653F	B	GCA_002944125.1	61.9
CVM N60190F	B	GCA 002944305.1	60.8
CVM_N60001F	B	GCA_002944615.1	62.7
CVM_N59624F	B	GCA_002944675.1	61.4

CVM_N59613F	B	GCA_002944735.1	63.9
CVM_N59745F	B	GCA_002944915.1	60.5
CVM_N59711F	B	GCA_002944955.1	62.2
CVM_N59947F	B	GCA_002944975.1	66.4
CVM_N59943F	B	GCA_002945095.1	68.5
CVM_N60417F	B	GCA_002945395.1	66.7
CVM_N59531F	B	GCA_002945495.1	60
CVM_N59559F	B	GCA_002945575.1	66.2
CVM_N59589F	B	GCA_002946055.1	61.4
CVM_N59513F	B	GCA_002946315.1	60.6
CVM_N52656	B	GCA_002946855.1	64.4
CVM_N55317	B	GCA_002947295.1	66.1
CVM_N55290	B	GCA_002947355.1	64.9
CVM_N55279	B	GCA_002947365.1	64.5
CVM_N54599	B	GCA_002947775.1	63.4
CVM_N52769	B	GCA_002948695.1	64.8
CVM_N52732	B	GCA_002948755.1	65.9
CVM_N54519	B	GCA_002948895.1	64.3
HPCN13	B	GCA_003172835.1	57
BM4105-RF	B	GCA_003269465.1	60.3
197EA1	B	GCA_003320195.1	65.8
283EA1	B	GCA_003320395.1	64.9
319EA1	B	GCA_003320555.1	64.3
55EA1	B	GCA_003320585.1	64.5
8EA1	B	GCA_003320815.1	63.6
KMB_624	B	GCA_003346195.1	66
HY07	B	GCA_003574925.1	63.7
JE1	B	GCA_003667965.1	65.9
Gr17	B	GCA_003711605.	64.6
P18_C_A35_2963_1	B	GCA_003795795.1	59.3
P5_CL_A35_2938_1	B	GCA_003795905.1	60
P5_CL_A_35_2938_1	B	GCA_003795915.1	61.3
P12_C_A28_2916_1	B	GCA_003795975.1	63.1
P7_C_A14_2837_1	B	GCA_003796465.1	60.6
P12_C_A35_2951_1	B	GCA_003796745.1	63.1
P3_C_A35_2933_1	B	GCA_003796865.1	59.8
P16_C_A14_2855_1	B	GCA_003797365.1	60.6
8S3	B	GCA_003862415.1	65.5
Com15	B	GCA_004006255.1	61
CBA7134	B	GCA_004015145.1	66.1
SP15	B	GCA_004101305.1	66.8
SRCM103341	B	GCA_004101385.1	65.5

SRCM103470	B	GCA_004103475.1	64.9
UAMSEF_24	B	GCA_004300245.1	64.1
UBEF-41	B	GCA_005116555.1	65.4
ARL09-409	B	GCA_005234835.1	68.4
BSD2780061688st2_C8	B	GCA_005844995.1	62.2
HB-1	B	GCA_006337145.1	66.2
R10	B	GCA_006375675.1	61.7
SN592	B	GCA_006375815.1	66.4
C603	B	GCA_006376115.1	67.8
TR10-27	B	GCA_006376275.1	65.4
W148	B	GCA_006541485.1	63.4
R2	B	GCA_006541615.1	64.2
W141	B	GCA_006541625.1	63.8
R26E	B	GCA_006541645.1	64.3
R4E	B	GCA_006541685.1	64.2
DMEA02	B	GCA_008330605.1	64.1
R_A73	B	GCA_008365415.1	62.7
BIOML-A3	B	GCA_009891505.1	68.2
DT1-1	B	GCA_011745645.1	64.4
FA3	B	GCA_011801455.1	64.5
BIOPOP-3ALE	B	GCA_012045365.1	65
BIOPOP-3WT	B	GCA_012045505.1	65
SCPM-O-B-8400	B	GCA_012935425.1	65.3
FS 86	B	GCA_013201055.1	64.4
SWEnt-1198	B	GCA_013248515.1	62.1
BIO4598	B	GCA_013249095.1	68.2
cau273	B	GCA_013280535.1	64.6
cau274	B	GCA_013280545.1	66.1
M750	B	GCA_013371105.1	67.4
M749	B	GCA_013371115.1	65.6
M5123	B	GCA_013371125.1	65.8
M755	B	GCA_013371155.1	67.7
M754	B	GCA_013371185.1	67.4
M641	B	GCA_013371205.1	66.2
M2124	B	GCA_013371245.1	65.9
M648	B	GCA_013371255.1	65.8
M213	B	GCA_013371305.1	63.6
M426	B	GCA_013371315.1	65.5
M212	B	GCA_013371345.1	65.5
M210	B	GCA_013371355.1	67.4
M1126	B	GCA_013371365.1	65.7
M208	B	GCA_013371415.1	65.7

IIFCSG-B5	B	GCA_014050505.1	64.4
TK-P5D	B	GCA_015377765.1	65.9
Hp_5-10_05	B	GCA_900143335.1	65.3
Hp_7-8_05	B	GCA_900143385.1	66.5
Hp 22-12 05	B	GCA 900143455.1	63.4
AT1E22	B	GCA_900166945.1	62.8
4928STDY7071357	B	GCA_902159035.1	66.8
4928STDY7071454	B	GCA 902159225.1	65.6
4928STDY7071453	B	GCA_902159235.1	67.6
4928STDY7071419	B	GCA_902159245.1	65.4
4928STDY7071541	B	GCA_902159305.1	65.6
4928STDY7071538	B	GCA_902159325.1	65.6
4928STDY7071598	B	GCA_902159345.1	62.3
4928STDY7071662	B	GCA_902159505.1	68.1
4928STDY7071455	B	GCA_902160635.1	65.7
4928STDY7071456	B	GCA_902160645.1	67.6
4928STDY7071457	B	GCA_902160705.1	67.5
4928STDY7071510	B	GCA_902161035.1	65.2
4928STDY7071509	B	GCA_902161175.1	65.2
4928STDY7071539	B	GCA 902161395.1	65.6
4928STDY7071568	B	GCA_902161615.1	65.2
4928STDY7071569	B	GCA_902161635.1	65.2
4928STDY7071585	B	GCA 902161725.1	65.2
4928STDY7071620	B	GCA_902161965.1	64.7
4928STDY7071623	B	GCA_902161985.1	66.1
4928STDY7071621	B	GCA_902162005.1	64
4928STDY7071265	B	GCA_902163095.1	65.2
4928STDY7071675	B	GCA_902163145.1	65.2
4928STDY7071689	B	GCA_902163255.1	65.3
4928STDY7071758	B	GCA_902164005.1	66.1
4928STDY7387679	B	GCA_902164045.1	65.1
4928STDY7387731	B	GCA_902164525.1	67.5
4928STDY7387787	B	GCA_902165155.1	66.4
4928STDY7387797	B	GCA 902165375.1	67
4928STDY7387840	B	GCA_902165765.1	64
4928STDY7387877	B	GCA_902165855.1	64.8
4928STDY7387886	B	GCA 902165955.1	64.7
4928STDY7387890	B	GCA_902165995.1	65
4928STDY7071283	B	GCA_902166385.1	56.9
4928STDY7071282	B	GCA 902166525.1	57
4928STDY7071735	B	GCA_902166655.1	56.8
4928STDY7387800	B	GCA_902166835.1	59.1

MGYG-HGUT-02320	B	GCA_902385775.1	61.9
CICC6078	L	GCA_009735405.1	64.4
CICC20089	L	GCA_009735435.1	65.5
CICC10840	L	GCA_009735445.1	66.5
CICC20680	L	GCA_009735475.1	63
CICC24101	L	GCA_009735495.1	65
APC_3837	L	GCA_009896605.1	65.1
APC_3836	L	GCA_009896635.1	65.1
APC_3835	L	GCA_009896685.1	65.1
APC_3831	L	GCA_009896695.1	65.1
APC_3833	L	GCA_009896715.1	65.1
APC_3830	L	GCA_009896725.1	65.1
APC_3832	L	GCA_009896765.1	65
APC_3828	L	GCA_009897405.1	65.1
APC_3880	L	GCA_009897565.1	65.1
APC_3827	L	GCA_009897695.1	65.1
APC_3826	L	GCA_009897725.1	65.1
s-7	L	GCA_013867815.1	65.2
CCM 8412	L	GCA_015751045.1	64.5
JCM 30200	L	GCA_015751065.1	61.5
L2672-1	L	GCA_015751085.1	66.3
KCTC 21015	L	GCA_015767715.1	63.2
LMG 25958	L	GCA_015904215.1	64.1
HPCN38	L	GCA_016599235.1	67.1
XJ28304	L	GCA_016767495.1	66.5
G67-2	L	GCA_016767515.1	65.2
Tb32-6	L	GCA_016767535.1	68
S10-4	L	GCA_016767545.1	65.3
XJ28301	L	GCA_016767575.1	66.5
XZ37302	L	GCA_016767595.1	67
XZ35303	L	GCA_016767615.1	67.1
NM31-5	L	GCA_016767635.1	65.6
NM30-4	L	GCA_016767645.1	65
NM29-3	L	GCA_016767675.1	64.4
AnGM4_AISHA	L	GCA_016863785.1	65.3
H53	L	GCA_017942505.1	65.2
B-4989	L	GCA_018069745.1	64.9
B-4492	L	GCA_018069825.1	64.5
105-1	L	GCA_018397225.1	64.4
106-1	L	GCA_018397255.1	67
102-1	L	GCA_018397275.1	64.8
104-1	L	GCA_018397285.1	67

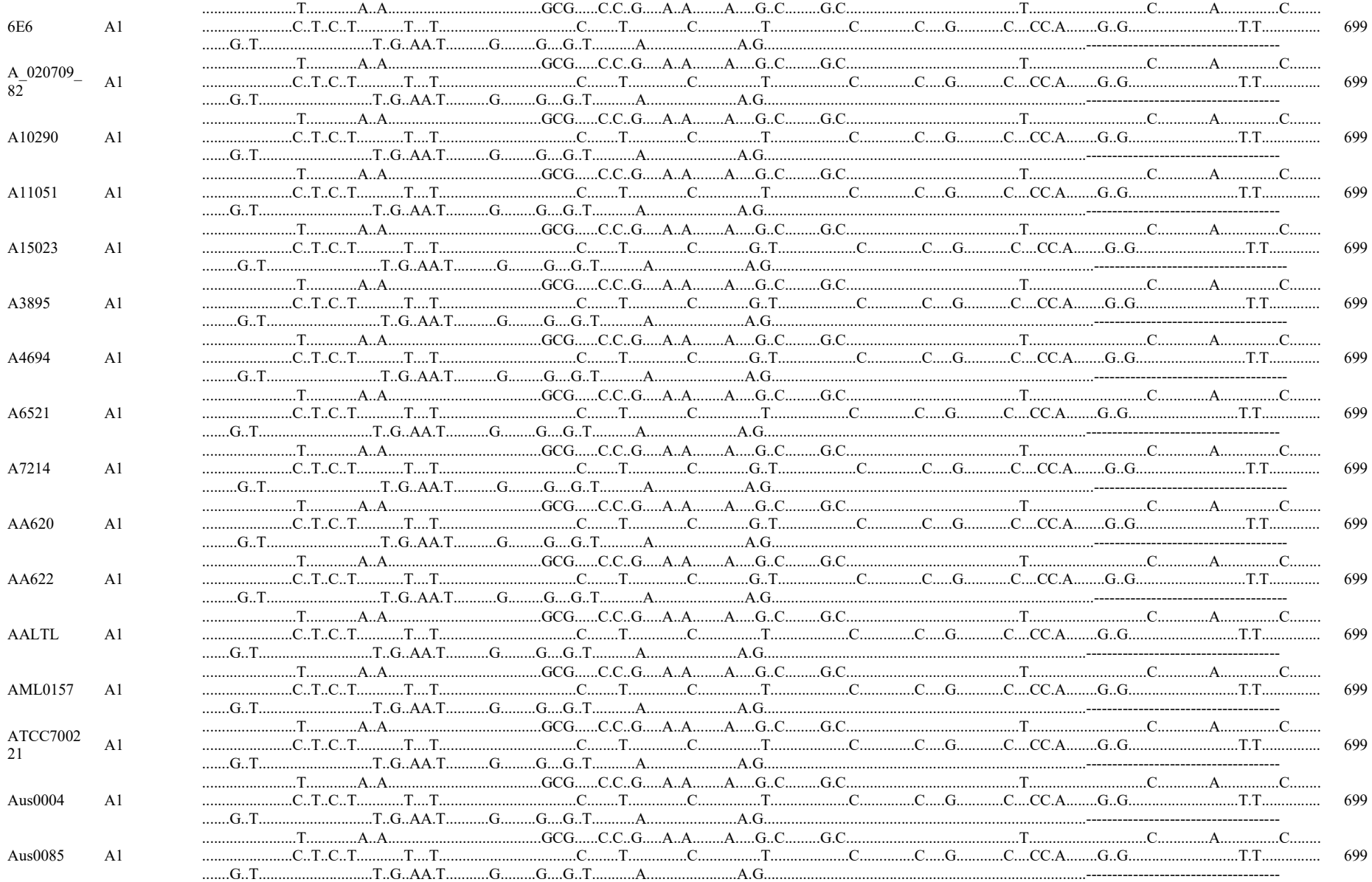
88-2	L	GCA_018397315.	64.4
97-1	L	GCA_018397325.1	64.6
77-1	L	GCA_018397355.1	63.2
86-1	L	GCA_018397375.1	64.4
76-1	L	GCA 018397385.1	65.2
88-1	L	GCA_018397395.1	67
73-1	L	GCA_018397435.1	67
56-1	L	GCA 018397455.1	64.4
55-1	L	GCA_018397485.1	64.4
65-1	L	GCA_018397515.1	64.4
54-1	L	GCA_018397535.1	64.4
46-1	L	GCA_018397545.1	64.9
44-1	L	GCA_018397575.1	66.2
45-1	L	GCA_018397595.1	64.9
32-1	L	GCA_018397615.1	65.5
26-1	L	GCA_018397625.1	66.3
28-1	L	GCA_018397655.1	64.4
18-1	L	GCA_018397675.1	64.4
25-1	L	GCA_018397695.1	64.4
7-1	L	GCA 018397715.1	66.2
6-1	L	GCA_018397725.1	63.5
12-1	L	GCA_018397745.1	64.4
4-1	L	GCA 018397765.1	66.2
5-1	L	GCA_018397795.1	66.2
1-1	L	GCA_018397815.1	64.4
3-1	L	GCA_018397835.1	65.1
JDM1	L	GCA_019203145.1	66.3
CX-6-2	L	GCA_019343125.1	64.3
EF220	L	GCA_019659145.1	65.8
EF218	L	GCA_019659185.1	65.7
EF216	L	GCA_019659205.1	63
EF221	L	GCA_019662145.	64.1
EF214	L	GCA_019662205.1	66.8
EF217	L	GCA 019662215.1	60.7
EF213	L	GCA_019662245.1	62.4
EF208	L	GCA_019662335.1	66.2
EF206	L	GCA 019662365.1	66.6
EF207	L	GCA_019662385.1	64.2
EF203	L	GCA_019662395.1	67.2
EF205	L	GCA 019662405.1	66
EF204	L	GCA_019662445.1	66.2
EF202	L	GCA_019662465.1	64

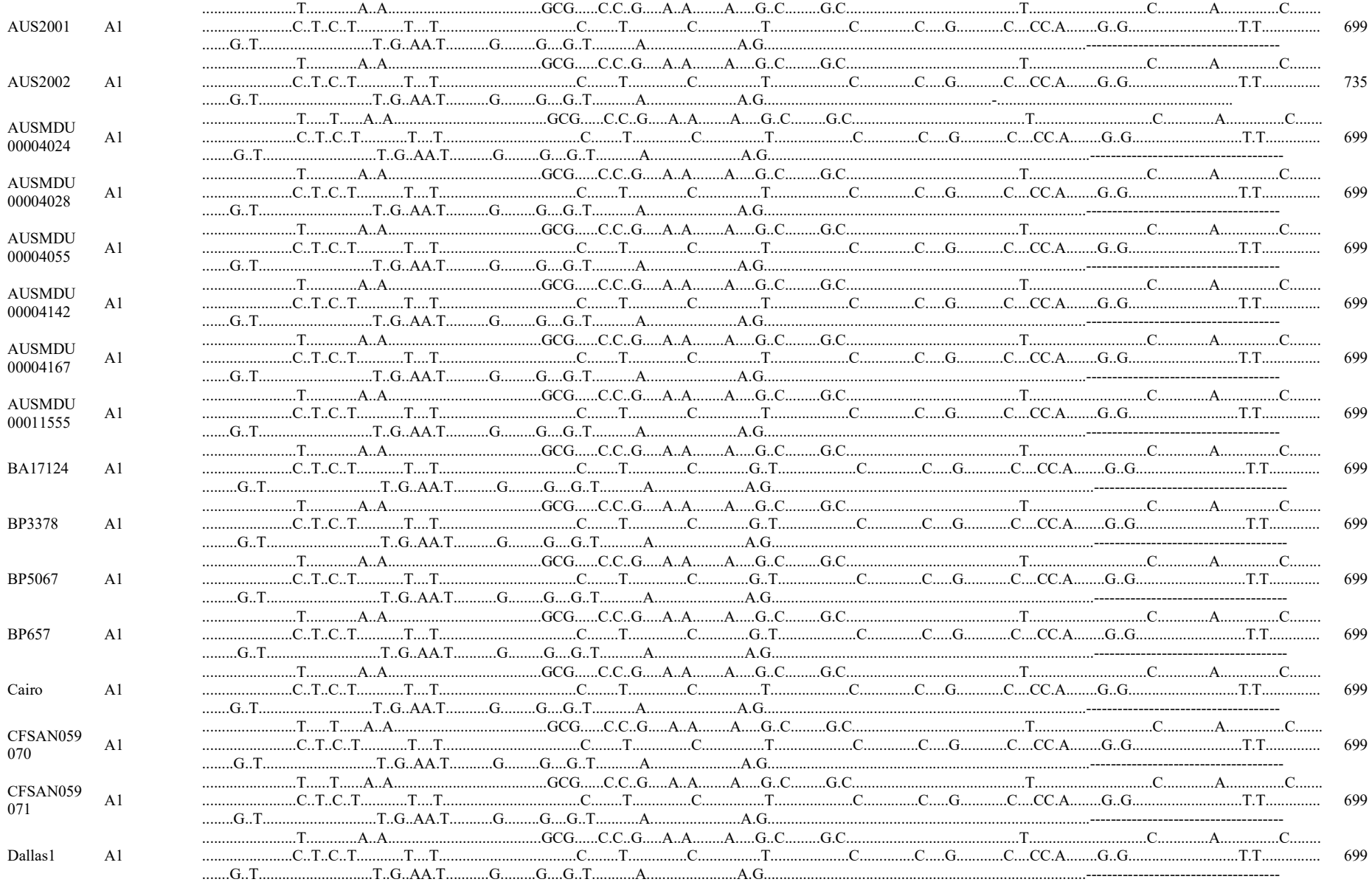
E843	L	GCA_019880345.1	66.6
HJS001	L	GCA_019967715.1	66.9
DH9003	L	GCA_020268645.1	63.4
SCPM-O-B-8947	L	GCA_020405655.1	68.9
SCPM-O-B-8939	L	GCA 020405675.1	66.3
SCPM-O-B-8948	L	GCA_020405685.1	64.2
SCPM-O-B-8931	L	GCA_020405775.1	66.8
SCPM-O-B-8932	L	GCA 020405785.1	66.9
SCPM-O-B-8943	L	GCA_020405815.1	66.8
SCPM-O-B-8929	L	GCA_020405935.1	66.8
SCPM-O-B-8930	L	GCA_020405965.1	66.9
SCPM-O-B-8933	L	GCA_020406915.1	66.8
SCPM-O-B-8953	L	GCA_020406935.1	65.3

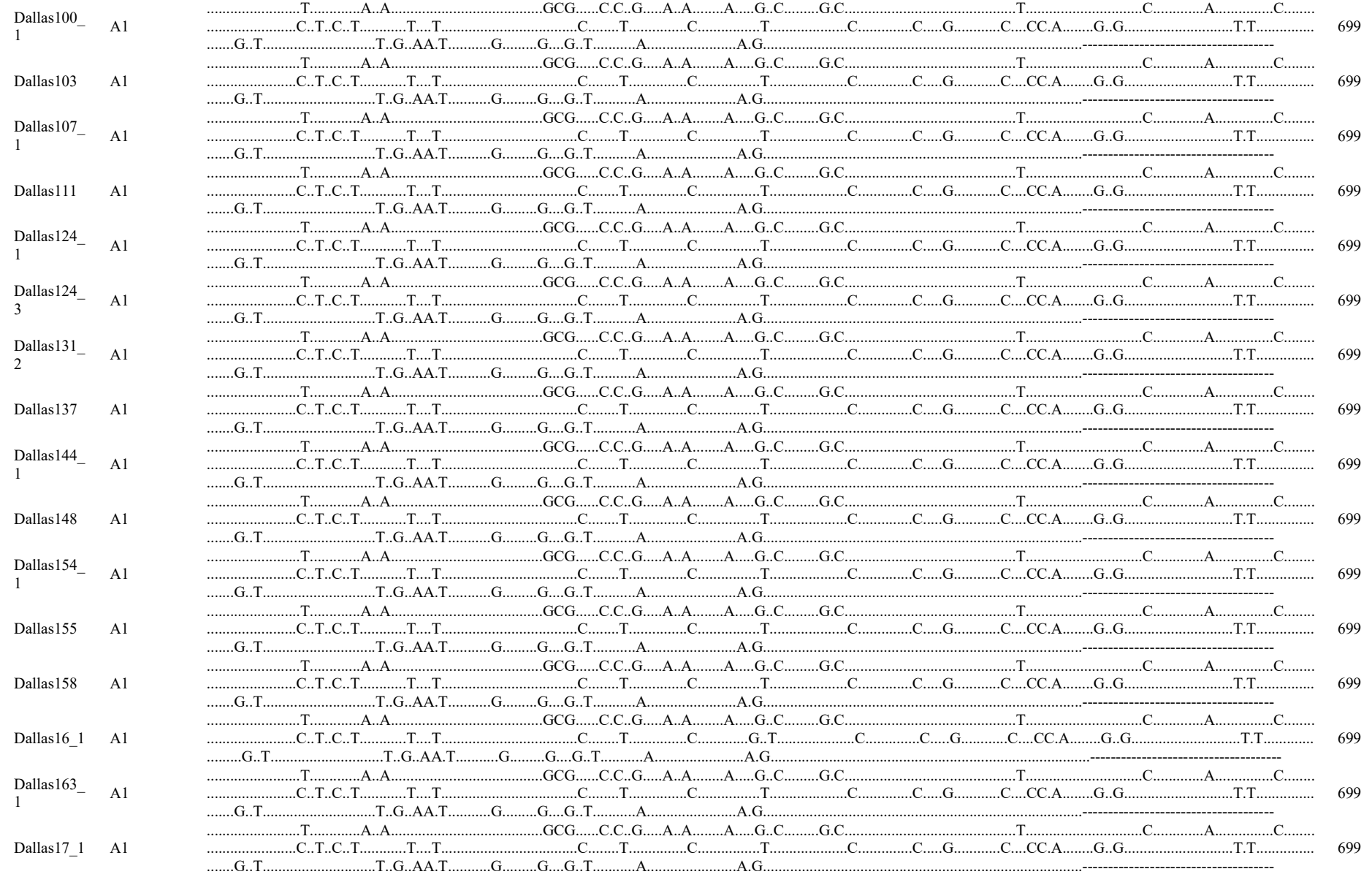
A1, A2 and B refer to *E. faecium* clades and L is referred to *E. lactis.*

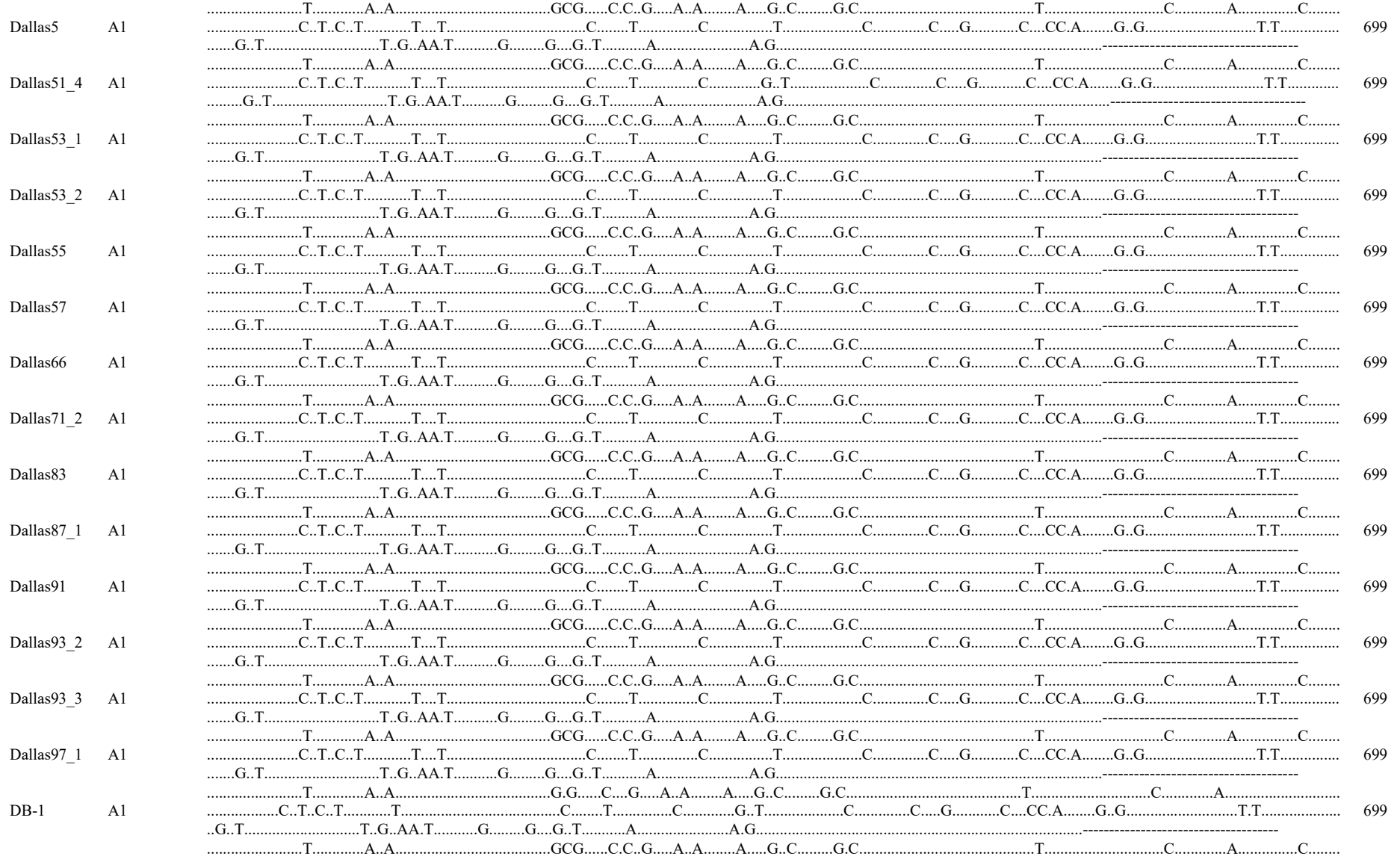
Table S2. Gene alignment of *gluP* across the 512 genomes included in this study.

