High-Quality Genome Sequence Assembly of R.A73 Enterococcus faecium Isolated from Freshwater fish mucus Exhibiting high Probiotic Potential

Rim Jeni, Kais Ghedira, Monia Bour, Sonia Abdelhak, Alia Kahla, Balkiss Bouhaouala-Zahar

High-Quality Genome Sequence Assembly of R.A73 Enterococcus faecium Isolated from Freshwater fish mucus Exhibiting high Probiotic Potential

Rim El Jeni
Institut National des Sciences et Technologies de la Mer, Salammbo Tunis / Institut Pasteur de Tunis
Kais Ghedira
Institut Pasteur de Tunis
Monia El Bour
Institut National des Sciences et Technologies de la Mer, Salammbo Tunis
Sonia Abdelhak
Institut Pasteur de Tunis
Alia Ben Kahla
Institut Pasteur de Tunis
Balkiss Bouhaouala-Zahar (balkiss.bouhaouala@fmt.utm.tn)
Institut Pasteur de Tunis https://orcid.org/