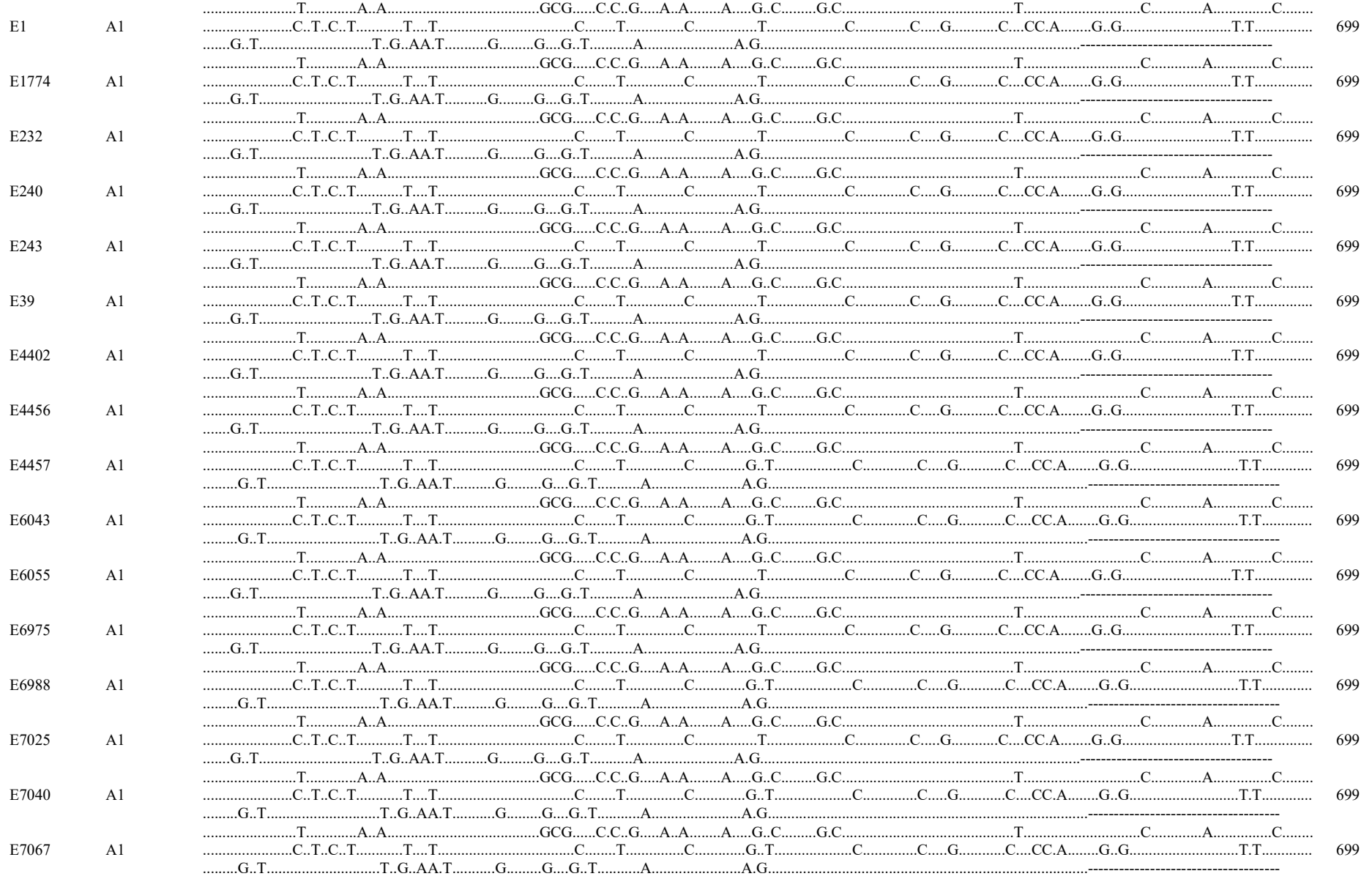
Strain	Clade	Reference Sequence: ATGAATTATCAACAGCAAATAAAAATGCGTGCATGGTTGAGGCGACCGTTATGGACCTATGCTTTTTTAGGGATACAGACAATAGTCTTCATTATAATG GAGTTGTTTCCGCGTCTTGAAATTCCTTATTATACTGGAATGTATGGTCCTTATCTTGTTCACTTCAATGAATGGTGGCGCTTAGTGACACCAATCTTC ATCCATTTTGGAGTGATGCATTTCGTCATGAACTCACTGATATTGTATTTTATGGGGCAGCAAATCGAGGCAATATACGGTCACTGGCGTTTTTTCCTG ATCTATATGTTCAGCGGGATAATGGGAAATACAGCAAGTTTTGCTTTCAACGAAGCGAATGTTTTGTCTGGTGGTGCAAGCACTTCGATATTCGGTTTG TTTGGTGCGTTGTTTATTTTAGGATTTCATTTTAAATATAATACGGCGATCCAACAACTAGTCAGACATTTTCTTTTATTTATTGCGATGACATTTGTTT TTGGTCTTCTTGACACCTCTGTAGATGTTTGGGGACATGTCGGCGGACTGGTGGGCGGATTAGTATTAGGGAATATTTTAGGGCTACCGAAACAGCAG ACAAGCTATTCGATACACCAGAAGATTCTTTCTACACTAGTTTTTGTATTTCTTTTTGTAATATGTATTTTACTGGGTTTAAAAAAATATGGTTTACTTG TATAATATGAAGGAAGTGCTTAAAGATGGAGAGCTTATATGA	736 (bp)
7150	A1	.....T.....A..A.....GC.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
15-307-1	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
16-021	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
18-042	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
18-204	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....-	735
18-276	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
2014-VREF-114	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
2014-VREF-268	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
2014-VREF-41	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
2014-VREF-63	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
4995-20	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
515	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699

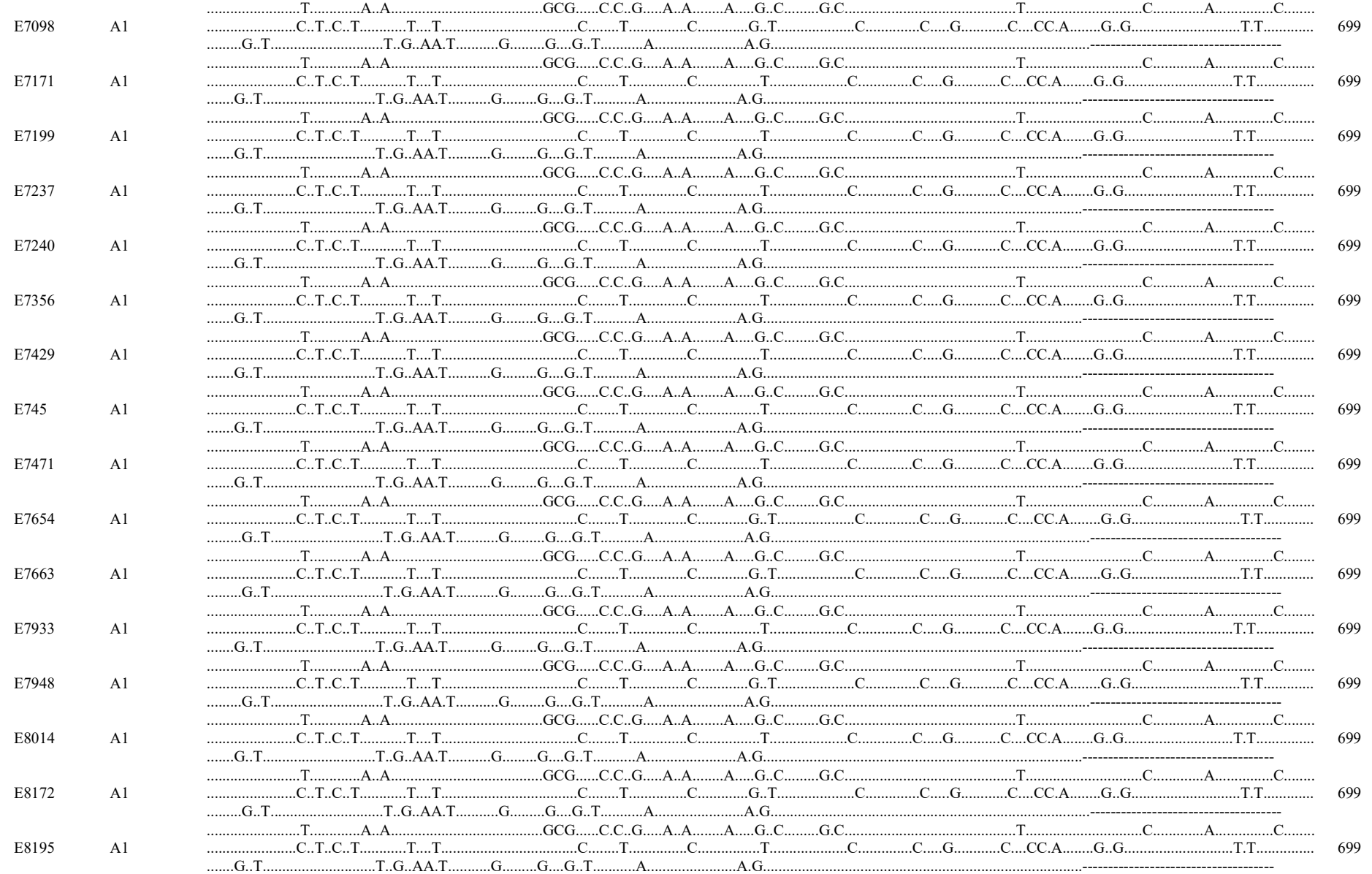

6E6 A1 699
A_020709_82 A1 699
A10290 A1 699
A11051 A1 699
A15023 A1 699
A3895 A1 699
A4694 A1 699
A6521 A1 699
A7214 A1 699
AA620 A1 699
AA622 A1 699
AALTL A1 699
AML0157 A1 699
ATCC700221 A1 699
Aus0004 A1 699
Aus0085 A1 699

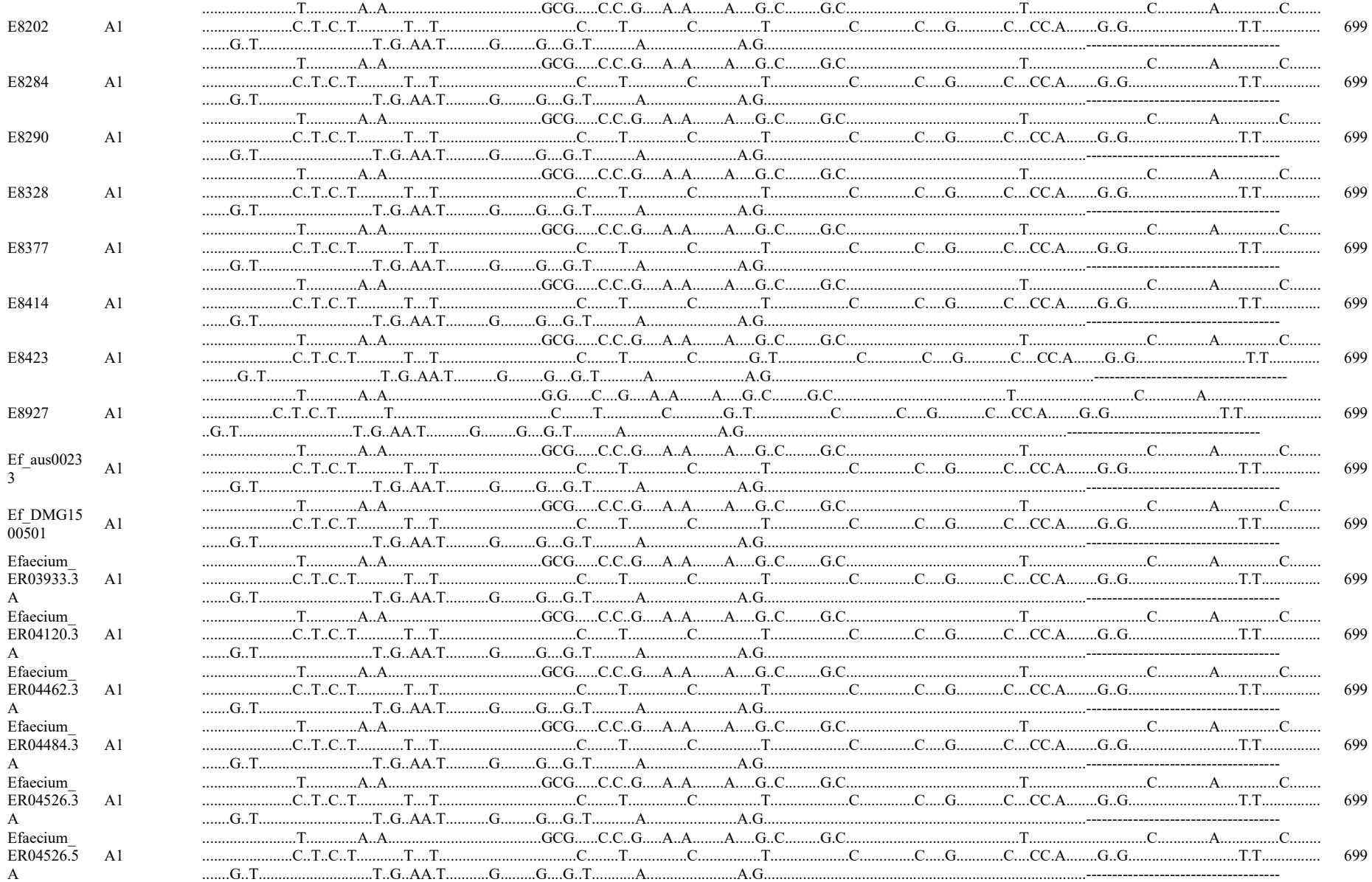

AUS2001 A1 699
AUS2002 A1 735
AUSMDU 00004024 A1 699
AUSMDU 00004028 A1 699
AUSMDU 00004055 A1 699
AUSMDU 00004142 A1 699
AUSMDU 00004167 A1 699
AUSMDU 00011555 A1 699
BA17124 A1 699
BP3378 A1 699
BP5067 A1 699
BP657 A1 699
Cairo A1 699
CFSAN059 070 A1 699
CFSAN059 071 A1 699
Dallas1 A1 699


Dallas100_1 A1 699
Dallas103 A1 699
Dallas107_1 A1 699
Dallas111 A1 699
Dallas124_1 A1 699
Dallas124_3 A1 699
Dallas131_2 A1 699
Dallas137 A1 699
Dallas144_1 A1 699
Dallas148 A1 699
Dallas154_1 A1 699
Dallas155 A1 699
Dallas158 A1 699
Dallas16_1 A1 699
Dallas163_1 A1 699
Dallas17_1 A1 699

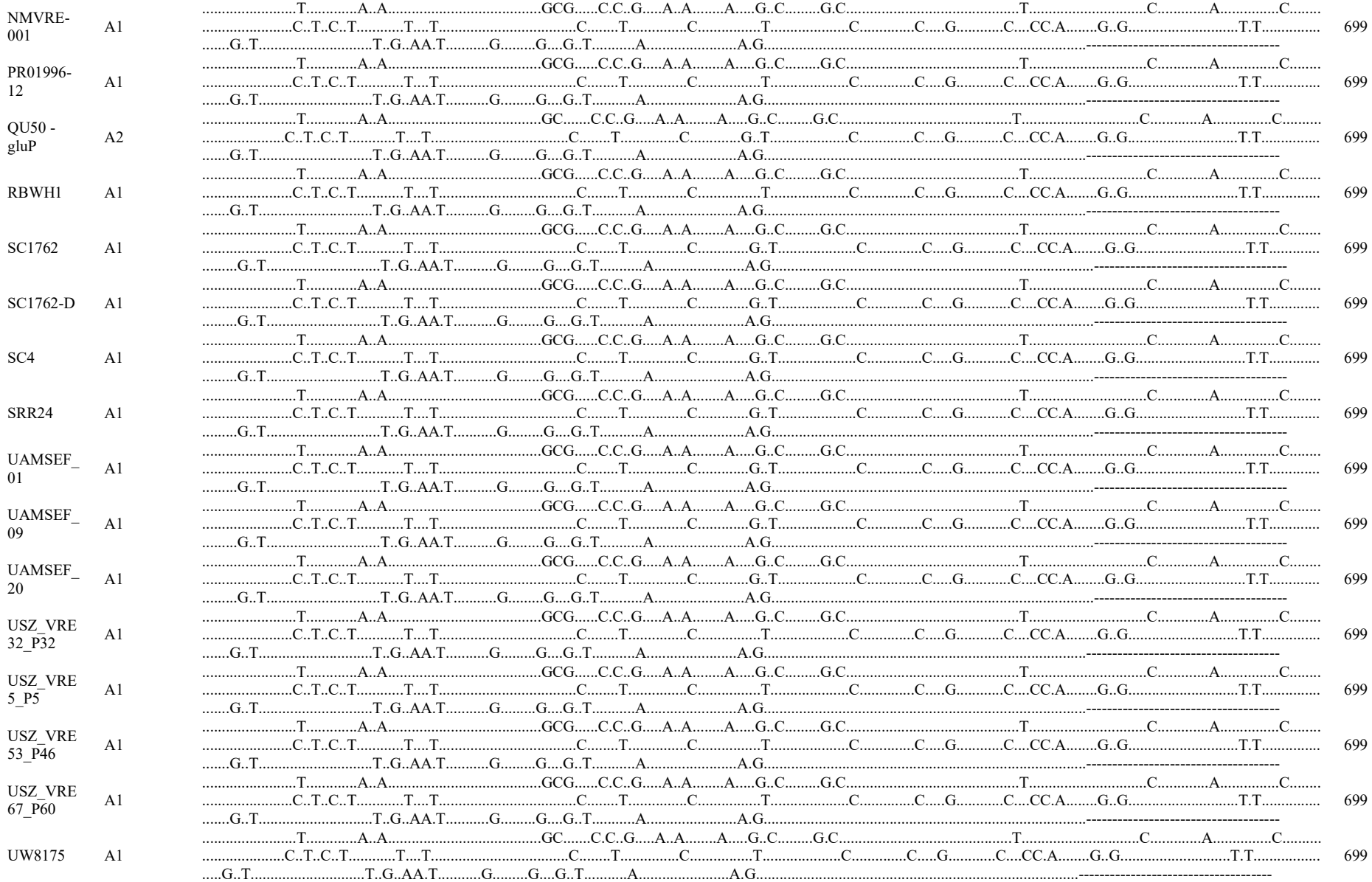

Dallas5 A1 699
Dallas51_4 A1 699
Dallas53_1 A1 699
Dallas53_2 A1 699
Dallas55 A1 699
Dallas57 A1 699
Dallas66 A1 699
Dallas71_2 A1 699
Dallas83 A1 699
Dallas87_1 A1 699
Dallas91 A1 699
Dallas93_2 A1 699
Dallas93_3 A1 699
Dallas97_1 A1 699
DB-1 A1 699
DO A1 699

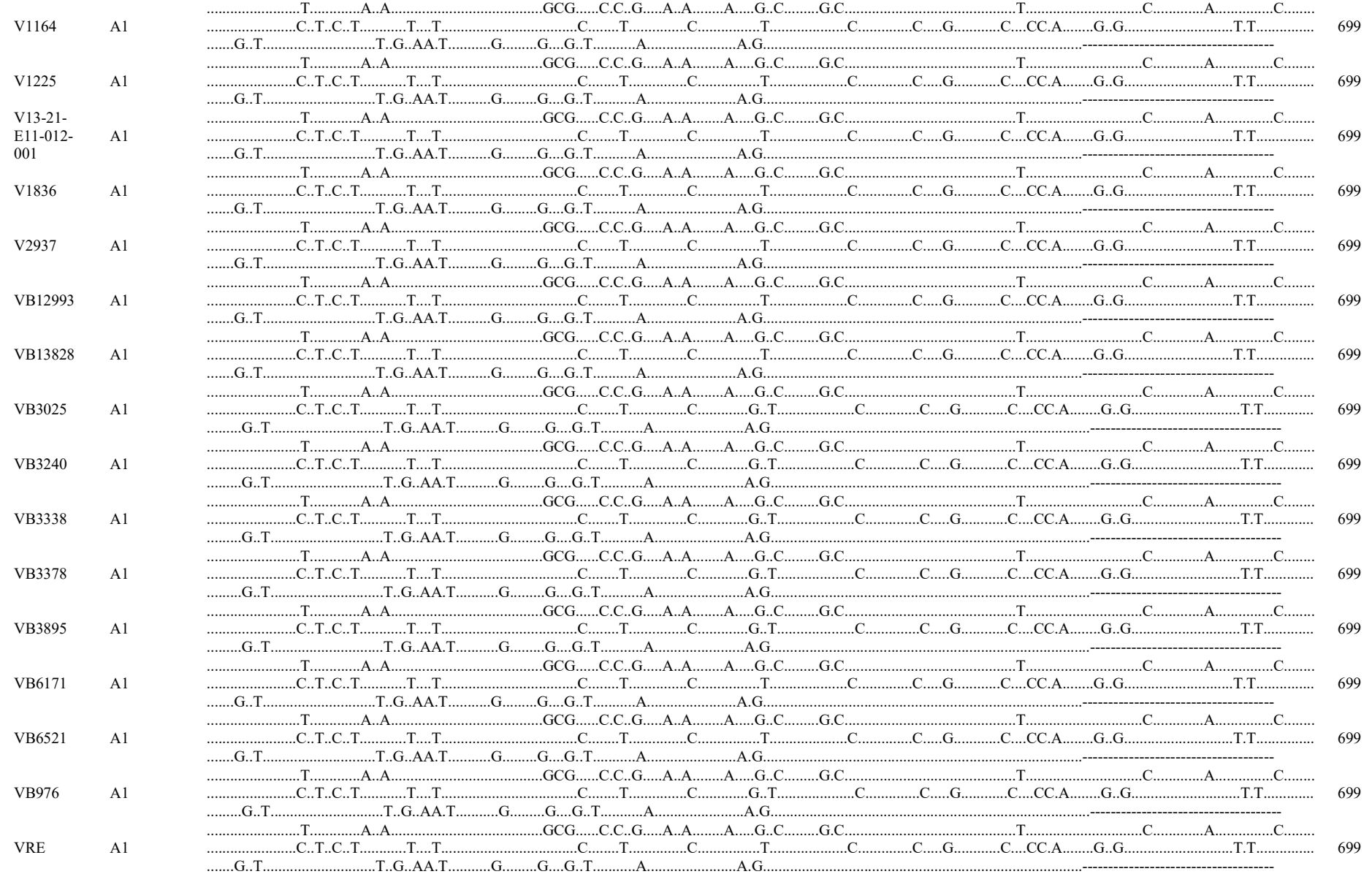

E1 A1 699
E1774 A1 699
E232 A1 699
E240 A1 699
E243 A1 699
E39 A1 699
E4402 A1 699
E4456 A1 699
E4457 A1 699
E6043 A1 699
E6055 A1 699
E6975 A1 699
E6988 A1 699
E7025 A1 699
E7040 A1 699
E7067 A1 699


E7098 A1 699
E7171 A1 699
E7199 A1 699
E7237 A1 699
E7240 A1 699
E7356 A1 699
E7429 A1 699
E745 A1 699
E7471 A1 699
E7654 A1 699
E7663 A1 699
E7933 A1 699
E7948 A1 699
E8014 A1 699
E8172 A1 699
E8195 A1 699

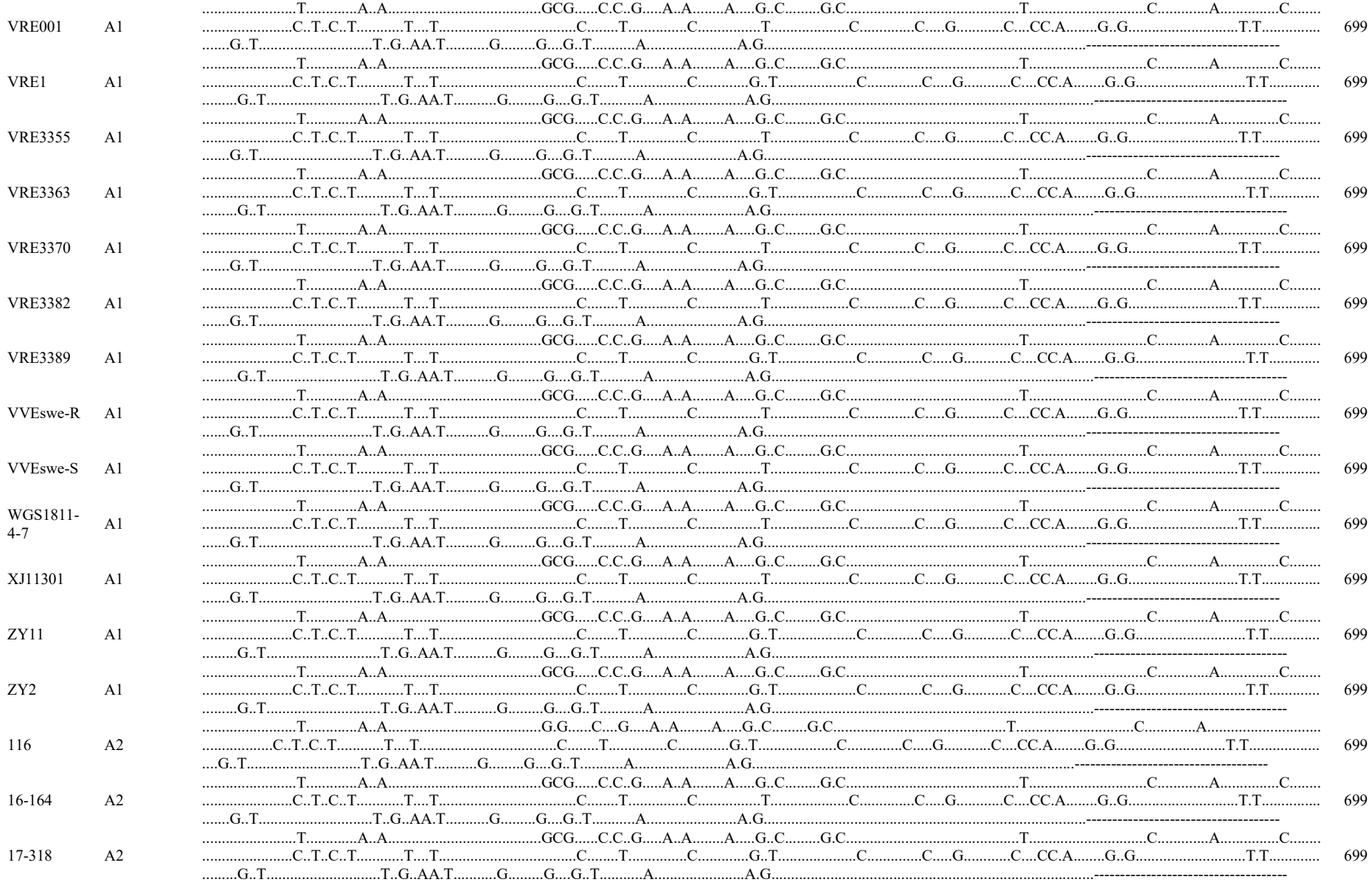

E8202 A1 699
E8284 A1 699
E8290 A1 699
E8328 A1 699
E8377 A1 699
E8414 A1 699
E8423 A1 699
E8927 A1 699
Ef_aus00233 A1 699
Ef_DMG1500501 A1 699
Efaecium_ER03933.3A A1 699
Efaecium_ER04120.3A A1 699
Efaecium_ER04462.3A A1 699
Efaecium_ER04484.3A A1 699
Efaecium_ER04526.3A A1 699
Efaecium_ER04526.5A A1 699

Efaecium_ER04619.3A	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
EFE10021	A1	.....T.....A..A.....G.G.....C....G....A..A.....A....G..C.....G.C.....T.....C.....A..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... ....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
EFE11651	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
FB-1	A1	.....T.....A..A.....G.G.....C....G....A..A.....A....G..C.....G.C.....T.....C.....A..... .....C..T..C..T.....T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... ..G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
HJP554	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
HOU503	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
ISMMS_VRE_1	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
ISMMS_VRE_11	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....-...G....G..T.....A.....A.G.....----------	633
ISMMS_VRE_12	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
ISMMS_VRE_7	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
ISMMS_VRE_9	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
K60-39	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
KUHS13	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
LAC7.2	A1	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
M20887	B	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699

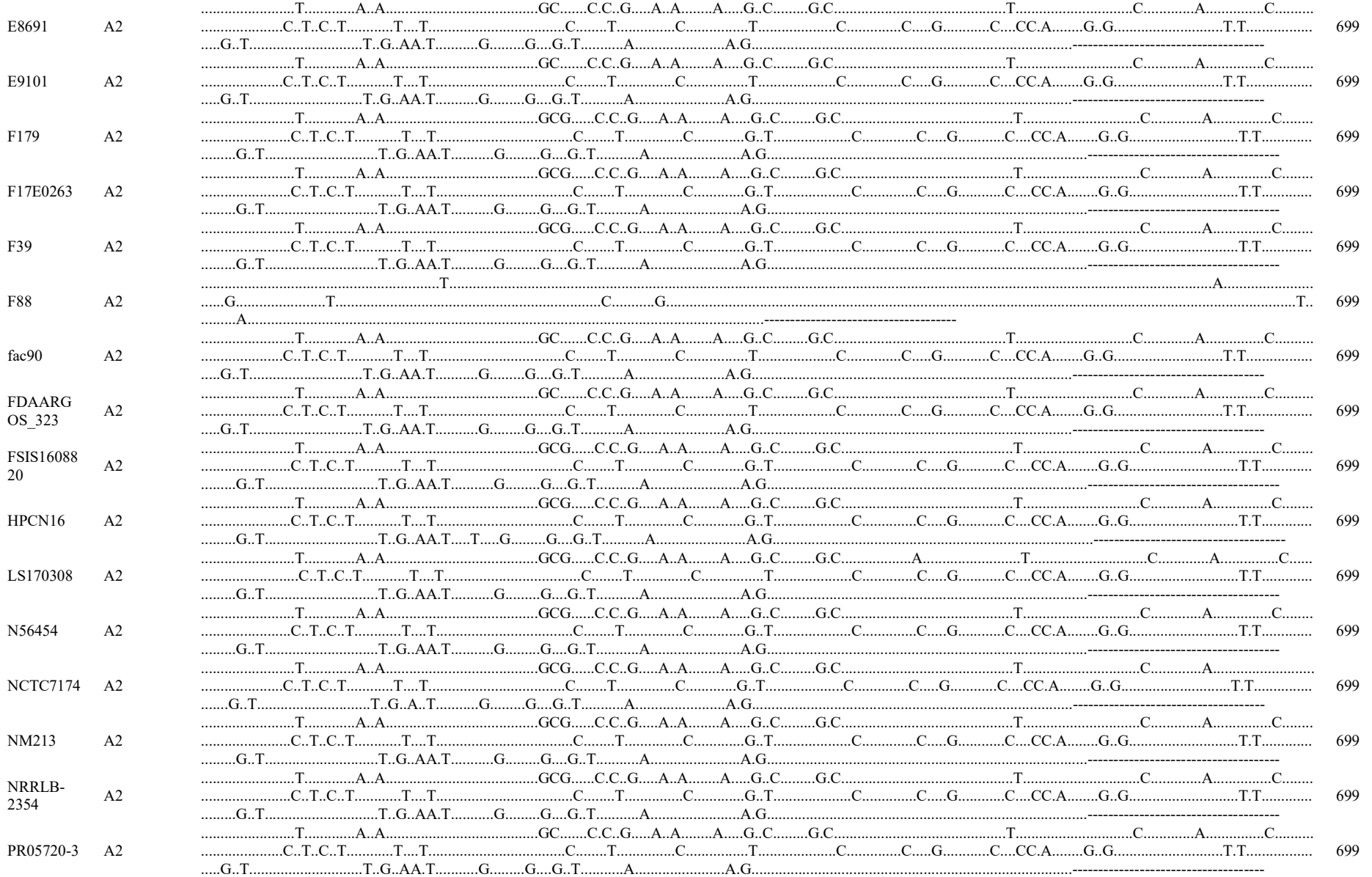

NMVRE-001 A1 699
PR01996-12 A1 699
QU50 - gluP A2 699
RBWH1 A1 699
SC1762 A1 699
SC1762-D A1 699
SC4 A1 699
SRR24 A1 699
UAMSEF_01 A1 699
UAMSEF_09 A1 699
UAMSEF_20 A1 699
USZ_VRE32_P32 A1 699
USZ_VRE5_P5 A1 699
USZ_VRE53_P46 A1 699
USZ_VRE67_P60 A1 699
UW8175 A1 699

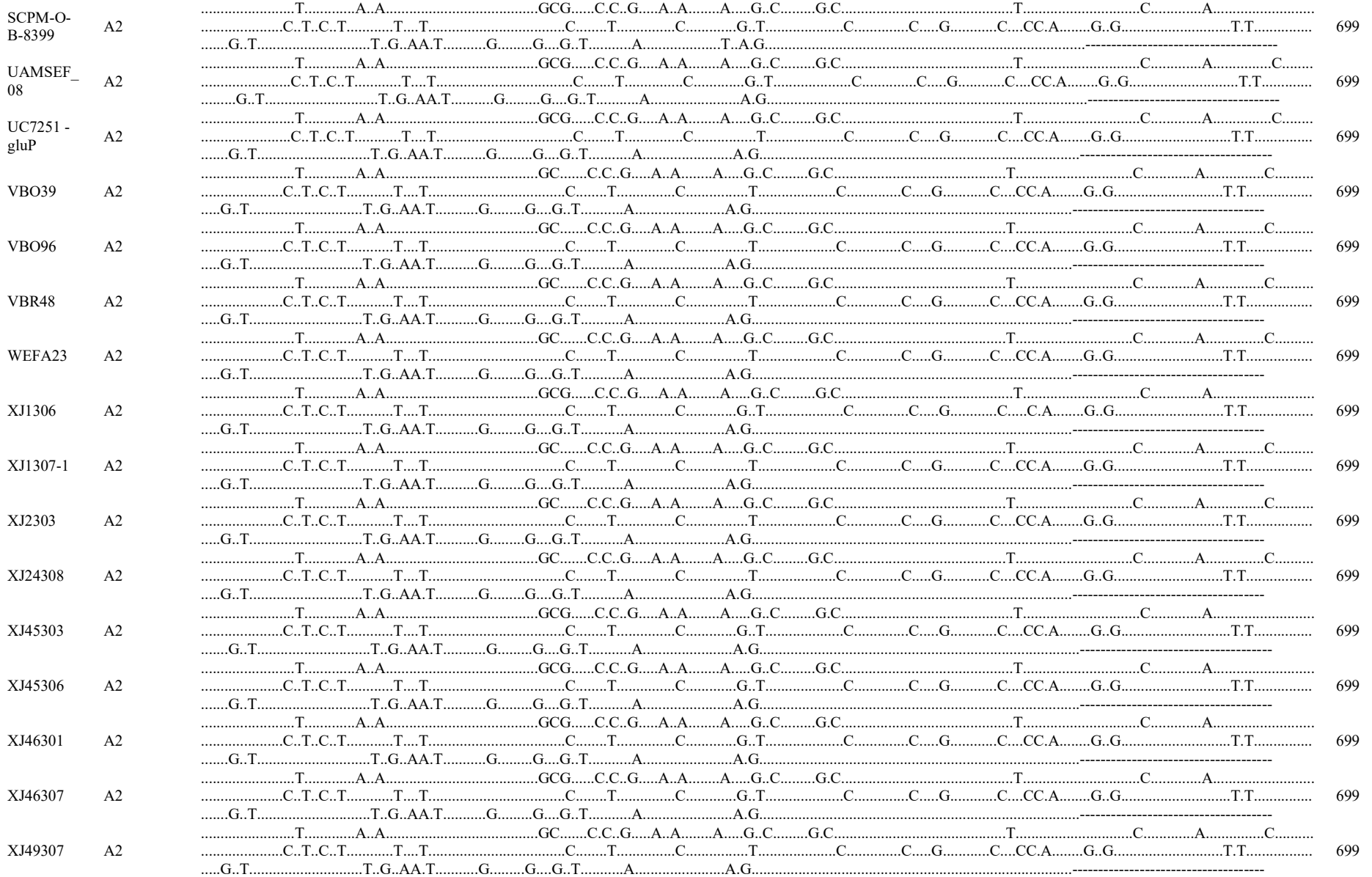


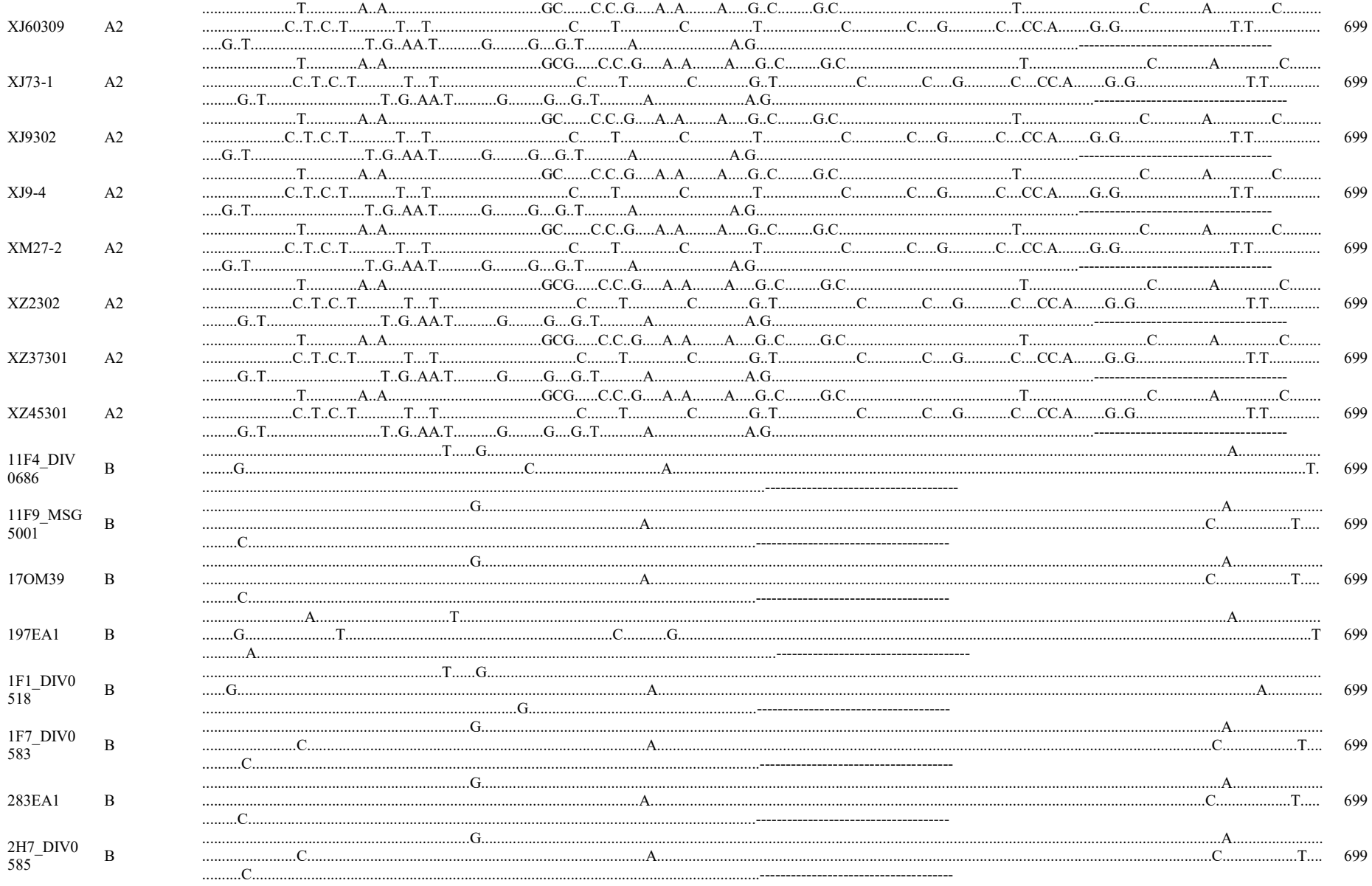
V1164	A1		699
V1225	A1		699
V13-21-E11-012-001	A1		699
V1836	A1		699
V2937	A1		699
VB12993	A1		699
VB13828	A1		699
VB3025	A1		699
VB3240	A1		699
VB3338	A1		699
VB3378	A1		699
VB3895	A1		699
VB6171	A1		699
VB6521	A1		699
VB976	A1		699
VRE	A1		699

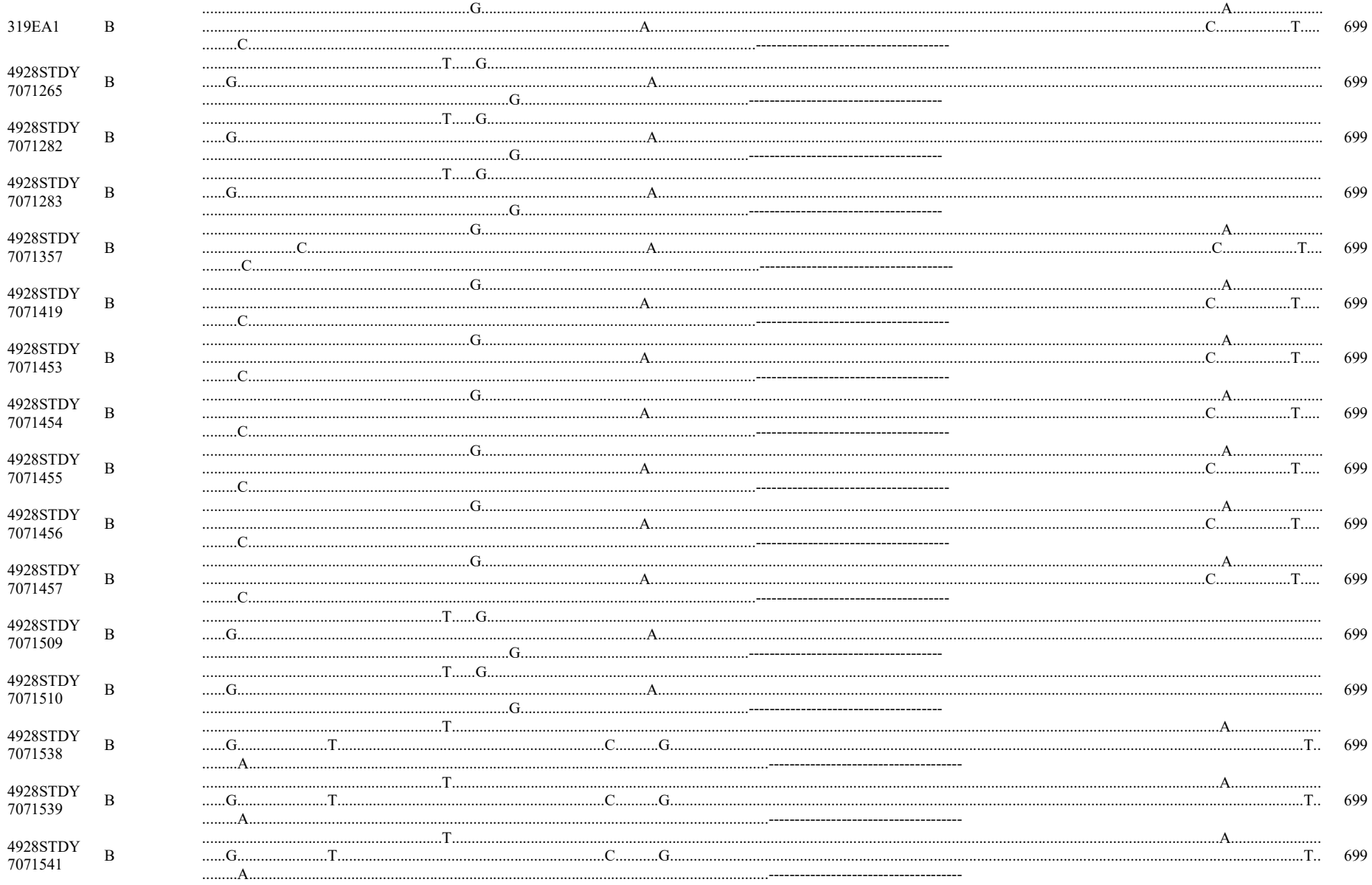

VRE001 A1 699
VRE1 A1 699
VRE3355 A1 699
VRE3363 A1 699
VRE3370 A1 699
VRE3382 A1 699
VRE3389 A1 699
VVEswe-R A1 699
VVEswe-S A1 699
WGS1811-4-7 A1 699
XJ11301 A1 699
ZY11 A1 699
ZY2 A1 699
116 A2 699
16-164 A2 699
17-318 A2 699

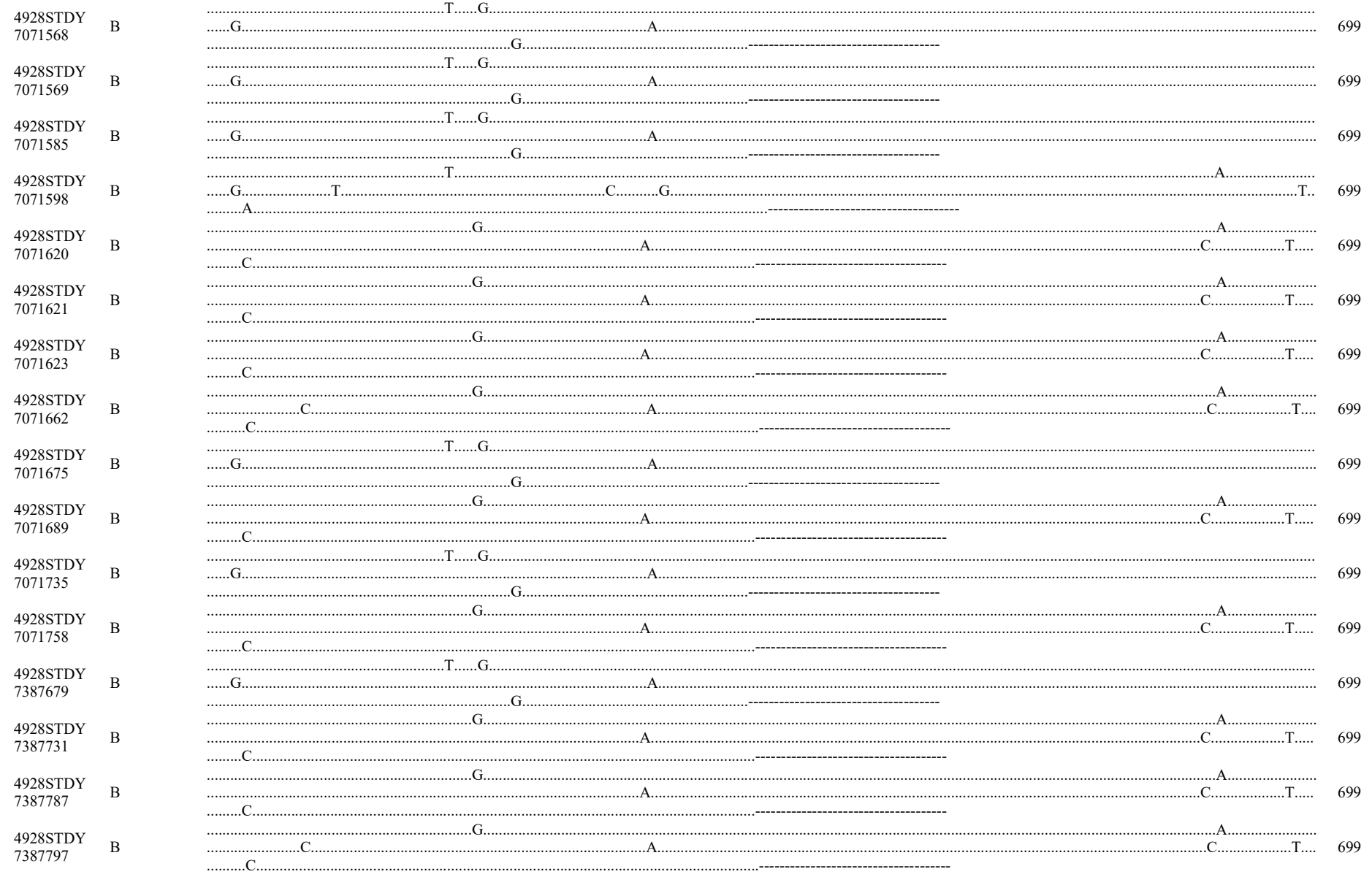
17-508	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
18-133	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....-	735
18-201	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
18-465	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
3012STDY 6244127	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
64-3	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
AVS0243	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E0139	A2	.....T.....A..A.....GC.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E0595	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E1334	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E2079	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E4227	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E4239	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....G..T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E4413	A2	.....T.....A..A.....GC.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E4438	A2	.....T.....A..A.....GC.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699
E843xGE-1-TC1	A2	.....T.....A..A.....GCG.....C.C..G....A..A.....A....G..C.....G.C.....T.....C.....A.....C..... .....C..T..C..T.....T...T.....C.....T.....C.....T.....C.....C....G.....C...CC.A.....G..G.....T.T..... .....G..T.....T..G..AA.T.....G.....G....G..T.....A.....A.G.....----------	699


E8691 A2 699
E9101 A2 699
F179 A2 699
F17E0263 A2 699
F39 A2 699
F88 A2 699
fac90 A2 699
FDAARGOS_323 A2 699
FSIS1608820 A2 699
HPCN16 A2 699
LS170308 A2 699
N56454 A2 699
NCTC7174 A2 699
NM213 A2 699
NRRLB-2354 A2 699
PR05720-3 A2 699


SCPM-O-B-8399 A2 699
UAMSEF_08 A2 699
UC7251 - gluP A2 699
VBO39 A2 699
VBO96 A2 699
VBR48 A2 699
WEFA23 A2 699
XJ1306 A2 699
XJ1307-1 A2 699
XJ2303 A2 699
XJ24308 A2 699
XJ45303 A2 699
XJ45306 A2 699
XJ46301 A2 699
XJ46307 A2 699
XJ49307 A2 699


XJ60309 A2 699
XJ73-1 A2 699
XJ9302 A2 699
XJ9-4 A2 699
XM27-2 A2 699
XZ2302 A2 699
XZ37301 A2 699
XZ45301 A2 699
11F4_DIV0686 B 699
11F9_MSG5001 B 699
17OM39 B 699
197EA1 B 699
1F1_DIV0518 B 699
1F7_DIV0583 B 699
283EA1 B 699
2H7_DIV0585 B 699


319EA1 B 699
4928STDY 7071265 B 699
4928STDY 7071282 B 699
4928STDY 7071283 B 699
4928STDY 7071357 B 699
4928STDY 7071419 B 699
4928STDY 7071453 B 699
4928STDY 7071454 B 699
4928STDY 7071455 B 699
4928STDY 7071456 B 699
4928STDY 7071457 B 699
4928STDY 7071509 B 699
4928STDY 7071510 B 699
4928STDY 7071538 B 699
4928STDY 7071539 B 699
4928STDY 7071541 B 699


4928STDY 7071568 B 699
4928STDY 7071569 B 699
4928STDY 7071585 B 699
4928STDY 7071598 B 699
4928STDY 7071620 B 699
4928STDY 7071621 B 699
4928STDY 7071623 B 699
4928STDY 7071662 B 699
4928STDY 7071675 B 699
4928STDY 7071689 B 699
4928STDY 7071735 B 699
4928STDY 7071758 B 699
4928STDY 7387679 B 699
4928STDY 7387731 B 699
4928STDY 7387787 B 699
4928STDY 7387797 B 699

4928STDY7387800	B	.....G.....A..... .....C.....A.....C.....T.... .....C.....----------	699
4928STDY7387840	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
4928STDY7387877	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
4928STDY7387886	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
4928STDY7387890	B	.....T.....G..... .....G.....A..... .....G.....----------	699
55EA1	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
6H2_DIV0141	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
7H8_DIV0219	B	.....T.....G.....A....T...A..... .....G.....A..... .....C.....----------	699
8EA1	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
8S3	B	.....T.....G..... .....G.....A..... .....G.....----------	699
ARL09-409	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
BIO4598	B	.....G.....A..... .....C.....A.....C.....T.... .....C.....----------	699
BIOML-A3	B	.....G.....A..... .....C.....A.....C.....T.... .....C.....----------	699
BIOPOP-3ALE	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
BIOPOP-3WT	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
BM4105-RF	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699

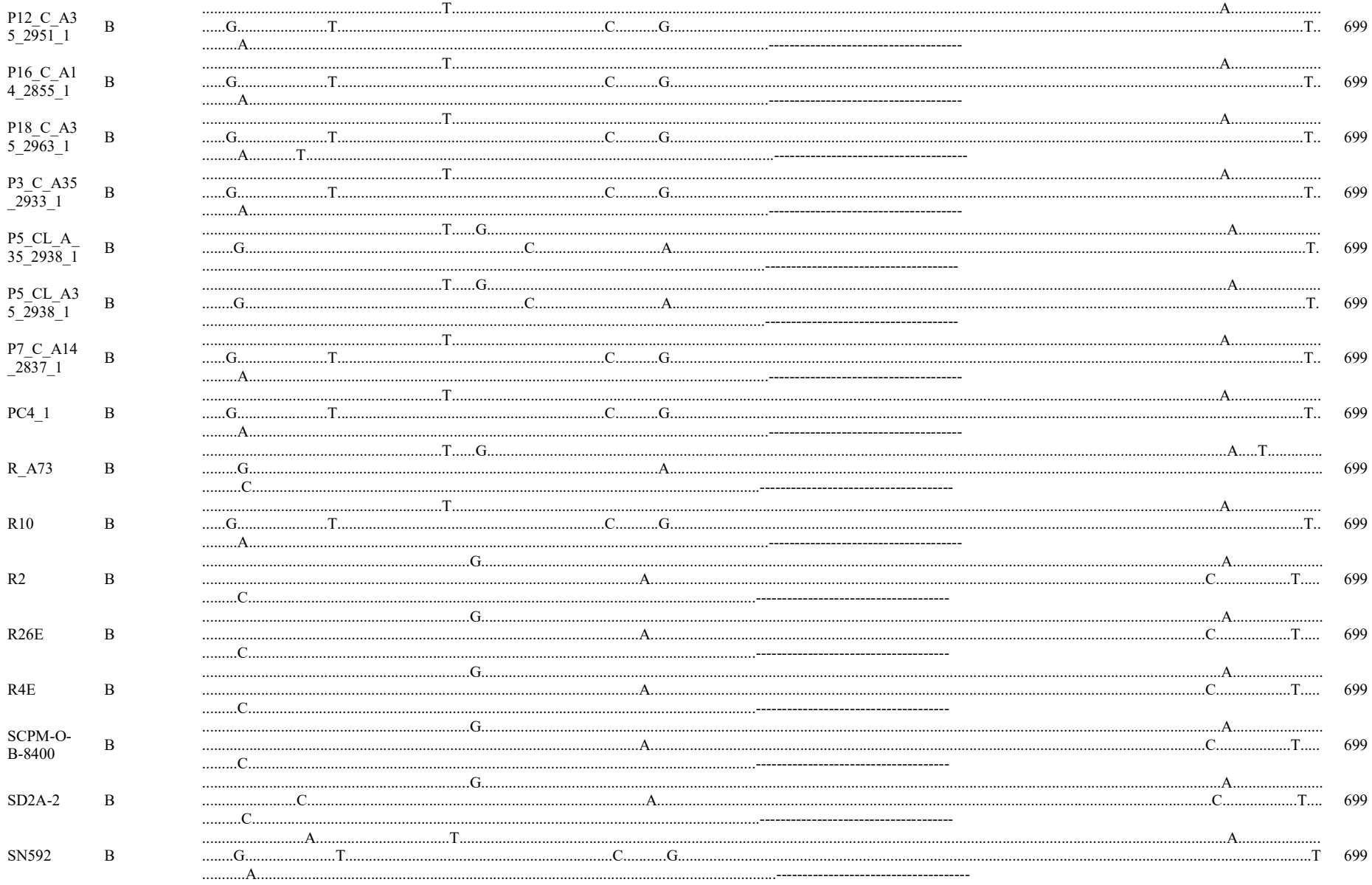
BM4107	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
BMECCcat_4	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
BMpECCcat_2	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
BSD2780061688st2_C8	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
C603	B	.....G.....A..... .....C.....A.....C.....T.... .....C.....----------	699
cau273	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
cau274	B	.....G.....A..... .....C.....A.....C.....T.... .....C.....----------	699
CBA7134	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
com12	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
Com15	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N52656	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N52732	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N52769	B	.....G.....A..... .....A.....T..... G.....C.....T.....----------	699
CVM_N54519	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N54599	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N55279	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699

CVM_N55290	B	.....G.....A..... .....A.....T..... G.....C.....T.....----------	699
CVM_N55317	B	.....G.....A..... .....A.....T..... G.....C.....T.....----------	699
CVM_N59513F	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
CVM_N59531F	B	.....G.....A..... .....A.....T..... .....C.....A.....----------	699
CVM_N59559F	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N59589F	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
CVM_N59613F	B	.....T.....G.....A..... .....G.....T.....C.....G.....T .....A.....----------	699
CVM_N59624F	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
CVM_N59653F	B	.....T.....G.....A..... .....G.....T.....C.....G.....T .....A.....----------	699
CVM_N59711F	B	.....T.....A..... .....G.....T..T.....C.....G.....T. .....A.....----------	699
CVM_N59745F	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
CVM_N59943F	B	.....G.....A..... .....A.....C.....T..... .....C.....A.....----------	699
CVM_N59947F	B	.....T.....G.....A..... .....G.....T.....C.....G.....T .....A.....----------	699
CVM_N60001F	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
CVM_N60190F	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
CVM_N60417F	B	.....G.....A..... .....C.....A.....C.....T.... .....C.....----------	699

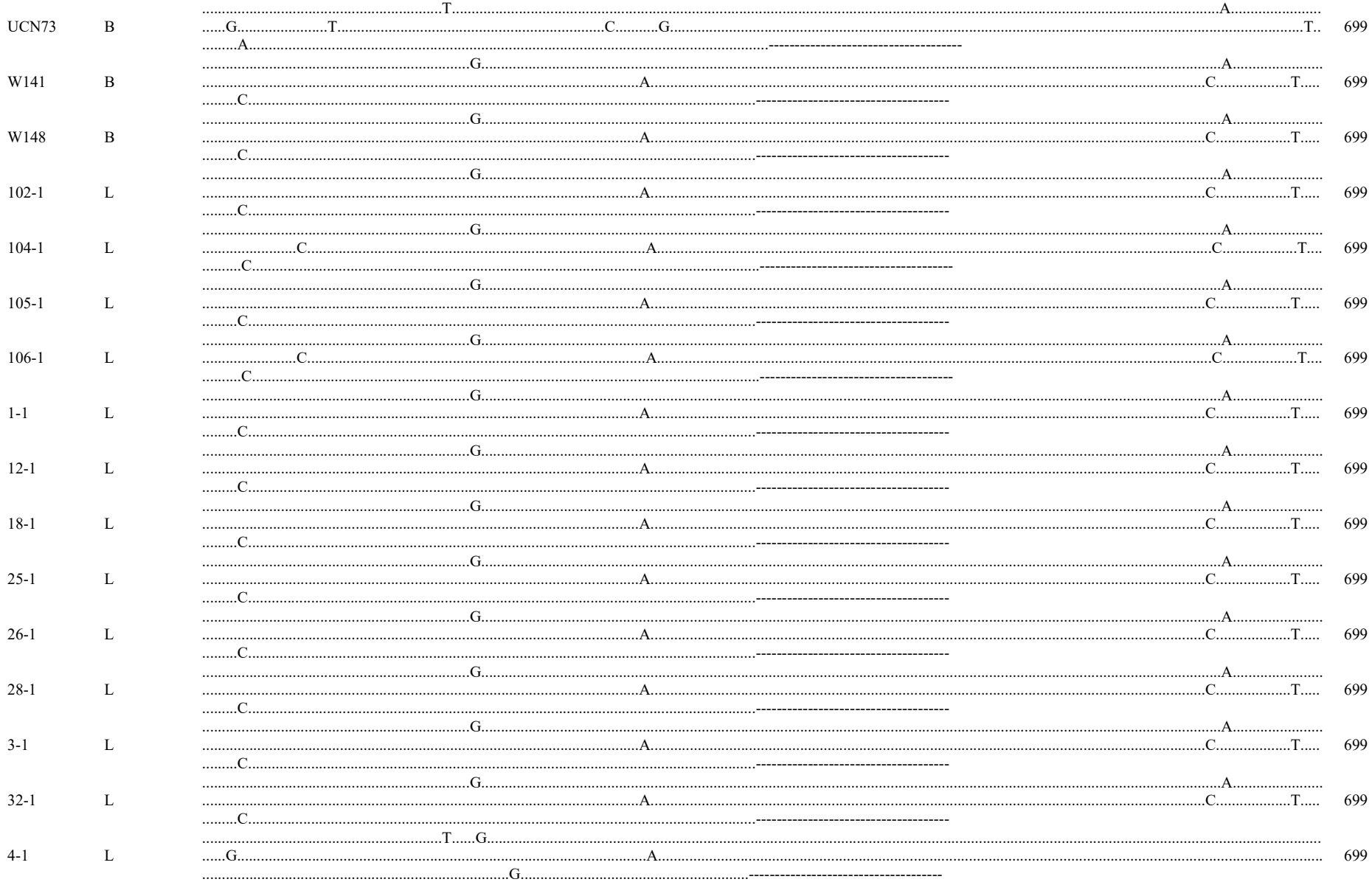
DMEA02	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
DT1-1	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
FA3	B	.....G.....A..... .....A.....A.....C.....T... .....C.....----------	699
FS86 - gluP	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
Gr17	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
HB-1	B	.....G.....A..... .....C.....A.....C.....A.....T.. .....C.....----------	699
Hp_22-12_05	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
Hp_5-10	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
Hp_5-10_05	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
Hp_7-8	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
Hp_7-8_05	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
HPCN13	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
HY07	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
IIFCSG-B5	B	.....T.....G.....A..... .....G.....T.....C.....G.....T .....A.....----------	699
JB00008	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
JE1	B	.....T.....G..... .....G.....A..... .....G.....----------	699

KACC15689	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC15700	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC15711	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC15960	B	.....T.....G.....A..... .....A.....C.....T... .....C.....----------	699
KACC15962	B	.....G.....A..... .....A.....T.....C.....T.... .....C.....----------	699
KACC16076	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC16093	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC16097	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC16100	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KACC16106	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KMB_624	B	.....T.....G..... .....G.....A..... .....G.....----------	699
LCT-EF20	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
LCT-EF258	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
LCT-EF90	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
L-X	B	.....T.....G..... .....G.....A..... .....G.....----------	699
M1126	B	.....G.....A..... .....A.....T.....C.....T.... .....C.....-.....	735

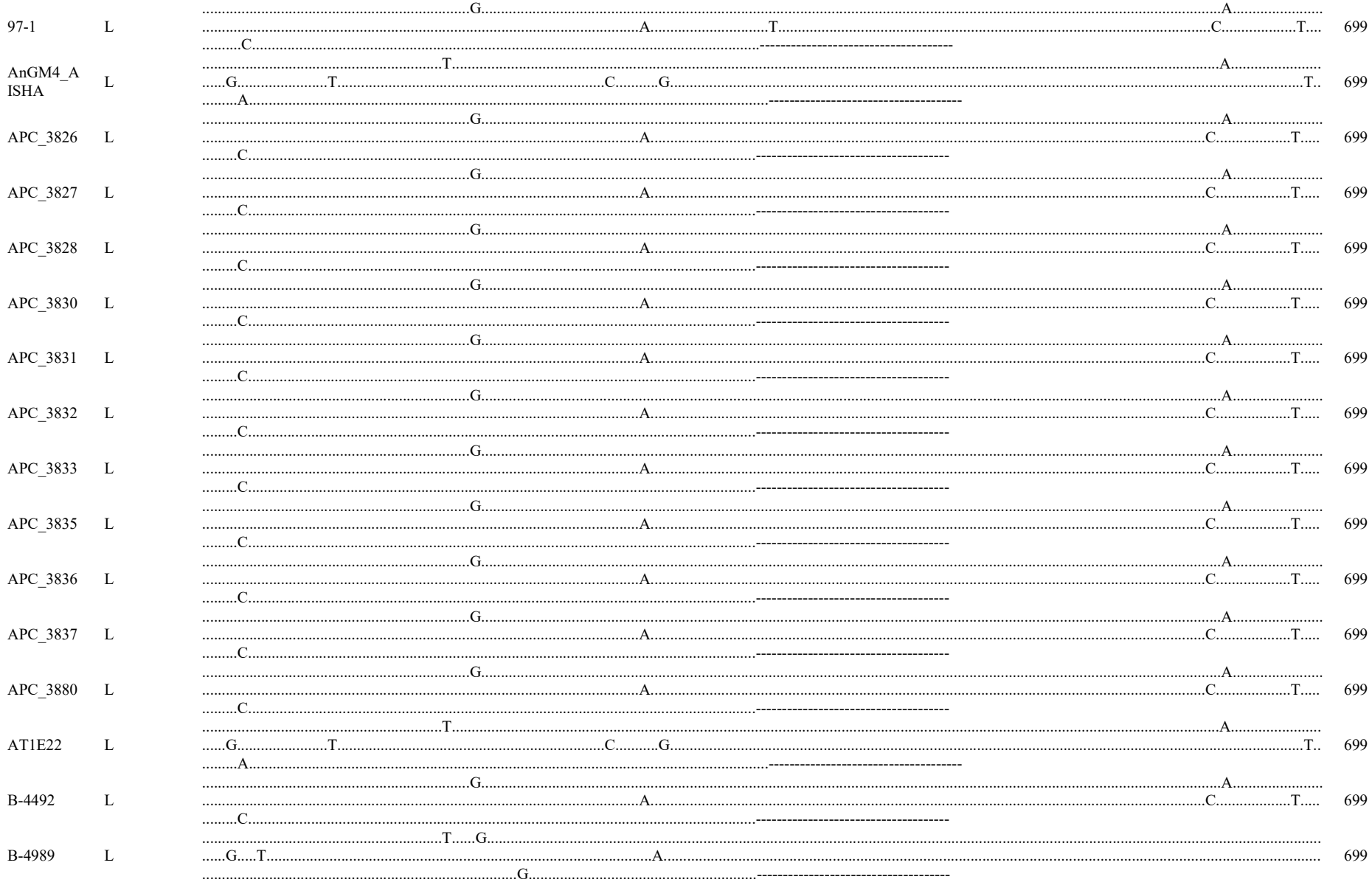
M208	B	.....G.....A..... .....A.....T.....C.....T.... .....C.....-.....	735
M210	B	.....G.....A..... .....A.....C.....T.... .....C.....-.....	735
M212	B	----------.....-G.....	684
M2124	B	.....G.....A..... .....A.....T.....C.....T.... .....C.....-.....	735
M213	B	.....G.....A..... .....A.....C.....T.... .....C.....-.....	735
M3K31	B	.....G.....A..... .....A.....C.....T.... .....C.....----------	699
M426	B	.....G.....A..... .....A.....C.....T.... .....C.....-.....	735
M5123	B	.....G.....A..... .....A.....C.....T.... .....C.....----------	699
M641	B	.....G.....A..... .....A.....T.....C.....T.... .....C.....-.....	735
M648	B	.....G.....A..... .....A.....C.....T.... .....C.....----------	699
M749	B	----------.....-G.....	684
M750	B	.....G.....A..... .....A.....C.....T.... .....C.....-.....	735
M754	B	.....G.....A..... .....A.....C.....T.... .....C.....-.....	735
M755	B	.....G.....A..... .....A.....C.....T.... .....C.....----------	699
MGYG-HGUT-02320	B	.....T.....A..... .....G.....T..T.....C.....G.....T. .....A.....----------	699
MJR8396B	B	.....G.....A..... .....A.....C.....T.... .....C.....----------	699
P12_C_A28_2916_1	B	.....T.....A..... .....G.....T.....C.....G.....A.....T .....A.....----------	699

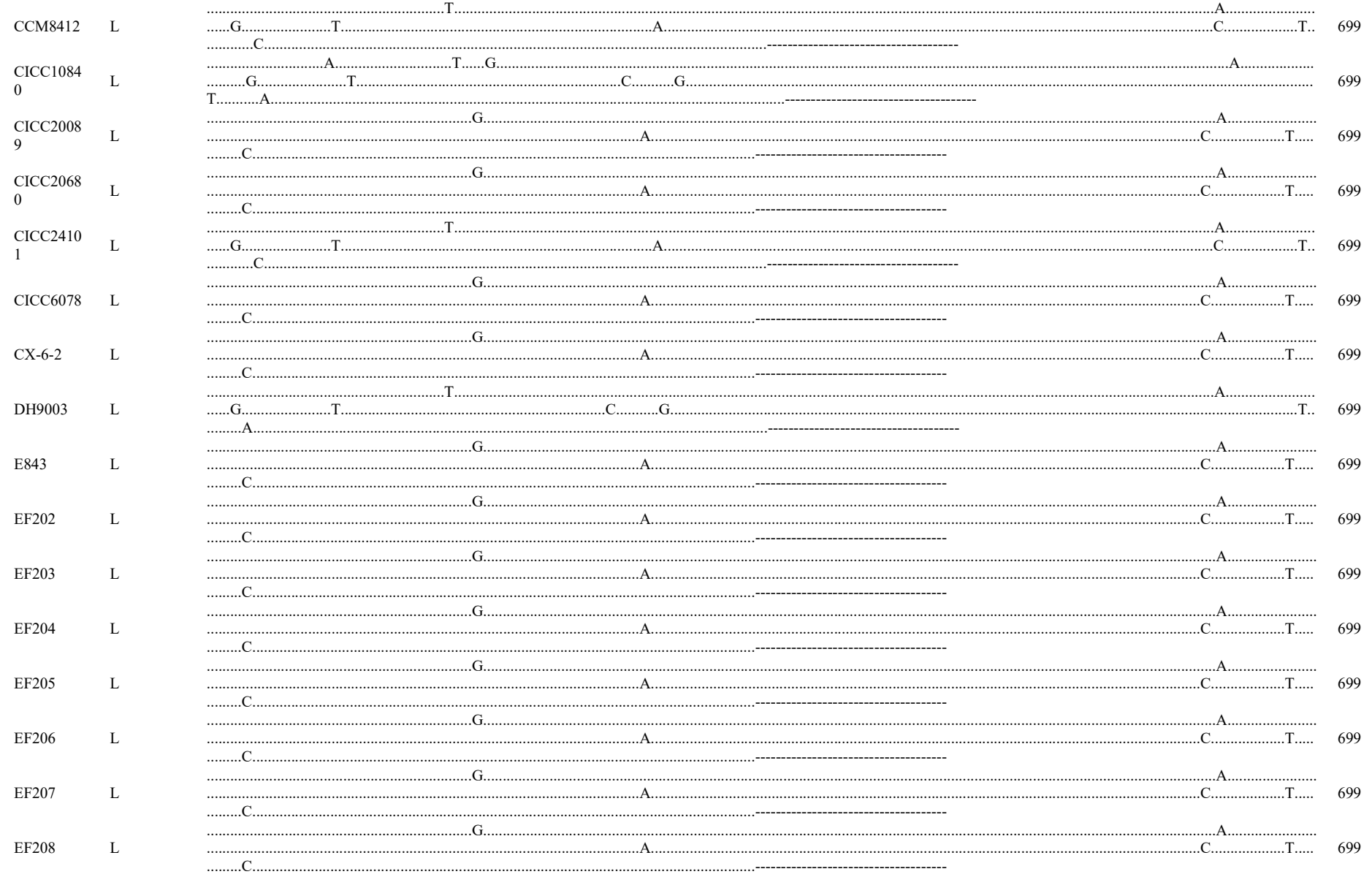

P12_C_A35_2951_1 B 699
P16_C_A14_2855_1 B 699
P18_C_A35_2963_1 B 699
P3_C_A35_2933_1 B 699
P5_CL_A_35_2938_1 B 699
P5_CL_A35_2938_1 B 699
P7_C_A14_2837_1 B 699
PC4_1 B 699
R_A73 B 699
R10 B 699
R2 B 699
R26E B 699
R4E B 699
SCPM-O-B-8400 B 699
SD2A-2 B 699
SN592 B 699

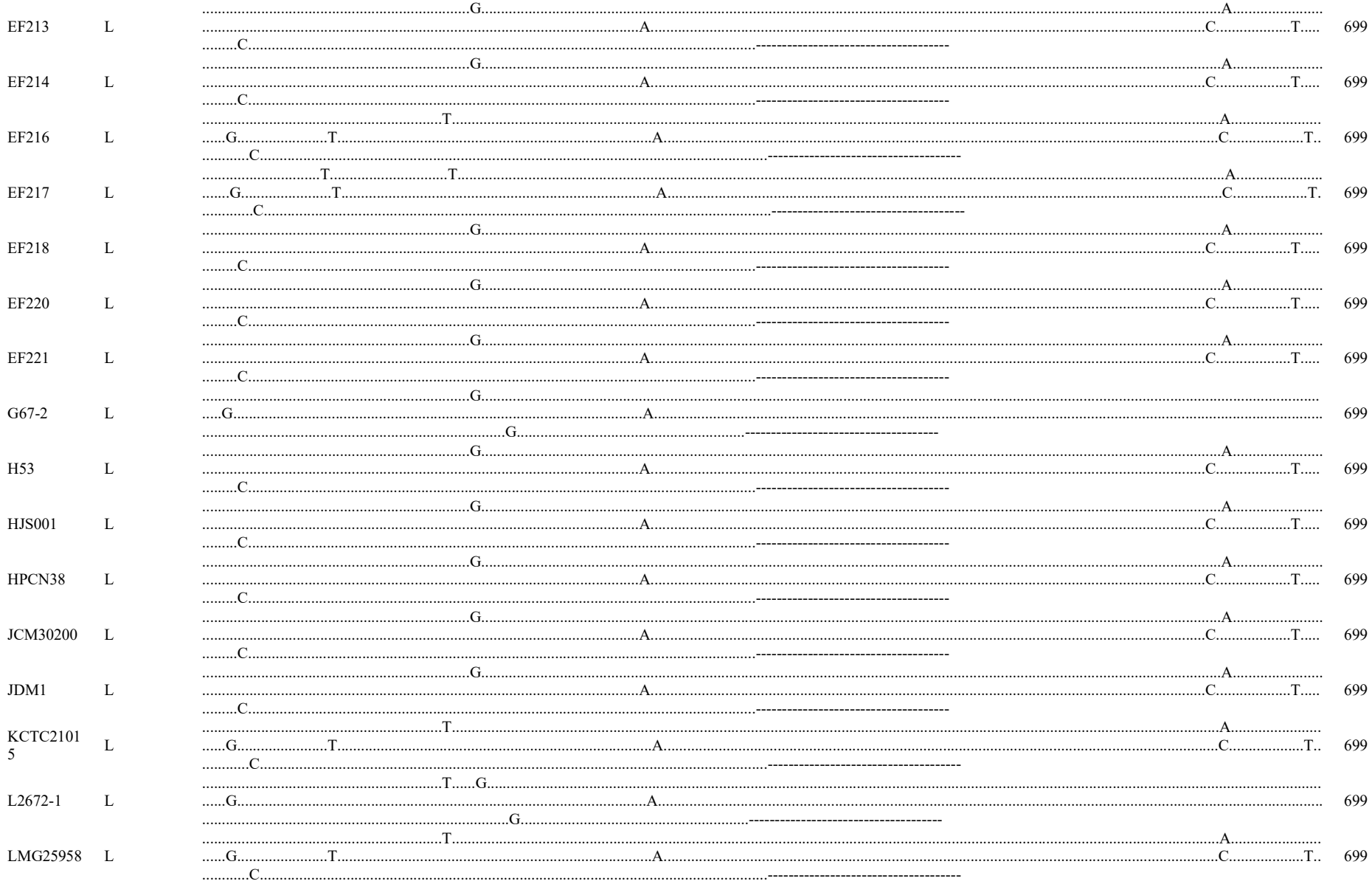
SP15	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SRCM103341	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SRCM103470	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SWEnt-1198	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
T110	B	.....T.....G..... .....G.....A..... .....G.....----------	699
TK-P5D	B	.....T.....G..... .....G.....A..... .....G.....----------	699
TR10-27	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
TX1330	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
TX1337RF	B	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
UAA1280	B	.....T.....G.....A..... .....G.....A..... .....----------	699
UAMSEF_24	B	.....T.....A..... .....G.....T.....A.....C.....T.. .....C.....T.....----------	699
UBA4566	B	.....T.....G.....A..... .....G.....T.....C.....G.....T .....A.....----------	699
UBEF-41	B	.....T.....G..... .....G.....A..... .....G.....----------	699
UC7256	B	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
UC8668	B	.....G.....A..... .....A.....T.....T..... .....C.....----------	699
UCN72	B	.....T.....A..... .....G.....T.....C.....G.....C.....T .....A.....----------	699


UCN73 B 699
W141 B 699
W148 B 699
102-1 L 699
104-1 L 699
105-1 L 699
106-1 L 699
1-1 L 699
12-1 L 699
18-1 L 699
25-1 L 699
26-1 L 699
28-1 L 699
3-1 L 699
32-1 L 699
4-1 L 699

		.....T.....G.....	
44-1	L	.....G.....A.....	699
		.....G.....----------	
		.....T.....A.....	
45-1	L	.....G.....T.....C.....G.....T..	699
		.....A.....----------	
		.....T.....A.....	
46-1	L	.....G.....T.....C.....G.....T..	699
		.....A.....----------	
		.....T.....G.....	
5-1	L	.....G.....A.....	699
		.....G.....----------	
		.....G.....A.....	
54-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....G.....A.....	
55-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....G.....A.....	
56-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....T.....G.....	
6-1	L	.....G.....A.....	699
		.....G.....----------	
		.....G.....A.....	
65-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....T.....G.....	
7-1	L	.....G.....A.....	699
		.....G.....----------	
		.....G.....A.....	
73-1	L	.....C.....A.....C.....T....	699
		.....C.....----------	
		.....G.....A.....	
76-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....G.....A.....	
77-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....G.....A.....	
86-1	L	.....A.....C.....T.....	699
		.....C.....----------	
		.....G.....A.....	
88-1	L	.....C.....A.....C.....T....	699
		.....C.....----------	
		.....G.....A.....	
88-2	L	.....A.....C.....T.....	699
		.....C.....----------	


97-1 L 699
AnGM4_AISHA L 699
APC_3826 L 699
APC_3827 L 699
APC_3828 L 699
APC_3830 L 699
APC_3831 L 699
APC_3832 L 699
APC_3833 L 699
APC_3835 L 699
APC_3836 L 699
APC_3837 L 699
APC_3880 L 699
AT1E22 L 699
B-4492 L 699
B-4989 L 699


CCM8412 L 699
CICC10840 L 699
CICC20089 L 699
CICC20680 L 699
CICC24101 L 699
CICC6078 L 699
CX-6-2 L 699
DH9003 L 699
E843 L 699
EF202 L 699
EF203 L 699
EF204 L 699
EF205 L 699
EF206 L 699
EF207 L 699
EF208 L 699



EF213	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
EF214	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
EF216	L	.....T.....A..... .....G.....T.....A.....C.....T.. .....C.....----------	699
EF217	L	.....T.....T.....A..... .....G.....T.....A.....C.....T. .....C.....----------	699
EF218	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
EF220	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
EF221	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
G67-2	L	.....G..... .....G.....A..... .....G.....----------	699
H53	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
HJS001	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
HPCN38	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
JCM30200	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
JDM1	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
KCTC21015	L	.....T.....A..... .....G.....T.....A.....C.....T.. .....C.....----------	699
L2672-1	L	.....T.....G..... .....G.....A..... .....G.....----------	699
LMG25958	L	.....T.....A..... .....G.....T.....A.....C.....T.. .....C.....----------	699

NM29-3	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
NM30-4	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
NM31-5	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
S10-4	L	.....G..... .....G.....A..... .....G.....----------	699
s-7	L	.....T.....A..... .....G.....T.....A.....C.....T.. .....C.....----------	699
SCPM-O-B-8929	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8930	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8931	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8932	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8933	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8939	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8943	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8947	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
SCPM-O-B-8948	L	.....T.....A..... .....G.....T.....C.....G.....T.. .....A.....----------	699
SCPM-O-B-8953	L	.....G.....A..... .....A.....C.....T..... .....C.....----------	699
Tb32-6	L	.....G.....A..... .....A.....T....T..... .....C.....----------	699

XJ28301	L	.....T.....A..... .....G.....T.....A.....C.T.....T. .....C.....----------	699
XJ28304	L	.....T.....A..... .....G.....T.....A.....C.T.....T. .....C.....----------	699
XZ35303	L	.....G.....A..... .....A.....T....T..... .....C.....----------	699
XZ37302	L	.....G.....A..... .....A.....T....T..... .....C.....----------	699

Table S3. List of well-characterized enterococci (n=137) used for PCR validation of primers.

Strain	Country	Origin	Year	MLST	ABR profile	Clade A	Clade B
CCP102	Portugal	Hospital	1996	ST18	VAN, TEC, AMP, CIP, ERY, TET, GEN, STR	+	-
CCP103	Portugal	Hospital	1997	ST670	VAN, TEC, AMP, ERY, TET, GEN, STR	+	-
CCP104	Portugal	Hospital	1997	ST8	VAN, TEC, AMP, CIP, ERY	+	-
CCP105	Portugal	Hospital	1998	ST190	VAN, TEC, AMP, CIP, ERY, TET, GEN, STR	+	-
CCP106	Portugal	Hospital	1999	ST132	VAN, TEC, AMP, CIP, ERY, TET, GEN, STR	+	-
CCP107	Portugal	Hospital	1999	ST366	VAN, TEC, AMP, CIP, ERY, GEN, STR	+	-
CCP108	Portugal	Hospital	2000	ST16	VAN, TEC, AMP, ERY, TET, STR	+	-
CCP109	Portugal	Hospital	2000	ST18	VAN, TEC, AMP, CIP, ERY, TET, GEN	+	-

CCP110	Portugal	Hospital	2000	ST132	VAN, TEC, AMP, CIP, ERY, TET, GEN	+	-
CCP111	Portugal	Hospital	2000	ST367	VAN, TEC, AMP, CIP, ERY, GEN, STR	+	-
CCP112	Portugal	Hospital	2000	ST18	VAN, TEC, AMP, CIP, ERY, STR	+	-
CCP113	Portugal	Hospital	2001	ST1054	CIP	+	-
CCP114	Portugal	Hospital	2001	ST719	VAN, TEC, AMP, ERY	+	-
CCP115	Portugal	Hospital	2001	ST8	VAN, TEC, AMP, CIP, ERY, TET, GEN, STR	+	-
CCP116	Portugal	Hospital	2001	ST18	VAN, TEC, AMP, CIP, ERY, TET	+	-
CCP117	Portugal	Hospital	2001	ST132	VAN, TEC, AMP, CIP, ERY, GEN	+	-
CCP118	Portugal	Hospital	2001	ST132	VAN, TEC, AMP, CIP, ERY, GEN, STR	+	-
CCP119	Portugal	Hospital	2001	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP120	Portugal	Hospital	2002	ST18	VAN, TEC, AMP, CIP, ERY, TET	+	-
CCP121	Portugal	Hospital	2002	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP122	Portugal	Hospital	2002	ST390	AMP, CIP, ERY, TET, STR	+	-
CCP123	Portugal	Hospital	2002	ST280	AMP, ERY, TET	+	-
CCP124	Portugal	Hospital	2002	ST5	VAN, TEC, CIP, ERY	+	-
CCP125	Portugal	Hospital	2002	ST132	VAN, AMP, CIP, ERY, GEN	+	-
CCP126	Portugal	Hospital	2002	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP127	Portugal	Hospital	2002	ST18	VAN, TEC, AMP, CIP, ERY	+	-

CCP128	Portugal	Hospital	2002	ST132	VAN, TEC, AMP, CIP, ERY	+	-
CCP129	Portugal	Hospital	2002	ST17	VAN, TEC, AMP, CIP, ERY	+	-
CCP130	Portugal	Hospital	2003	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP131	Portugal	Hospital	2003	ST280	VAN, TEC, AMP, CIP, ERY	+	-
CCP132	Portugal	Hospital	2003	ST280	VAN, TEC, AMP, CIP, ERY, GEN	+	-
CCP133	Portugal	Hospital	2003	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP134	Portugal	Hospital	2003	ST18	VAN, TEC, AMP, CIP, ERY, STR	+	-
CCP135	Portugal	Hospital	2003	ST132	VAN, TEC, AMP, CIP, ERY, GEN	+	-
CCP136	Tunisia	Hospital	2003	ST17	VAN, TEC, AMP, CIP, ERY, STR	+	-
CCP137	Portugal	Hospital	2006	ST280	VAN, TEC, AMP, CIP, ERY	+	-
CCP138	Portugal	Hospital	2007	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP139	Portugal	Hospital	2007	ST515	VAN, TEC, AMP, CIP, ERY	+	-
CCP140	Portugal	Hospital	2007	ST78	AMP, CIP, ERY, GEN	+	-
CCP141	Portugal	Hospital	2007	ST391	VAN, TEC, AMP, CIP, ERY	+	-
CCP142	Portugal	Hospital	2007	ST18	VAN, TEC, AMP, CIP, ERY	+	-
CCP143	Portugal	Hospital	2008	ST656	VAN, TEC, AMP, CIP, ERY, TET, STR	+	-
CCP144	Portugal	Hospital	2011	ST412	VAN, TEC, AMP, CIP, ERY, TET	+	-
CCP145	Portugal	Hospital	2011	ST412	VAN, TEC, AMP, CIP, ERY, TET	+	-

CCP146	Portugal	Hospital	2012	ST412	VAN, AMP, CIP, ERY, TET	+	-
CCP147	Portugal	Hospital	2012	ST117	VAN, AMP, CIP, ERY, TET, GEN	+	-
CCP148	Portugal	Hospital	2013	ST78	VAN, AMP, CIP, ERY, TET	+	-
CCP149	Portugal	Hospital	2013	ST117	VAN, AMP, CIP, ERY, TET, GEN	+	-
CCP150	Portugal	Hospital	2014	ST117	AMP, CIP, ERY, LZD	+	-
CCP151	Portugal	Hospital	2014	ST117	VAN, TEC, AMP, CIP, ERY, TET, GEN	+	-
CCP152	Portugal	Hospital	2014	ST117	AMP, CIP, ERY, LZD	+	-
CCP153	Portugal	Hospital	2014	ST117	VAN, TEC, AMP, CIP	+	-
CCP154	Portugal	Hospital	2019	ST117	VAN, AMP, CIP, ERY, STR	+	-
CCP155	Portugal	Hospital	2019	ST80	VAN, TEC, AMP, CIP, ERY	+	-
CCP156	Portugal	Hospital	NK	ST18	VAN, TEC, AMP, CIP, ERY, TET, GEN	+	-
CCP157	Portugal	Hospital	NK	ST18	VAN, TEC, AMP, CIP, ERY, TET	+	-
CCP158	Portugal	Healthy volunteers	2001	ST845	ERY	+	-
CCP159	Portugal	Healthy volunteers	2001	ST32	ERY, TET, STR	+	-
CCP160	Portugal	Healthy volunteers	2001	ST845	CIP, ERY, TET, STR	+	-
CCP161	Portugal	Healthy volunteers	2001	ST18	VAN, TEC, AMP, CIP, ERY, TET, GEN, STR	+	-
CCP162	Portugal	Healthy volunteers	2001	ST89	VAN, TEC, ERY, TET, STR	+	-

CCP163	Portugal	Healthy volunteers	2001	ST846	TET, ERY, CIP, STR	+	-
CCP164	Tunisia	Cow milk	2003	ST1058	CIP, ERY, TET, LZD	+	-
CCP165	Tunisia	Cow milk	2003	ST1058	CIP, ERY, TET, LZD	+	-
CCP166	Tunisia	Cow milk	2003	ST1058	CIP, ERY, TET, LZD	+	-
CCP167	Portugal	Piggery	2006	ST428	ERY, TET, GEN	+	-
CCP168	Portugal	Piggery	2006	ST393	AMP, ERY, TET, GEN, STR	+	-
CCP169	Portugal	Piggery	2006	ST848	ERY, TET, STR	+	-
CCP170	Portugal	Piggery	2006	ST393	AMP, CIP, ERY, TET, GEN, STR	+	-
CCP171	Portugal	Piggery	2006	ST5	ERY, TET	+	-
CCP172	Portugal	Piggery	2006	ST264	AMP, TET, STR	+	-
CCP173	Portugal	Piggery	2006	ST264	AMP, CIP, ERY, TET, STR	+	-
CCP174	Portugal	Piggery	2006	ST32	TET	+	-
CCP175	Portugal	Piggery	2006	ST430	AMP, ERY, TET, STR	+	-
CCP176	Portugal	Piggery	2006	ST132	VAN, TEC, AMP, ERY, GEN	+	-
CCP177	Portugal	Piggery	2006	ST859	AMP, CIP, ERY, TET, GEN, STR	+	-
CCP178	Portugal	Piggery	2006	ST430	AMP, CIP, ERY, TET, STR	+	-
CCP179	Portugal	Piggery	2006	ST431	AMP, CIP, ERY, TET, STR	+	-
CCP180	Portugal	Piggery	2007	ST432	ERY, TET	+	-
CCP181	Portugal	Piggery	2007	ST185	VAN, TEC, TET	+	-
CCP182	Portugal	Piggery	2007	ST185	VAN, TEC, CIP, TET	+	-
CCP183	Portugal	Piggery	2007	ST133	ERY, TET, STR	+	-

CCP184	Angola	Poultry environment	2013	ST29	-	+	-
CCP185	Portugal	Poultry carcass	1999	ST196	VAN, TEC,TET	+	-
CCP186	Portugal	Poultry carcass	1999	ST421	ERY, TET	+	-
CCP187	Portugal	Poultry carcass	1999	ST157	VAN, TEC, CIP, ERY, TET	+	-
CCP188	Portugal	Poultry carcass	1999	ST405	VAN, TEC, CIP, ERY, STR	+	-
CCP189	Portugal	Poultry carcass	1999	ST236	VAN, TEC, CIP, ERY, TET, STR	+	-
CCP190	Portugal	Poultry carcass	2001	ST27	TET, STR	+	-
CCP191	Portugal	Poultry carcass	2001	ST148	AMP, CIP, ERY, TET, GEN, STR	+	-
CCP192	Portugal	Poultry carcass	2001	ST450	VAN, TEC, CIP, ERY, TET, STR	+	-
CCP193	Portugal	Poultry carcass	2001	ST406	VAN, TEC, CIP, ERY, TET, GEN, STR	+	-
CCP194	Portugal	Poultry carcass	2001	ST9	VAN, TEC, ERY, TET, STR	+	-
CCP195	Portugal	Swine faeces	2007	ST150	AMP, CIP, ERY, TET, STR	+	-
CCP196	Angola	Swine faeces	2013	ST971	AMP, CIP, TET, STR	+	-
CCP197	Portugal	Trout	2012	ST32	ERY, TET	+	-
CCP198	Portugal	Trout	2012	ST32	CIP, ERY, TET	+	-
CCP199	Portugal	Trout	2012	ST683	TET, STR	+	-
CCP200	Portugal	Trout	2012	ST1059	-	+	-
CCP201	Portugal	Aquaculture	2010	ST101	TET	+	-
CCP202	Portugal	Aquaculture	2011	ST30	TET, STR	+	-
CCP203	Portugal	Ready-to-eat salad	2010	ST640	ERY, TET	+	-
CCP204	Portugal	Ready-to-eat salad	2010	ST640	-	+	-
CCP205	Portugal	River	2003	ST18	VAN, TEC, AMP, CIP, ERY, STR	+	-

CCP206	Portugal	River	2003	ST369	VAN, TEC, AMP, CIP, ERY, GEN, STR	+	-
CCP207	Portugal	River	2003	ST369	VAN, TEC, AMP, CIP, ERY, STR	+	-
CCP208	Portugal	Sewage	2001	ST132	AMP, CIP, GEN	+	-
CCP209	Portugal	Sewage	2001	ST368	VAN, AMP, CIP, ERY, GEN, STR	+	-
CCP210	Portugal	Sewage	2001	ST132	VAN, TEC, AMP, CIP, ERY, STR	+	-
CCP211	Portugal	Sewage	2001	ST17	AMP, CIP, ERY, TET, GEN, STR	+	-
CCP212	Portugal	Hospital	2019	ST2215	-	-	+
CCP213	Portugal	Hospital	2019	ST329	ERY	-	+
CCP214	Portugal	Hospital	2020	ST361	ERY	-	+
CCP215	Portugal	Hospital	2022	ST994	ERY	-	+
CCP216	Portugal	Hospital	2022	ST118	-	-	+
CCP217	Portugal	Healthy volunteers	2001	ST798	CIP, ERY, STR	-	+
CCP218	Portugal	Healthy volunteers	2001	ST361	ERY	-	+
CCP219	Portugal	Healthy volunteers	2022	ST75	ERY, TET	-	+
CCP220	Portugal	Healthy volunteers	2022	ST717	ERY	-	+
CCP221	Portugal	Healthy volunteers	2022	ST361	ERY, TET	-	+
CCP222	Portugal	Healthy volunteers	2022	ST2097	ERY	-	+
CCP223	Portugal	Healthy volunteers	2022	ST2223	ERY	-	+
CCP224	Portugal	Healthy volunteers	2022	ST2097	ERY	-	+
CCP225	Portugal	Healthy volunteers	2022	ST800	ERY	-	+

CCP226	Portugal	Healthy volunteers	2022	ST2224	ERY	-	+
CCP227	Portugal	Healthy volunteers	2022	ST2021	ERY	-	+
CCP228	Portugal	Healthy volunteers	2022	ST640	ERY, TET, STR	-	+
CCP229	Portugal	Healthy volunteers	2022	ST233	ERY	-	+
CCP230	Portugal	Healthy volunteers	2022	ST107	-	-	+
CCP231	Portugal	Healthy volunteers	2022	ST94	ERY	-	+
CCP232	Portugal	Piggery	2006	ST695	ERY, TET	-	+
CCP233	Portugal	Piggery	2006	ST108	ERY, TET	-	+
CCP234	Portugal	Trout	2012	ST1060	CIP	-	+
CCP235	Portugal	LTCF	2016	ST1284	ERY, STR	-	+
CCP236	Portugal	LTCF	2016	ST74	ERY	-	+
CCP237	Portugal	Sewage	2002	ST123	AMP, CIP, ERY, TET, STR	-	+
CCP238	Portugal	River	2003	ST1055	-	-	+

Abbreviations: AMP, ampicillin; CCP, Culture Collection of Porto (https://ccp.ff.up.pt/); CIP, ciprofloxacin; ERY, erythromycin; GEN, gentamicin; LIN, linezolid; LTCF, long-term care facility; STR, streptomycin; ST, sequence type; TEC, teicoplanin; TET, tetracycline.VAN, vancomycin.

Isolates CCP212, CCP213, CCP214, CCP215 and CCP216 correspond to HPH55b, HPH67, HPH133, HPH282 and HPH288 respectively.

Table S4. Virulence gene content of *E. lactis* clinical isolates.

Isolate	Surface-Exposed Cell-Wall Anchored Proteins and Miscellaneous					Carbohydrate Metabolism, Regulation, Transport	PGC 1 *(fms21-20)*			PGC 2 *(fms14-fms17-fms13)*		PGC 3 (EmpABC)			PGC 4 *(fms11-fms19-fms16)*			General stress proteins		
	acm	*scm*	*sgrA*	*fibronectin*	*sagA*	*ccpA*	*bepA*	*fms21*	*fms20*	*fms14*	*fms17*	*fms13*	*ebpA* Efm	*ebpB* Efm	*ebpC* Efm	*fms16*	*fms19*	*fms11*	*gls20*	*gls33*	*glsB*
HPH 55b	*Δacm*			*Δfibronectin*		*ccpA*	*bepA*	*fms21*	*fms20*	*fms14*	*Δfms17*	*fms13*	*ΔebpAEfm*	*ΔebpBEfm*	*ebpCEfm*				*gls20*	*gls33*	*glsB*
HPH 67	*acm*	*scm*		*Δfibronectin*	*ΔsagA*	*ccpA*	*bepA*	*fms21*	*fms20*	*fms14*	*fms17*	*fms13*	*ΔebpAEfm*	*ΔebpBEfm*	*ebpCEfm*				*gls20*	*Δgls33*	*glsB*
HPH 133	*acm*		*ΔsgrA*	*Δfibronectin*	*ΔsagA*	*ccpA*	*bepA*			*fms14*	*fms17*	*fms13*	*ΔebpAEfm*	*ebpBEfm*	*ebpCEfm*	*fms16*	*fms19*	*Δfms11*	*gls20*	*Δgls33*	*glsB*
HPH 282	*acm*		*sgrA*	*Δfibronectin*	*ΔsagA*	*ccpA*	*bepA*	*Δfms21*	*fms20*	*Δfms14*	*fms17*	*Δfms13*							*gls20*	*Δgls33*	*glsB*
HPH 288	*acm*		*ΔsgrA*	*Δfibronectin*	*ΔsagA*	*ccpA*	*bepA*	*fms21*	*fms20*	*fms14*	*fms17*	*fms13*				*fms16*	*fms19*	*fms11*	*gls20*	*Δgls33*	*glsB*

Abbreviations: The virulence genes are presented in different grey tones according to the highest homology they present with. Truncated forms are represented with Δ.

Homology

	100%
	≥99%<100%
	≥97%<99%
	≥95%<97%
	≥90%<95%
	≥80%<90%

CHAPTER 5

Exploring the antimicrobial activity of Platinum Nanoparticles in Gram-positive and Gram-negative bacteria.

Chapter 5

Exploring the antimicrobial activity of Platinum Nanoparticles in Gram-positive and Gram-negative bacteria.

This study has been partially published.

Role of platinum nanozymes in the oxidative stress response of *Salmonella* Typhimurium.

Mireya Viviana Belloso Daza[1], Anna Scarsi[2,4], Francesca Gatto[2], Gabriele Rocchetti[3], Pier Paolo Pompa[2], Pier Sandro Cocconcell[i1]

1 Dipartimento di Scienze e Tecnologie Alimentari per una Filiera Agro-Alimentare Sostenibile (DISTAS), Università Cattolica del Sacro Cuore, Piacenza, Italy.

2 Nanobiointeractions & Nanodiagnostics, Istituto Italiano di Tecnologia (IIT), 16163 Genova, Italy.

3 Department of Animal Science, Food and Nutrition (DiANA), Facoltà di Scienze Agrarie, Alimentari e Ambientali, Università Cattolica del Sacro Cuore, 29122 Piacenza, Italy

4 Department of Chemistry and Industrial Chemistry, University of Genoa, Via Dodecaneso, 31 – 16146 Genova, Italy

Corresponding authors: Pier Paolo Pompa pierpaolo.pompa@iit.it , Pier Sandro Cocconcelli pier.cocconcelli@unicatt.it

Keywords: Platinum nanozymes, reactive oxygen species, oxidative stress, Salmonella, antimicrobial nanoparticles

This manuscript will be submitted to MDPI Nanomaterials.

1 Abstract

Platinum nanoparticles (PtNPs) are being explored due to their minimal side effects and potent catalytic activities like oxidases and peroxidases granting strong antimicrobial characteristics. This capability has been attributed to both their ROS-promoting ability (through their peroxidase and oxidase functionalities) and/or to their membrane-penetration capacity, however the precise mode of action is still unclear. Considering this, we initially investigated the oxidative stress response of Gram-positive *Enterococcus faecium* and *Staphylococcus aureus* and Gram-negative *Salmonella* ser. Typhimurium and *Escherichia coli* bacteria. Additionally, we analysed the effect of different bacterial survival mechanisms such as biofilm formation, quorum sensing, transcriptional regulation, and stress response. For this purpose we used *E. faecium* mutant strain lacking the enterococcal surface protein *esp,* responsible for biofilm formation and a well-known virulence marker; *S. aureus* lacking *sarA* , an important transcriptional regulator involved in many cellular processes; *E. coli* lacking *luxS*, part of the quorum sensing machinery; and *Salmonella* Typhimurium lacking ROS-coping enzymes to analyse the effect on the oxidative stress response. Our results showed that PtNPs presented limited biocidal effects against Gram-positive bacteria *E. faecium* and *S. aureus* where the effect on biofilm formation capacity and transcriptional regulation of PtNPs was not clear. The effect on *E. coli* was bacteriostatic showing link between the antimicrobial properties of PtNPs and quorum sensing mechanism deficiency. We deepened the investigation concerning the oxidative stress response of *Salmonella enterica* Typhimurium when exposed to PtNPs and their role in ROS-scavenging. The knock-out mutant HpxF$^-$ (Δ*katE* Δ*katG* Δ*katN* Δ*ahpCF* Δ*tsaA*) and its parental wild-type strain allowed us to investigate the role of the *Salmonella* ROS copy machinery. In line with other studies, our results demonstrated that, at high doses, PtNPs exert antibacterial activity, principally due to their oxidase-like properties, with significantly stronger effect on the HpxF$^-$ mutant strain compared to the limited activity on the wild-type ones, especially in aerobic conditions. Moreover, metabolomic analyses of oxidative stress markers, including lipid, glutathione, and DNA oxidation, indicated that 12023 HpxF$^-$ was not able to cope with PtNPs-based oxidative stress as efficiently as the wild-type. The more rapid ROS accumulation of the mutant strain impacts on the expression and function of efflux pumps together with the modification of the outer membrane fatty acid composition. This oxidase-induced effect compromised the bacterial membrane integrity and led to DNA damage. Conversely, when combined with other ROS like H_2O_2, PtNP peroxidase function promotes ROS scavenging, protecting bacterial cells from oxidative damage. This mechanistic study shed light on the understanding of the mechanisms of PtNPs enzyme-like activity in view of potential antimicrobial applications.

2 Introduction

The constant increase of bacterial resistance to antibiotics represents one of the greatest worldwide health challenges in the recent years, posing the necessity to find new longer-term solutions for successful control of bacterial infections. In this regard, there is a growing interest in developing and exploiting engineered nanomaterials that could integrate novel antibacterial functionalities (1). Among the currently available nanotechnology tools, metallic and metallic oxide nanostructures have been proposed as potential candidates to overcome the drawbacks of antibiotics due to their peculiar chemical properties, high surface-to-volume ratio and their potential antibacterial activity (2, 3). Several nanoparticles (NPs), like Ag, Pd, Au, Cu, ZnO and TiO_2 have shown promising results (4), even if concerns about their cytotoxicity limit their practical use (5). Silver nanoparticles (AgNPs) possess intrinsic broad-spectrum antimicrobial characteristics (6–9) and have been widely adopted as effective bactericidal agents (Buckley et al., 2010; Klasen, 2000; Tripathi & Goshisht, 2022; Rizzello & Pompa, 2014). Nevertheless, the increasing use of silver in a great number of commercial and medical tools is leading to the development of bacterial molecular strategies of resistance. Mechanisms by which bacteria become resistant to silver involve the reduction of Ag^+ to its less toxic neutral oxidation state, or the employment of active efflux from cell. Moreover, it has been recently reported that, after repeated long-term exposure to subinhibitory concentrations of AgNPs, Gram-negative bacteria can promote the aggregation of silver NPs by the production of the bacterial flagellum protein flagellin, thus evading the antibacterial effect (14). Furthermore, AgNPs present an important threat due to their toxicity to human cells (15, 16). Therefore, it is essential to exploit new antimicrobial nanomaterials with a proper biocompatibility.

Gold, platinum, and palladium-based NPs have been revealed as safer antimicrobial candidates. Evidence suggests that the antibacterial properties of noble metal-based NPs are usually attributed to their oxidase- and peroxidase-like activity. In particular, reactive oxygen species (ROS) , are able to oxidize diverse cellular components (17). As a promising alternative to natural enzymes, nanozymes are catalytic nanomaterials possessing several advantages, such as low-cost synthesis, room temperature stability, and the possibility to be employed in harsh conditions keeping high efficiency (18). Furthermore, the catalytic activity and thus the enzyme-like behaviour of NPs could be controlled tuning several features such as size, shape, and exposed facets. For example, shape and facet-dependent antibacterial activities of Pd nanocrystals have been recently reported, exhibiting high oxidase- -like properties (19). This catalytic activity efficiently inhibited the proliferation of both Gram-positive and Gram-negative bacteria, although at high NP concentrations. This behaviour could be attributed to the different membrane-penetration capacity of Pd nanocrystals, even though the precise mechanisms involved need further clarifications.

In this framework, the aim of our investigation was to perform a mechanistic study to understand the dynamics of the enzyme-like activity of 5 nm PtNPs using four risk-associated foodborne pathogens: two for Gram-positive bacteria, *Enterococcus faecium* and *Staphylococcus aureus*, and two for Gram-negative bacteria, *Escherichia coli* and *Salmonella enterica* ser. Typhimurium. These bacteria are highly associated with foodborne outbreaks and are a increasing concern of Public Health. Furthermore, we investigated the mode of action of PtNPs and how they affect biofilm formation, quorum sensing, virulence and oxidative stress by using different knock-out strains.

3 Methods and Materials

3.1 Bacterial Strains and Culture Conditions

A total of eight bacterial strains were subjected to treatment with PtNPs as described in Table 1, *E. faecium* E1162 and mutant E1162Δesp, *S. aureus* UAMS-1 and mutant strain UAMS-929, *E. coli* BW25113 and mutant strain JW2662 and *S. enterica* serovar Typhimurium 12023 and its mutant strain 12023 HpxF$^-$ were grown in brain heart infusion (BHI) at 37°C and supplemented with antibiotics when necessary.

Table 1. Bacterial strains used in this study including parental and mutant strains.

Bacteria	**Wild Type**	**Selective Agent**	**Mutant**	**Mutation**	**Selective agent**	**Function**	**Reference**
Enterococcus faecium	E1162		E1162	*Δesp*	Kan 50 mg/L	biofilm formation	Top et al. 2013
Staphylococcus aureus	UAMS-1		UAMS-929	*ΔsarA*	Kan 50 mg/L	virulence/adhesion	Beenken et al 2010
Escherichia coli	BW25113	Kan 25 mg/L	JW2662	*ΔluxS*	Kan 25 mg/L	quorum sensing	Anand and Griffiths, 2003
Salmonella typhimurium	12023		HpxF	*ΔkatE;ΔkatG;ΔkatN;ΔahpCF;ΔtsaA*	Kan 25 mg/L	oxidative stress	Hébrard et al 2009

Kan: Kanamycin

3.2 Platinum nanoparticle synthesis, functionalization, and characterization

5 nm spherical citrate-coated platinum nanoparticles were synthesized by wet chemical reduction, following a previously reported protocol with some optimizations (24). PtNP monodispersity was analyzed by Transmission Electron Microscopy (TEM) and Dynamic Light Scattering (DLS) (Fig. 1). Since we observed extensive aggregation of PtNPs when dissolved in BHI medium, PtNP stability was improved by functionalization with Bovine Serum Albumin (BSA) (Sigma-Aldrich) (25, 26). BSA-coating was performed through physical adsorption of the protein at the surface of the nanoparticles. First, the pH of the aqueous solution of 5 nm PtNPs was raised to 7.5, and then the BSA solution was added. The reaction flask was kept under stirring for 30 minutes at room temperature, and then the solution was washed using 30 K Amicon Ultra centrifugal filters. The final dose of PtNPs was determined by ICP-OES analysis. For PtNPs characterization, TEM analysis was performed by using a JEOL JE-1011 microscope with thermionic source (W filament). Accelerating voltages: 100 kV; conventional TEM imaging: bright field; TEM resolution = 4.0 Å (100 kV). For DLS analysis, PtNPs and PtNPs-BSA suspensions were diluted in MilliQ water and BHI medium up to optimal dose, and the spectra were recorded at room temperature by Zetasizer Nano Range (Malvern-PANalytical) as frequency distribution of intensity.

3.3 Bacterial growth inhibition

For the determination of the bacterial growth inhibition, different doses of PtNPs (0, 5, 10, 20, 50 and 100 mg/L) were prepared in BHI and added to a 96-well microtiter plate. Next, 0.1 OD of the selected bacteria (Table 1) were inoculated to each well containing the different PtNPs doses and grown at 37°C overnight. After incubation, the optical density of the cultures was measured spectrophotometrically at 620nm. The nalysis was performed in triplicate.

3.4 Biocidal Effect of PtNPs in *Salmonella*

In the second part of our study, we proceeded to test *Salmonella* Typhimurium parental and mutant strain with a different approach under aerobic and anaerobic conditions, performing plate counts. For the determination of the bacterial growth inhibition, different doses of PtNPs (0, 5, 10, 20, 50 and 100 mg/L) were prepared in 10ml of BHI. Next, 100µl of the selected bacteria (Table 1) were inoculated at log phase to each tube containing the different PtNPs doses and grown at 37°C overnight. After incubation, the cultures were serially diluted and plated on BHI, with respective antibiotics when necessary. Plate counts were performed in triplicate. The biocidal effect was determined in *Salmonella* as follows: log phase cultures of parental and knock-out strain were treated PtNPs at doses of 20 and 50 mg/L. After 1h of incubation, PtNPs were removed by filtering the suspension through a 0.22um cellulose filter (Merck MF-Millipore). Filters were washed with sterile saline solution and then resuspended in initial volume of BHI, serially diluted, and plated onto respective agar plates. A negative control was included.

3.5 Hydrogen peroxide scavenging and sensitivity in *Salmonella*

To test the scavenging capacity and potential sensitivity to H_2O_2, *Salmonella* strains were subjected to H_2O_2 in the presence of PtNPs. Briefly, overnight cultures of 12023 and 12023 HpxF were diluted to an OD of 0.1 and mixed with 10 µg/mL PtNPs. Next different concentrations of H_2O_2, 0, 0.001, 0.01, 0.1, 0.5, 1, 2, 5 and 10 mM were added to the solution. A control with 0 µg/mL PtNPs was included. Cells were incubated at 37 °C and optical density was measured every two hours overnight. All samples were prepared in triplicate.

3.6 Untargeted metabolomics by UHPLC-HRMS in *Salmonella*

Metabolomic profiling of the oxidative stress of *Salmonella* Typhimurium induced by PtNPs was determined using high-resolution mass spectrometry HRMS) performed on a Q-Exactive™ Focus Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific, Waltham, MA, USA) coupled to a Vanquish ultra-high-pressure liquid chromatography (UHPLC) pump and equipped with heated electrospray ionization (HESI)-II probe (Thermo Scientific, USA). Parental and mutant strains were incubated overnight with PtNPs doses of 0, 20 and 50 mg/L. Next, cell cultures were centrifuged for 10 min at 10,000 × g. Bothpellet and supernatant were kept for analysis. For the extraction step, pellet and supernatant were treated using a 1:20 ratio with the extraction buffer, consisting in 80% methanol (Carlo Erba) and 0.1% formic acid (Carlo-Erba). Thereafter, they were incubated for 10 min at maximum power using an ultrasound assisted extraction step. Samples were centrifuged at 4 °C for 10 min at 10,000 x g. For subsequent instrumental analysis, extracted pellet samples were added to UHPLC vials and supernatant samples were filtered using 0.22 micron syringe-filters before adding them to UHPLC vials. The chromatographic separation was achieved under a water-acetonitrile (both LC-MS grade, from Sigma-Aldrich, Milan, Italy) gradient elution (6–94% acetonitrile in 35 min) using 0.1% formic acid as phase modifier, on an Agilent Zorbax Eclipse Plus C18 column (50 × 2.1 mm, 1.8 µm). The HRMS conditions were adapted as described previously (27). The flow rate was 200 µL/min, and full scan MS analysis was chosen, with a positive ionization mode and a mass resolution of 70,000 at m/z 200. The injection volume was 6 µL, using an m/z range of 100–1200. The automatic gain control target (AGC target) and the maximum injection time (IT) were 1e6 and 200 ms, respectively. Randomized injections of pooled quality control (QC) samples were acquired in a data-dependent (Top N= 3) MS/MS mode with full scan mass resolution reduced to 17,500 at m/z 200, with an AGC target value of 1e5, maximum IT of 100 ms, and isolation window of 1.0 m/z, respectively. The Top N ions were selected for fragmentation under stepped (10, 20, 40 eV) Normalized Collisional Energy. The HESI parameters for both MS and MS/MS were as follows: sheath gas flow 40 arb (arbitrary units), auxiliary gas flow 20 arb, spray voltage 3.5 kV, capillary temperature 320 ◦C. Before data collection, the mass spectrometer was calibrated using Pierce™ positive ion calibration solution (Thermo Fisher Scientific, San Jose CA, USA). To avoid possible bias, the sequence of injections was randomized. The raw data

(.RAW files) were converted into .abf format using the Reifycs Abf Converter and then further processed using the software MSDIAL (version 4.38) (28). Automatic peak finding, LOWESS normalization, and annotation via spectral matching (against the database MoNA – Mass Bank of North America) were performed. The mass range 100–1200 m/z was searched for features with a minimum peak height of 10,000 cp*s.* The MS and MS/MS tolerance for peak centroiding was set to 0.01 and 0.05 Da, respectively. Retention time information was excluded from the calculation of the total score. Accurate mass tolerance for identification was 0.01 Da for MS and 0.05 Da for MS/MS. The identification step was based on mass accuracy, isotopic pattern, and spectral matching. In MS-Dial, these criteria were used to calculate a total identification score. The total identification score cut off was 60%, considering the most common HESI + adducts. Gap filling using peak finder algorithm was performed to fill in missing peaks, considering 5 ppm tolerance for m/z values. The software MS-Finder (29) was used for in-silico fragmentation of the not annotated mass features, using Lipid Maps and FoodDB libraries, thus reaching a level 2 of confidence in annotation (30). To this aim, the compounds presenting an *in-silico* prediction score > 5 were retained. Finally, the information regarding the ontology of each annotated compound was provided by the annotation softwares MS-Dial and MS-Finder.

3.7 Malondialdehyde TBARS assay in *Salmonella*

The ratio between GSH and oxidized glutathione (GSSG) clarifies on how 12023 and $HpxF^-$ would react to ROS in terms of cell membrane oxidation in the presence of sublethal PtNPs doses. Lipid oxidation was determined by the TBARS assay. Parental and mutant strains were incubated overnight with PtNPs doses of 0, 20 and 50 mg/L. After, cell cultures were centrifuged for 10 min at 10,000 × g. For the extraction step both the pellet and supernatant were treated with a 1:2 ratio of extraction buffer, that consists of 0.1% (w/v) Trichloracetic acid (TCA) in dH_2O. Samples were then incubated under ultrasound at maximum power for 10 min. Followed by a centrifugation step (10,000 rpm for 10 min) both the supernatant and pellet were added to one solution containing 20 % of TCA (w/v) and 0.65% of TBA (Thiobarbituric acid) (w/v), and to a second solution containing 20% of TCA . Samples were then mixed by inversion, incubated for 15 min at 95°C and cooled down to stop the reaction. After centrifugation for 10 min at 10,000 × g, samples were analysed using a spectrophotometer at an optical density of 532 nm. For MDA determination, a molar extinction coefficient of 155 cm−1 mM−1 was used. Results were finally expressed as nM MDA equivalents (n = 3).

3.8 Statistical Analysis

ANOVA analysis, with subsequent Tukey's significant difference test (p-value of 0.05), was used to compare the data obtained from each experiment using IBM SPSS Statistics (Version 25; IBM,

Armonk, NY, USA). All the data obtained in triplicate was reported as mean values ± standard deviation (SD).

4 Results and Discussion

4.1 Synthesis and characterization of platinum nanoparticles

The synthesized 5 nm PtNPs were monodispersed and homogeneous in shape and size, as illustrated in the TEM image (fig. 1A) and corresponding size distribution (fig. 1B). Moreover, they were stable in water, and showed a hydrodynamic radius around 8-9 nm (fig. 1C, red curve). However, when added to BHI medium, PtNPs showed a rapid and significant aggregation, causing a strong peak shift in the DLS spectrum (fig. 1C, blue curve). To increase their stability in this medium, PtNP surface was coated with BSA, through physical adsorption of the protein in aqueous solution. As illustrated in figure 1D, the presence of the adsorbed protein led to a peak shift of about 20 nm in water (red curve) and a similar peak is showed in BHI medium (blue curve), confirming that BSA coating ensured a relatively high stability of PtNPs.

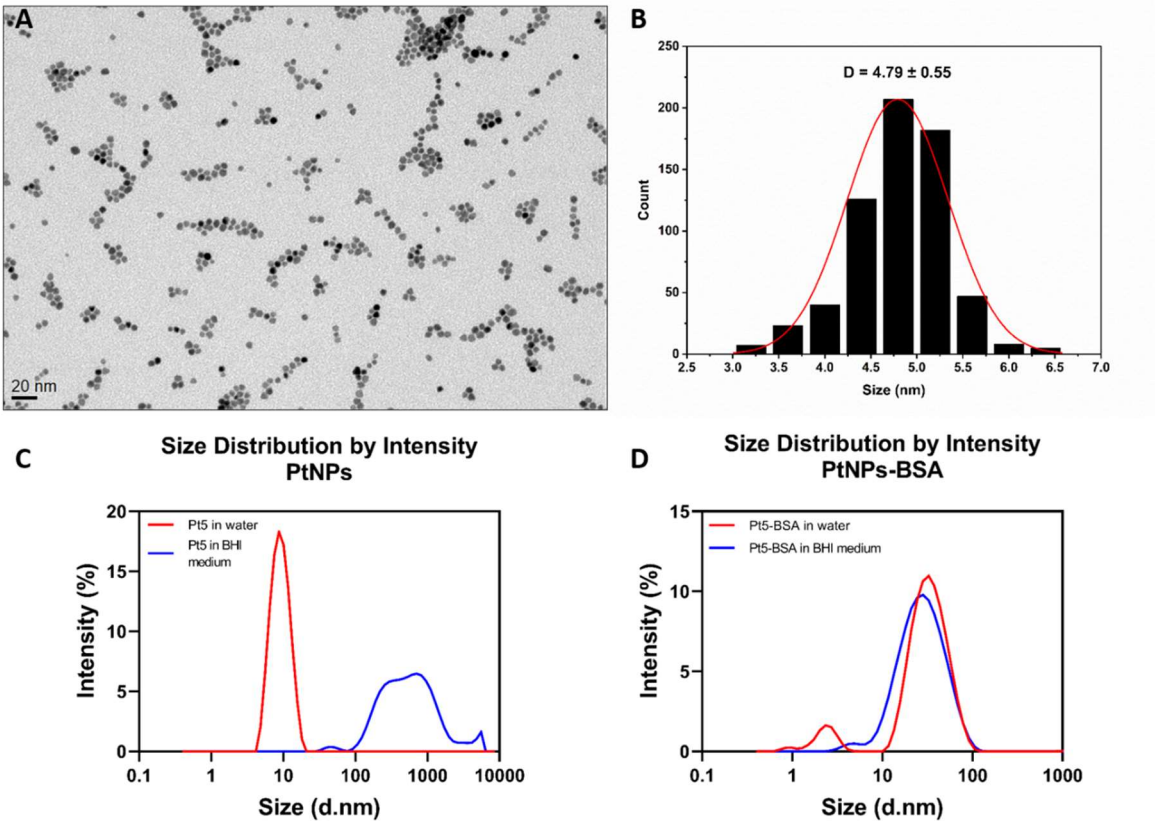


Figure 1. Characterization of synthesized nanoparticles and their stability in BHI medium. Representative TEM image (A) and relative size distribution analysis (B). DLS spectra of (C) 5 nm platinum nanoparticles and (D) 5 nm platinum nanoparticles coated with BSA, both dissolved in water (red curves) and BHI medium (blue curves). PtNPs alone are stable and monodispersed in water, but

they aggregate when added to BHI medium (C). Conversely, PtNPs-BSA show a similar peak when dissolved in water and in BHI medium (D, centered at around 30 nm), confirming the stability ensured by BSA-coating.

4.2 Antimicrobial activity of PtNPs directly affects redox homeostasis

In the initial phase of our study, we tested the bactericidal or bacteriostatic effect of PtNPs in terms of differential cellular functions (Table 1) in eight different knock-out bacteria including: *E. faecium* Δ*esp*, *S. aureus* Δ*sarA*, *E. coli* Δ*luxS, S. enterica* Typhimurium *ΔkatE katG katN ahpCF tsaA,* and respective parental strains (fig.1). In the case of *E. faecium,* both strains showed similar results without strong growth inhibition capacity at doses ≤ 64 mg/L, yet at 512 mg/L, both the parental and mutant strain show a reduction of 0.3 OD with respect to the control. These results indicate that, although there is an efficient inhibition effect of PtNPs against *E. faecium*, the absence of *esp,* involved in adherence and virulence*,* does not enhance the antibacterial action of PtNPs. Similar to our results, Vaidya and colleagues (2017) showed promising results with Pt ions against enterococcal biofilm in doses of 500 and 1000 mg/L (27). In the case of *E. faecium* exposed to PtNPs, it is important to understand the effect of a sublethal dose of PtNPs and the flatus of the *esp* gene. High doses of PtNPs (>500 mg/L) are still bactericidal and it is important to further understand the mechanisms of biofilm formation including not only the *esp* gene, but also the entire machinery for the formation of extracellular polymeric substances to assess the mode of action of PtNPs against the biofilm. Observing the results for *S. aureus*, the parental strain showed a reduction of 0.2 at 8 mg/L, with a continuous reduction to 0.7 at 512mg/L compared to the control. This is in accordance to another investigation, where an effective bactericidal effect of a dose of 101 mg/L PtNPs against *S. aureus,* especially *S. aureus* biofilms, was shown (28). Moreover, in our study, the corresponding mutant strain UAMS-929 showed a similar turbidity reduction compared to the parental strain. In this case, the deletion of *sarA* was expected to boost the bactericidal capacity of PtNPs related to oxidative stress, as the inactivation of this gene may affect mechanisms involved in oxidative stress by downregulating super-oxide dismutase sodA (29). Since SOD scavenges ROS, it was expected that the inactivation of *sarA* would downregulate the ROS scavenging mechanisms including the SOD activity (30), and so boosting the antimicrobial activity of the PtNPs . Since the reduction is similar in the mutant and the parental, there are more mechanisms that play a role in the compensation for the lack of *sarA* that protects the cell from oxidative stress. Concerning the results of *E. coli*, the parental strain remained almost unchanged with a slight decrease at 0.1 at 256 mg/L and a stronger decrease of 0.4 at 512 mg/L, when compared to the control. The mutant strain showed lower growth values, with an OD of 1.3 for the control and decreasing slightly to 0.9 and 0.8 at 128, 256 and 512 mg/L PtNPs, respectively. Several studies have proven the efficacy of PtNPs to kill *E. coli* (28,31,32), comparable to our results. The *luxS* knock-out clearly inactivates protection mechanisms and compensatory reaction chains are not strong enough to bypass the oxidative stress induced by the PtNPs inhibiting the growth, however the effect is rather bacteriostatic. This is

known as quorum sensing (QS) and uses signal molecules, termed autologous inducers (AIs). Genes *luxS* and *pfs* play an important role in QS, involved in AI-2 synthesis. The pathogenic association between bacteria and host is complex and diverse; numerous studies demonstrate that the QS system of pathogenic bacteria aids in a variety of biological functions, including biofilm formation, virulence factor production, drug resistance and adhesion (33). Finally, *S.* Typhimurium strains showed the strongest reduction compared to other bacteria. In fact, parental strain showed an OD reduction of 0.4 at 512 mg/L when compared to the control. The mutant showed an even stronger reduction of 0.8 OD at 512 mg/L when compared to the control. We deduce that the lack of ROS-coping enzymes led to the fatal outcome of the exposure of PtNPs to mutant cells. Given the fact that PtNPs are best known for the induced oxidative stress, the following chapters of our study will focus on *S.* Typhimurium.

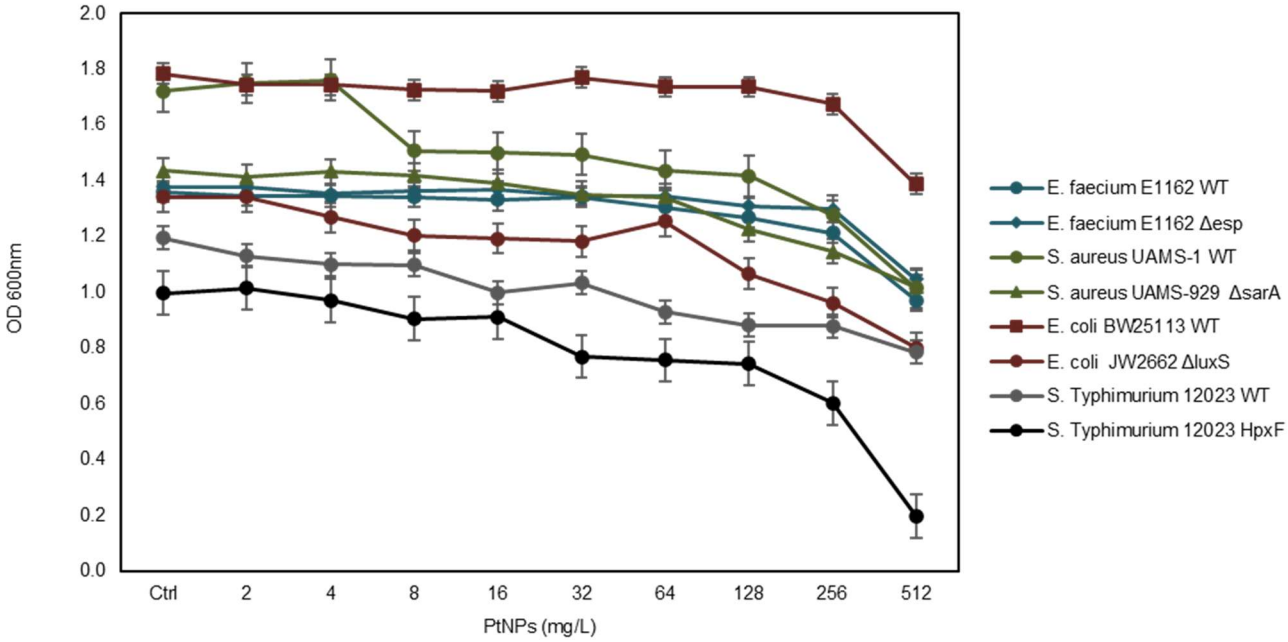


Figure 2. Minimal bacteriostatic/bactericidal doses of PtNPs measured as a function of optical density at 620nm. E. faecium E1162 wild type (blue rhomboid), E1162 Δesp (blue circle), S. aureus UAMS-1 (green circle), S. aureus UAMS-929 (green triangle), E. coli BW25113 (red square), E. coli JW2662 (red circle), S. Typhimurium 12023 (black circle) and S. Typhimurium HpxF$^-$(grey triangle).

4.3 Effect of PtNPs on *Salmonella* Typhimurium growth.

The second step of our assessment involved the evaluation of the bacterial growth inhibition capacity of PtNPs on *Salmonella* Typhimurium 12023 and its derivative mutant 12023 HpxF$^-$, to investigate if the lack of ROS-coping enzymes may result in different cell viability under aerobic and anaerobic conditions. Fig. 3 shows that the growth of 12023 was only partially limited up to 20 mg/L particle

concentration. A slightly more effective reduction is observed at 50 and 100 mg/L with 0.53 and 0.64 Log CFU reduction, respectively. No significant differences were detected between growth under aerobic or anaerobic conditions. On the contrary, a statistically significant reduction ($p<0.05$) in cell numbers was observed in 12023 HpxF$^-$ as a function of particle concentration. The mutant strain seemed to be more sensitive to the PtNPs in aerobic conditions with respect to anaerobiosis. In a second experiment, we analysed the PtNPs effect on cells exposed for 1 hour to sublethal doses of PtNPs (20 and 50mg/L) (fig. 4). After 1 hour of exposure at 20 mg/L cell viability was reduced of 0.46 and 0.45 Log CFU for parental and mutant strains, respectively. At 50 mg/L PtNPs a statistically difference was observed between the parental strain 12023 and mutant 12023 HpxF$^-$, with a 1.24 Log CFU reduction.

It has been reported that PtNPs have outstanding catalytic activity and exhibit typical kinetics of oxidases (31, 32). Interestingly, Song and colleagues reported that PtNPs coupled on deposited Multiwalled Carbon Nanotubes presented oxidase activity generating superoxide $O_2^{\bullet-}$, from dissolved O_2, rather than hydroxyl radicals $OH^{\bullet}$ or singlet oxygen $^1O_2$ (33). In fact, this catalytic activity mimics one of the best characterized sources of ROS during host cell-pathogen interactions, namely the NADPH oxidase (34). The accumulation of superoxide causes bacterial membrane lipid peroxidation, which increases the cell permeability causing the uncontrolled transport of intra- and extracellular molecules, finally leading to cell death (35). Our data indicate that the strain 12023 HpxF$^-$, with impaired response to oxidative stress, is significantly more susceptible than its parental strain. The combined effect of *Salmonella* Typhimurium ROS defence enzymes like alkyl hydroperoxide reductases AhpCF, TsaA and catalases KatE, KatG and KatN enables possibility of two types of scavenging systems that contribute to oxidative stress survival. Hence, while catalases act as the first line of oxidative stress defense by scavenging H_2O_2, alkyl hydroperoxide reductases eliminate micromolar concentrations of H_2O_2 and other hydroperoxides (21). These reductases are part of peroxiredoxins, which reduce organic hydroperoxides to alcohols and hydrogen peroxide to water at the expense of NADH or NADPH (36). Additionally, Hébard and colleagues (2009) showed that *S.* Typhimurium Kat$^-$ mutants (Δ*katE* Δ*katG* Δ*katN*) and reductase Ahp$^-$ mutants (Δ*ahpCF* Δ*tsaA*) were still able to scavenge H_2O_2 due to the compensatory regulation of the other enzymes (23). Previous studies have shown similar results with *Salmonella enterica* ser. Infantis exposed to 50 mg/L of PtNPs had a reduction of 1 Log CFU (37).

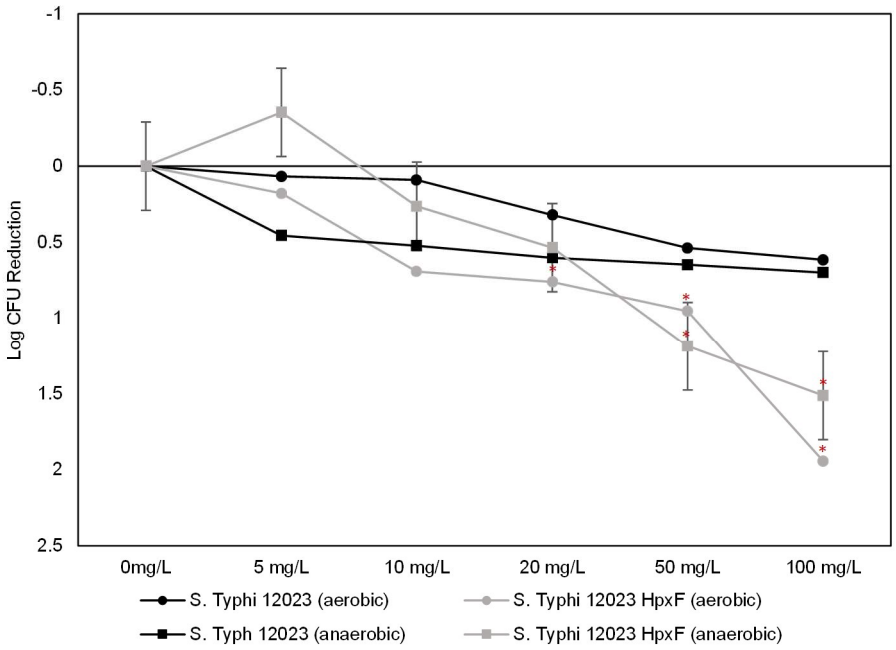


Figure 3. Growth reduction effect of PtNPs on Salmonella Typhimurium. Parental strain 12023 (black) showed a weak decrease of the growth. Mutant strain 12023 HpxF$^-$ (grey) growth decreased continuously as a function of PtNP dose with a maximum Log CFU reduction of 2 with 100 mg/L of PtNPs.

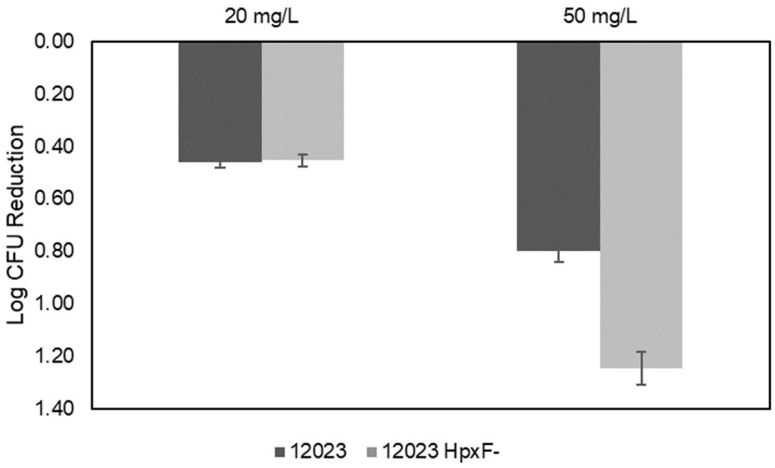


Figure 4. Biocidal effect of sublethal concentrations of PtNPs on Salmonella Typhimurium 12023 (black) and 12023 HpxF (grey). The exposure to 20 mg/L PtNPs for 1h caused a 0.2 and 0.45 Log CFU reduction for wild type and mutant strain, respectively. The 1h exposure to 50 mg/L showed an increased biocidal effect with 0.46 and 0.8 Log CFU reduction for wild type and mutant strain, respectively.

4.4 The combined effect of PtNPs and H_2O_2 on *Salmonella* Typhimurium.

After analysing the inhibitory effect of PtNPs-induced oxidative stress on *Salmonella* Typhimurium, we tested if the presence of exogenous H_2O_2 would cause the PtNPs to act as ROS scavengers and protect the cell, or as ROS enhancers by boosting the combined oxidative stress effect. To this aim, we exposed cells to increasing concentrations (0.001-10 mM) of H_2O_2 in presence or absence of the sub-inhibitory dose of 10 μg/mL PtNPs (fig.5). As shown in fig. 5A, the growth of strain 12023 was not affected by H_2O_2 equal or below 1 mM, whereas at higher concentrations (2, 5, 10 mM) of H_2O_2, the growth was significantly inhibited. Interestingly, when 10 μg/mL of PtNPs were added, a protection effect was observed, being 12023 able to grow with 2, 5 and 10 mM of H_2O_2 (fig. 5B). Conversely, the mutant 12023 HpxF$^-$ presented substantially higher susceptibility to hydrogen peroxide, with growth inhibition at H_2O_2 concentrations higher than 0.001 mM. These data are consistent with a previous study (Hébrard et al., 2009) which showed that 12023 HpxF$^-$ accumulates H_2O_2 during aerobic growth much faster than its parental strain, and when exposed to exogenous hydrogen peroxide, it was able to survive at a concentration of < 0.001 mM H_2O_2 (fig. 5C). Notably, the presence of 10 μg/mL PtNPs in the growth medium exerted a protection effect enabling the mutant strain to grow at a 10-fold higher concentration of H_2O_2 (0.01 mM, fig. 5D). This indicates that the addition of low doses of H_2O_2, does not increase oxidative stress, but protects the cells by shifting the catalytic ability of the nanozymes to scavenge exogenous ROS mimicking peroxidase activity, rather than generating them. In fact, other studies have shown that a PtNPs catalytic activity can be modulated depending on external conditions, showing peroxidase-like activity under acidic conditions and catalase or SOD-like activity under neutral and alkaline conditions (38, 39). In any case, the change in electrochemical behaviour of PtNPs after the addition of H_2O_2 was previously studied. Briefly, in the presence of PtNPs, H_2O_2 dissociates and is absorbed to $OH^{\bullet}$ on the PtNP surface. A subsequent reduction step, H_2O_2 becomes H_2O and O_2 , this capacity is attributed to the peroxidase properties of PtNPs (40). The protection capacity of PtNPs against oxidative stress has been observed in mammalian cells, particularly lung cancer cells when exposed to 100 μg/mL PtNPs and 350 μM H_2O_2, and rat skeletal L6 cells when exposed to 10 μg/mL PtNPs and 10 μM H_2O_2 (41). To the best of our knowledge, there are no studies that show the protection capacity of PtNPs in the presence H_2O_2 towards bacteria. This effect is more evident in the parental strain, being protected with the *Salmonella* innate ROS coping machinery, together with the peroxidase-like remotion of H_2O_2 by the nanoparticles. Furthermore, the results obtained with the mutant strain clarify that the peroxidase -mimicking activity of the PtNPs was enough to protect *katE katG katN ahpCF tsaA*-deficient cells by correcting the stress effects that preserve the bacterial growth. In a previous study, it was demonstrated that the failure of the induction of *ahpC*, *katG*,and *katE* gene expression in *E. coli* resulted in higher susceptibility against H_2O_2-induced oxidative stress (42). Likewise, another report by Lui and colleagues, when exposing *Salmonella* Enteritidis to 3 mM of

H_2O_2, genes *katG* and *ahpCF* boosted about 31-40 and 41-50 fold-change increase in differential expression (43).

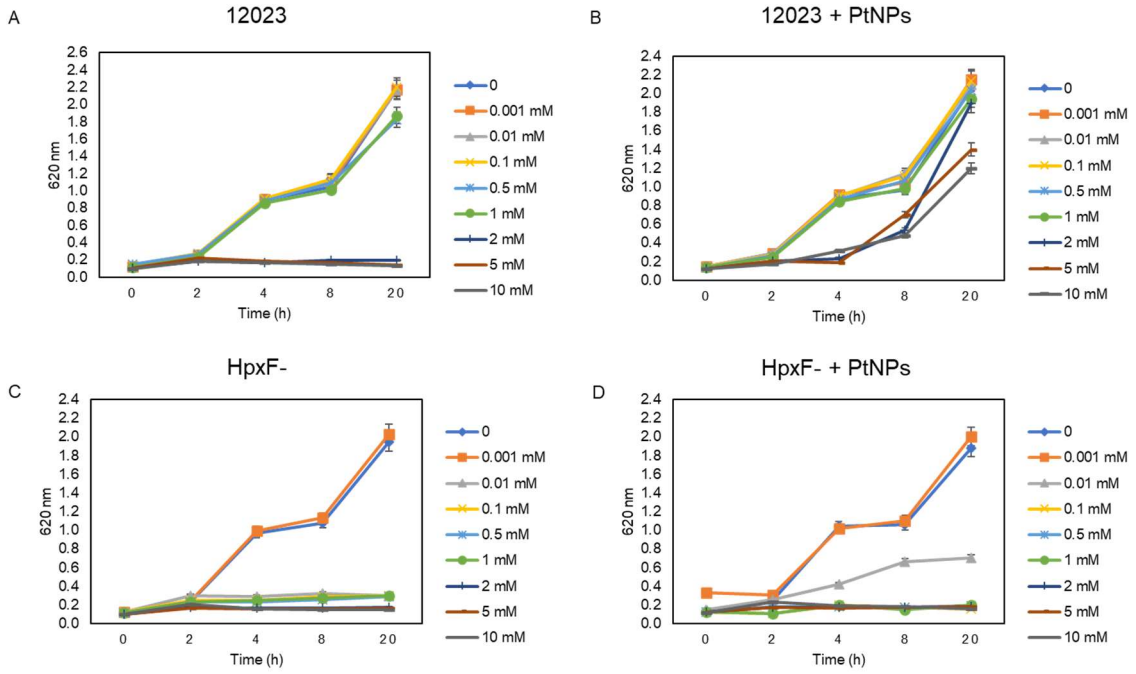


Figure 5. Exposure of S. Typhimurium parental strain 12023 to millimolar concentrations of H_2O_2 only (A) and in combination with 10 µg/mL PtNPs (B); and exposure of mutant strain 12023 HpxF$^-$ to millimolar concentrations of H_2O_2 only (C), and in combination with 10 µg/mL PtNPs (D).

4.5 PtNPs-induced ROS affect membrane lipids and oxidize DNA

Taking into consideration the previous results, untargeted metabolomics using UHPLC-HRMS was undertaken to explore the effect of PtNP-induced oxidative stress on cell metabolism. In particular, we analysed the metabolomic profile of cell extracts and supernatants of both *Salmonella* Typhimurium 12023 and 12023 HpxF$^-$ strains. As shown in fig. 5, OPLS-DA (Orthogonal Projections to Latent Structures Discriminant Analysis) scatter plot of extracts from supernatant and pelleted cells are reported. A dose response effect on the metabolomic profiles of cells and supernatants was observed, as all samples of both strains exposed to the same doses of PtNPs clustered together and were distinct from other samples treated with different doses of PtNPs. No significant differences were observed in the metabolome profiles between the two studied strains, by used this analytical approach.

In contrast, when the oxidative stress markers for Glutathione oxidation, DNA oxidation, and lipid peroxidation were analysed, the two strains responded differently to oxidative stress (fig. 7). Remarkably, our results show that the exposure dose to 20 and 50 mg/L caused the GSH/GSSG ratio to

decrease in the parental strain, suggesting that GSH oxidizes quicklier to combat oxidative stress effectively. On the contrary, the decrease of the ratio in 12023 HpxF$^-$ is slower due to a decompensation of ROS-metabolism machinery, showing a less effective way of preventing oxidative stress caused by ROS. In the presence of GSH, peroxidase-catalysed one-electron oxidation leads to ROS formation with GSSG production. Therefore, decrease of GSH levels could be supported by the pro-oxidant property of the PtNPs as well as indicating disruption of the intracellular redox state (44). Glutathione reduction systems has a key role in maintaining the reduced environment in *Enterobacteriaceae* and when ROS are generated, GSH is oxidized to GSSG, resulting in a decrease in GSH and increase in GSSG content in cellular extracts (29).

Furthermore, the DNA oxidation was analysed by measuring three different markers of oxidative damage, namely 8-oxo-G, 8-oxo-2dG and 8-oxo-2dA (45). The concentration of the three oxidative markers increased proportionally to the dose of PtNPs. However, this effect was higher in mutant strain 12023 HpxF$^-$ as shown by the ratio between the concentration of oxidative markers in the mutant and parental strain, as reported in fig. 6. Moreover, the intracellular interactions of the PtNPs with bacterial DNA have been previously reported with *Salmonella* Enteritidis, and it was shown that PtNPs affected the DNA through ROS damage (46). . A commonly used marker for oxidative stress damage is 8-oxo-7,8-dihydroguanosine (8-oxo-G), the most commonly observed oxidation product of a guanine base (Figure 1). Guanine has a lower reduction potential than the other DNA bases. Nearby oxidized bases can therefore readily be repaired by an electron transfer from the guanine base to the oxidized base (Candeias and Steenken, 1993), effectively transferring the oxidation site onto the guanine. 8-oxo-G can base-pair with both a cytosine or adenine base, thus affecting the coding potential of DNA (Cheng et al., 1992). Crystal structures of a *Bacillus stearothermophilus* DNA polymerase with 8-oxo-G pairing to both the cognate and mismatched base showed that under these conditions the proofreading mechanisms of the DNA polymerase are no longer effective (Hsu et al., 2004). In fact, Hsu *et al.* observed an inversion of the ordinary mismatch recognition. The in theory canonical 8-oxo-G:cytosine base-pair behaved as a mismatch, whereas the 8-oxo-G:adenine base-pair was recognized as cognate, which ultimately leads to a G to T transversion in the replicated DNA strand. (47)

Stress-related lipids oxidation was qualitatively assessed by untargeted metabolomics with markers for cell envelope fatty acid compounds in *Salmonella,* as function of the fold change of mutant over parental strain (fig. 6). Overall, at 50 mg/L PtNPs the fold change is higher when compared to the lower PtNP doses. In addition to this, previous studies have stated that H_2O_2 up-regulates efflux pumps conferring an additional protection towards ROS (48). Thus, the mutant strain lacking the catalases and reductases accumulates ROS in a more rapid manner downregulating the expression and the function of efflux pumps, and so causing terminal damage to the cell envelope. Other studies show that, under exogenous stress, *Salmonella* Typhimurium modifies its outer membrane fatty acid composition, increasing

unsaturated oleic, and linoleic acid and resulting to a enhance fluidity of the cellular membrane that affects permeabilization (49, 50),

This is in accordance with our results when analysing the high degree of lipid peroxidation of the mutant strain in comparison to its parental strain (fig. 7). Moreover, the determination of MDA, the benchmark method for ROS-induce lipid-peroxidation (51) showed that in parental strain 12023, no MDA was detected at any dose of PtNPs, whereas in mutant strain 12023 HpxF$^-$, gradually increasing values of 0.001, 0.002 and 0.004 nM of MDA were detected at 0, 20 and 50 mg/L, respectively.

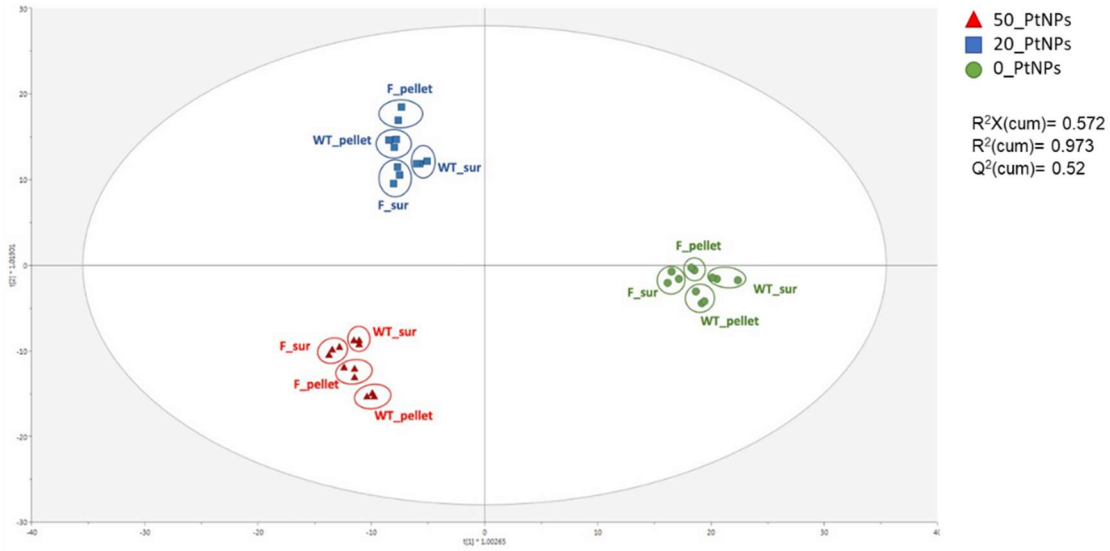


Figure 6. Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot based on the metabolomic profile of treated and untreated supernatant and cells of parental strain 12023 and mutant strain 12023 HpxF$^-$. The exposure was carried out with 0 mg/L (green), 20 mg/L (blue) and 50 mg/L (red) of PtNPs. WT_sur: parental strain supernatant sample; WT_pellet: parental strain cell sample; F_sur: mutant strain supernatant sample; F_pellet: mutant strain cell sample.

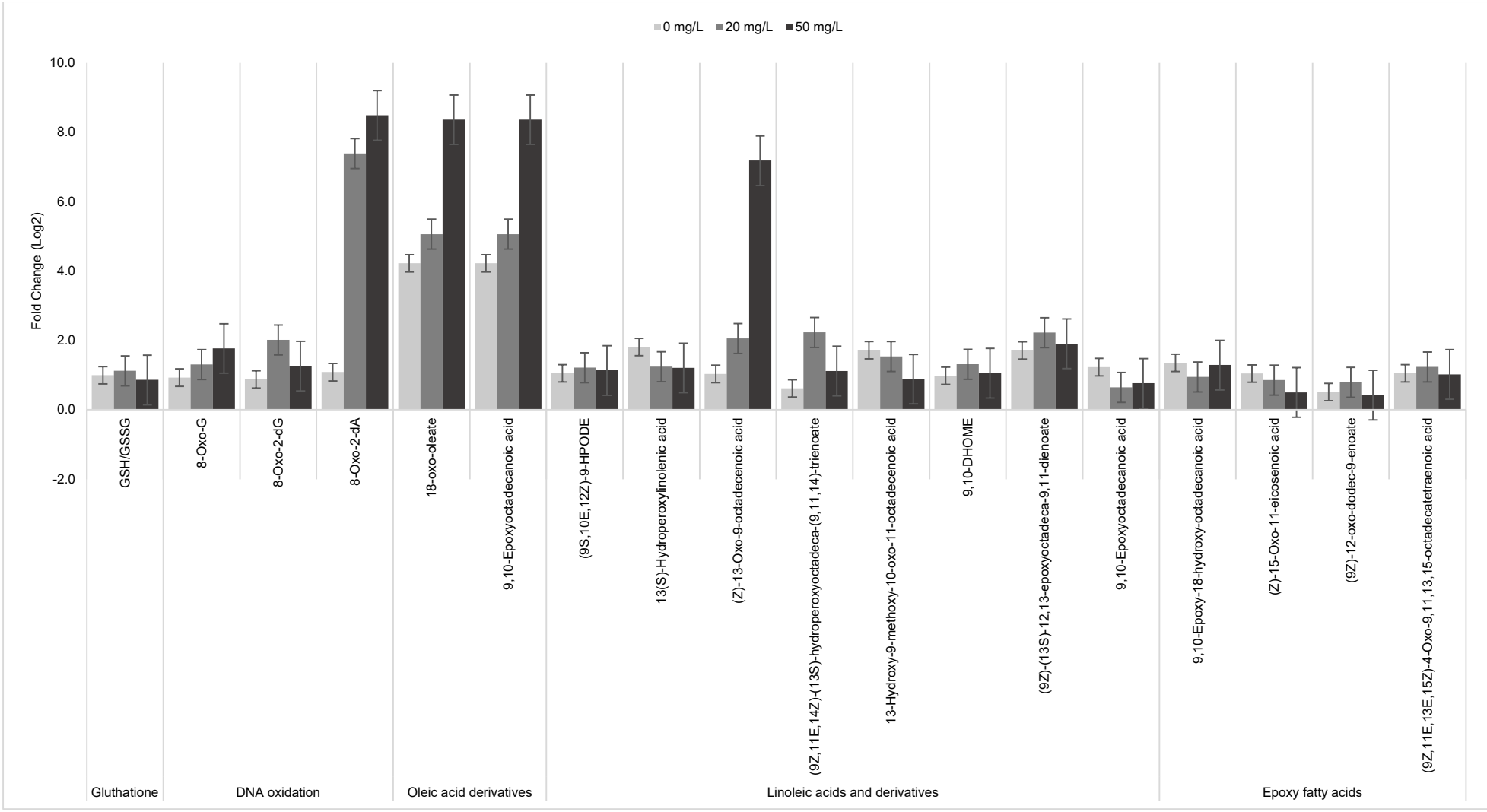


Figure 7. Fold change of the mutant 12023 HpxF$^-$ in comparison to parental strain 12023 at a PtNPs dose of 0 mg/L (light grey), 20 mg/L (grey) and 50 mg/L

(dark grey). Overall cellular oxidation state was determined as the ratio between reduced (GSH) and oxidize (GSSG) Glutathione. DNA oxidation was determined by markers 8-oxo-G, 8-oxo-2-dG and 8-oxo-2-dA. Lipid peroxidation was assessed with markers for oleic acid, linoleic acid derivatives and epoxy fatty acids.

5 Conclusion

Platinum nanoparticles are of high interest due to their strong catalytic activity that mimics oxidase and peroxidase enzymes, making them great candidates for antimicrobial agents. Here, we analysed the effect of different bacterial survival mechanisms such as biofilm formation, quorum sensing, transcriptional regulation, and stress response. The limited biocidal effects of PtNPs against Gram-positive bacteria *E. faecium* and *S. aureus* suggests that additional machinery is required to target biofilm formation and transcriptional regulation. The effect on *E. coli* was bacteriostatic showing link between the antimicrobial properties of PtNPs and quorum sensing mechanism deficiency. We further investigated the oxidative stress response of *Salmonella* enterica Typhimurium when exposed to PtNPs and their role in ROS scavenging. In agreement with other studies, our results showed that PtNPs exert antibacterial activity at high doses, mainly due to their oxidase-like properties, with a significantly higher effect on the mutant strain, lacking ROS-coping mechanisms, compared to the limited activity on the wild type, especially under aerobic conditions. This result was supported by metabolomic analyses of oxidative stress markers, including lipids, glutathione and DNA oxidation. In contrast, the modulation of the catalytic function of PtNP is shifted to a peroxidase-like behaviour in conjunction with H_2O_2, where ROS are scavenged and thus, protects bacterial cells from oxidative damage. This mechanistic study sheds light on understanding the mechanisms of PtNP enzymatic activity in view of its potential antimicrobial applications, which contrary to other evidence, does not pose exceptional antibacterial activity when combined with other oxidative stress causing agents.

6 Transparency Declaration

The authors declare that there are no conflicts of interest. This work received no specific grant from any funding agency. We thank Prof. Aussel from Aix-Marseille Université for kindly providing us with *Salmonella* Typhimurium mutant and parental strains; and Prof. Top from UMC Utrecht for providing *Enterococcus faecium* mutant and parental strain.

7 References

1. Ozdal M, Gurkok S. 2022. Recent advances in nanoparticles as antibacterial agent. ADMET DMPK 10:115.

2. Manke A, Wang L, Rojanasakul Y. 2013. Mechanisms of nanoparticle-induced oxidative stress and toxicity. Biomed Res Int 2013.

3. Sánchez-López E, Gomes D, Esteruelas G, Bonilla L, Lopez-Machado AL, Galindo R, Cano A, Espina M, Ettcheto M, Camins A, Silva AM, Durazzo A, Santini A, Garcia ML, Souto EB. 2020.

Metal-Based Nanoparticles as Antimicrobial Agents: An Overview. Nanomaterials 10:292.

4. Beyth N, Houri-Haddad Y, Domb A, Khan W, Hazan R. 2015. Alternative antimicrobial approach: nano-antimicrobial materials. Evid Based Complement Alternat Med 2015.

5. Xiong P, Huang X, Ye N, Lu Q, Zhang G, Peng S, Wang H, Liu Y. 2022. Cytotoxicity of Metal-Based Nanoparticles: From Mechanisms and Methods of Evaluation to Pathological Manifestations. Adv Sci (Weinheim, Baden-Wurttemberg, Ger 9.

6. Dror-Ehre A, Mamane H, Belenkova T, Markovich G, Adin A. 2009. Silver nanoparticle-E. coli colloidal interaction in water and effect on E. coli survival. J Colloid Interface Sci 339:521–526.

7. El Badawy AM, Silva RG, Morris B, Scheckel KG, Suidan MT, Tolaymat TM. 2011. Surface charge-dependent toxicity of silver nanoparticles. Environ Sci Technol 45:283–287.

8. Dallas P, Sharma VK, Zboril R. 2011. Silver polymeric nanocomposites as advanced antimicrobial agents: classification, synthetic paths, applications, and perspectives. Adv Colloid Interface Sci 166:119–135.

9. Navarro E, Piccapietra F, Wagner B, Marconi F, Kaegi R, Odzak N, Sigg L, Behra R. 2008. Toxicity of silver nanoparticles to Chlamydomonas reinhardtii. Environ Sci Technol 42:8959–8964.

10. Klasen HJ. 2000. A historical review of the use of silver in the treatment of burns. II. Renewed interest for silver. Burns 26:131–138.

11. Buckley JJ, Lee AF, Olivi L, Wilson K. 2010. Hydroxyapatite supported antibacterial Ag3PO4 nanoparticles. J Mater Chem 20:8056–8063.

12. Tripathi N, Goshisht MK. 2022. Recent Advances and Mechanistic Insights into Antibacterial Activity, Antibiofilm Activity, and Cytotoxicity of Silver Nanoparticles. ACS Appl Bio Mater 5:1391–1463.

13. Rizzello L, Pompa PP. 2014. Nanosilver-based antibacterial drugs and devices: Mechanisms, methodological drawbacks, and guidelines. Chem Soc Rev 43:1501.

14. Panáček A, Kvítek L, Smékalová M, Večeřová R, Kolář M, Röderová M, Dyčka F, Šebela M, Prucek R, Tomanec O, Zbořil R. 2018. Bacterial resistance to silver nanoparticles and how to overcome it https://doi.org/10.1038/s41565-017-0013-y.

15. Wu J, Hou S, Ren D, Mather PT. 2009. Antimicrobial properties of nanostructured hydrogel webs containing silver. Biomacromolecules 10:2686–2693.

16. Holt KB, Bard AJ. 2005. Interaction of silver(I) ions with the respiratory chain of Escherichia coli: an electrochemical and scanning electrochemical microscopy study of the antimicrobial mechanism of

micromolar Ag+. Biochemistry 44:13214–13223.

17. Gao F, Shao T, Yu Y, Xiong Y, Yang L. 2021. Surface-bound reactive oxygen species generating nanozymes for selective antibacterial action. Nat Commun 12:745.

18. Pedone D, Mauro Moglianetti abc, Elisa De Luca ab, Giuseppe Bardi ab, Paolo Pompa P. 2017. Platinum nanoparticles in nanobiomedicine. Chem Soc Rev 46:4951.

19. Fang G, Li W, Shen X, Perez-Aguilar JM, Chong Y, Gao X, Chai Z, Chen C, Ge C, Zhou R. 2018. Differential Pd-nanocrystal facets demonstrate distinct antibacterial activity against Gram-positive and Gram-negative bacteria. Nat Commun 9:129.

20. Rana K, Nayak SR, Bihary A, Sahoo AK, Mohanty KC, Palo S ku, Sahoo D, Pati S, Dash P. 2021. Association of quorum sensing and biofilm formation with Salmonella virulence: story beyond gathering and cross-talk. Arch Microbiol 2021 20310 203:5887–5897.

21. Fu J, Qi L, Hu M, Liu Y, Yu K, Liu Q, Liu X. 2017. Salmonella proteomics under oxidative stress reveals coordinated regulation of antioxidant defense with iron metabolism and bacterial virulence. J Proteomics 157:52–58.

22. Rhen M. 2019. Salmonella and Reactive Oxygen Species: A Love-Hate Relationship. J Innate Immun https://doi.org/10.1159/000496370.

23. Hébrard M, Viala JPM, Méresse S, Barras F, Aussel L. 2009. Redundant hydrogen peroxide scavengers contribute to Salmonella virulence and oxidative stress resistance. J Bacteriol 191:4605–14.

24. Moglianetti M, De Luca E, Pedone D, Marotta R, Catelani T, Sartori B, Amenitsch H, Retta SF, Pompa PP. 2016. Platinum nanozymes recover cellular ROS homeostasis in an oxidative stress-mediated disease model. Nanoscale 8:3739–3752.

25. De Luca E, Pedone D, Moglianetti M, Pulcini D, Perrelli A, Retta SF, Pompa PP. 2018. Multifunctional Platinum@BSA-Rapamycin Nanocarriers for the Combinatorial Therapy of Cerebral Cavernous Malformation. ACS Omega 3:15389–15398.

26. Yu S, Cui Y, Guo X, Chen S, Sun H, Wang L, Wang J, Zhao Y, Liu Z. 2019. Biocompatible bovine serum albumin stabilized platinum nanoparticles for the oxidation of morin. New J Chem 43:8774–8780.

27. Rocchetti G, Michelini S, Pizzamiglio V, Masoero F, Lucini L. 2021. A combined metabolomics and peptidomics approach to discriminate anomalous rind inclusion levels in Parmigiano Reggiano PDO grated hard cheese from different ripening stages. Food Res Int 149:110654.

28. Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, Vandergheynst J, Fiehn O, Arita M. 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods 2015 126 12:523–526.

29. Tsugawa H, Kind T, Nakabayashi R, Yukihira D, Tanaka W, Cajka T, Saito K, Fiehn O, Arita M. 2016. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. Anal Chem 88:7946–7958.

30. Salek RM, Steinbeck C, Viant MR, Goodacre R, Dunn WB. 2013. The role of reporting standards for metabolite annotation and identification in metabolomic studies. Gigascience 2.

31. Jin P, Niu X, Gao Z, Xue X, Zhang F, Cheng W, Ren C, Du H, Manyande A, Chen H. 2021. Ultrafine Platinum Nanoparticles Supported on Covalent Organic Frameworks As Stable and Reusable Oxidase-Like Catalysts for Cellular Glutathione Detection. ACS Appl Nano Mater 4:5834–5841.

32. Deng H, He S, Lin X, Yang L, Lin Z, Chen R, Peng H, Chen W. 2019. Target-triggered inhibiting oxidase-mimicking activity of platinum nanoparticles for ultrasensitive colorimetric detection of silver ion. Chinese Chem Lett 30:1659–1662.

33. Song H, Ma C, Wang L, Zhu Z. 2021. Platinum Nanoparticles-Deposited Multiwalled Carbon Nanotubes as a NADH Nanoscale characterization and applications † https://doi.org/10.1039/D0NR04060F.

34. Spooner R, Yilmaz Ö. 2011. The Role of Reactive-Oxygen-Species in Microbial Persistence and Inflammation. Int J Mol Sci 12:334.

35. Wang L, Hu C, Shao L. 2017. The antimicrobial activity of nanoparticles: present situation and prospects for the future. Int J Nanomedicine 12:1227.

36. Horst SA, Jaeger T, Denkel LA, Rouf SF, Rhen M, Bange FC. 2010. Thiol Peroxidase Protects Salmonella enterica from Hydrogen Peroxide Stress In Vitro and Facilitates Intracellular Growth. J Bacteriol 192:2929.

37. Chlumsky O, Purkrtova S, Michova H, Svarcova V, Slepicka P, Fajstavr D, Ulbrich P, Demnerova K. 2020. The effect of gold and silver nanoparticles, chitosan and their combinations on bacterial biofilms of food-borne pathogens. Biofouling 36:222–233.

38. Liu Y, Wu H, Chong Y, Wamer WG, Xia Q, Cai L, Nie Z, Fu PP, Yin JJ. 2015. Platinum Nanoparticles: Efficient and Stable Catechol Oxidase Mimetics. ACS Appl Mater Interfaces 7:19709–19717.

39. Liu Y, Wu H, Li M, Yin JJ, Nie Z. 2014. PH dependent catalytic activities of platinum nanoparticles

with respect to the decomposition of hydrogen peroxide and scavenging of superoxide and singlet oxygen. Nanoscale 6:11904–11910.

40. Mazzotta E, Di Giulio T, Mastronardi V, Pompa PP, Moglianetti M, Malitesta C. 2021. Bare Platinum Nanoparticles Deposited on Glassy Carbon Electrodes for Electrocatalytic Detection of Hydrogen Peroxide. ACS Appl Nano Mater 4:7650–7662.

41. Nakanishi H, Hamasaki T, Kinjo T, Yan H, Nakamichi N, Kabayama S, Teruya K, Shirahata S. 2013. Low concentration platinum nanoparticles effectively scavenge reactive oxygen species in rat skeletal L6 cells. Nano Biomed Eng 5:76–85.

42. Jung IL, Kim IG. 2003. Transcription of ahpC, katG, and katE genes in Escherichia coli is regulated by polyamines: Polyamine-deficient mutant sensitive to H2O2-induced oxidative damage. Biochem Biophys Res Commun 301:915–922.

43. Liu X, Omar M, Abrahante JE, Nagaraja K V., Vidovic S. 2020. Insights into the oxidative stress response of salmonella enterica serovar enteritidis revealed by the next generation sequencing approach. Antioxidants 9:1–20.

44. Fanoro OT, Parani S, Maluleke R, Lebepe TC, Varghese RJ, Mgedle N, Mavumengwana V, Oluwafemi OS. 2021. Biosynthesis of smaller-sized platinum nanoparticles using the leaf extract of combretum erythrophyllum and its antibacterial activities. Antibiotics 10:1275.

45. Seixas AF, Quendera AP, Sousa JP, Silva AFQ, Arraiano CM, Andrade JM. 2022. Bacterial Response to Oxidative Stress and RNA Oxidation. Front Genet 12:2789.

46. Sawosz E, Chwalibog A, Szeliga J, Sawosz F, Grodzik M, Rupiewicz M, Niemiec T, Kacprzyk K. 2010. Visualization of gold and platinum nanoparticles interacting with Salmonella Enteritidis and Listeria monocytogenes. Int J Nanomedicine 5:631.

47. Fasnacht M, Polacek N. 2021. Oxidative Stress in Bacteria and the Central Dogma of Molecular Biology. Front Mol Biosci 8:392.

48. Bogomolnaya LM, Andrews KD, Talamantes M, Maple A, Ragoza Y, Vazquez-Torres A, Andrews-Polymenis H. 2013. The ABC-Type Efflux Pump MacAB Protects Salmonella enterica serovar Typhimurium from Oxidative Stress. MBio 4.

49. Redfern BH. 2019. The Aquila Digital Community The Aquila Digital Community The Incorporation of Lipids into the Cellular Membrane of The Incorporation of Lipids into the Cellular Membrane of Salmonella Salmonella.

50. Wang RY, Ou Y, Zeng XA, Guo CJ. 2019. Membrane fatty acids composition and fluidity modification in Salmonella Typhimurium by culture temperature and resistance under pulsed electric

fields. Int J Food Sci Technol 54:2236–2245.

51. Chanana V, Majumdar S, Rishi P. 2007. Involvement of caspase-3, lipid peroxidation and TNF-α in causing apoptosis of macrophages by coordinately expressed Salmonella phenotype under stress conditions. Mol Immunol 44:1551–1558.

CHAPTER 6

General Conclusions

Chapter 6

General Conclusions

Food quality and safety is of paramount importance for both consumer health and the food industry. Microbes can contaminate food in different ways at different stages of food production, therefore, detection and monitoring are crucial. The increasing demand for sensitive, accessible, and cost-effective methods has driven technological innovation including DNA-based method, enabling the development of increasingly efficient molecular tests as they offer the ability to instantly detect and characterise microbes that are critical for food quality and safety. While traditional microbiological methods are labour-intensive and time-consuming, DNA-based methods are better suited for rapid and accurate detection. However, molecular methods also have their limitations, as they usually focus on one or a few characteristics of a microorganism and can be compromised by inhibitors in the food matrix. In recent years, improvements in the performance of molecular methods have been evaluated and developed, finding new strategies to overcome limitations in sensitivity and efficiency, and improving the rapid visualisation of results. In addition, whole genome sequencing has extended the great potential of molecular methods by making genome sequencing available and opening the possibility of in-depth investigation. The high discriminatory power of WGS is unmatched by any other technique. The choice between a rapid molecular technique and WGS depends on the goal to be achieved. Chapters two to four were mainly focused on the main problems of surveillance and timely detection of foodborne bacteria, namely taxonomic identification, pathogenic potential, and antimicrobial resistance, which are faced by the WGS approach. It is important to denote that the presence of multi-drug resistant strains in ready-to-eat fermented food pose a risk of public health for the spread of AMR determinants in the food chain and in the gut microbiota of consumers and that in silico bioinformatic accurately assess the safety of *Enterococcus faecium* strain UC7251, encouraging the development of innovative strategies for the mitigation of the risk related to antimicrobial resistance diffusion in food. Additionally, WGS-derived taxonomy and population structure analysis addressed the definition of species in bacteria and show how the genomic approach makes the difference in taxonomic assignment, especially for species identification, a difficult step in the study of microbes. A combined genome-based approach was successfully applied to the reclassification of the taxonomic unit *E. faecium* clade B, with a higher percentage similarity to *E. lactis* compared to *E. faecium* clades A1 and A2. The study also highlights the limitations of 16S rRNA-based taxonomy and points to the great potential of the application of WGS in alternative analyses like dDDH, ANI and pangenomics. The deep understanding of pathogenic potential and population structure permits the development of rapid detection strategies by designing DNA specific probes. In the fourth chapter, the designed and validated PCR to identify *E. faecium* and *E. lactis* species, based on the pangenome-derived *gluP* gene showed high resolution and discriminatory power. To note that published primers widely used for years and designed to identify *E. faecium* (e.g., ddl gene) lack enough discriminatory power to distinguish these species. The development of a precise differentiation method has direct implications in both the clinical and food safety fields and could draw the line between *E. faecium* strains currently being used in probiotics and feed that actually correspond to *E. lactis* and isolates associated with human infections when they are actually *E. lactis* having possible implications in

infection management and overall, in different Public Health contexts. Furthermore, genome-derived studies have paved the path to understand microbial dynamics against stress and to thrive in suboptimal conditions. In the final chapter we analysed the effect of different bacterial survival mechanisms such as biofilm formation, quorum sensing, transcriptional regulation, and stress response systems when exposing bacteria, lacking different genes coding for any of these mechanisms, to Platinum nanoparticles. The limited biocidal effects of PtNPs against Gram-positive bacteria *E. faecium* and *S. aureus* suggests that additional machinery is required to target biofilm formation and transcriptional regulation. The effect on *E. coli* was bacteriostatic showing link between the antimicrobial properties of PtNPs and quorum sensing mechanism deficiency. We further investigated the oxidative stress response of *Salmonella* enterica Typhimurium when exposed to PtNPs and their role in ROS scavenging. Especially in aerobic conditions, PtNPs has a higher effect on the mutant strain due to their oxidase-like properties and metabolomic analyses of oxidative stress markers supported the fact of lipids, glutathione and DNA oxidation. In contrast, the modulation of the catalytic function of PtNP is shifted to a peroxidase-like behaviour in conjunction with H_2O_2 protecting bacterial cells from oxidative damage. The mechanisms of PtNP enzymatic activity as potential antimicrobial applications, which contrary to other evidence, does not pose exceptional antibacterial activity when combined with other oxidative stress causing agents. To conclude, this work explored the power of genome derived information and analysis thereof for detection, surveillance and mitigation of foodborne pathogens and the need to further harmonize genomic investigation across the scientific community to achieve food safety and in a wider view to safeguard public health.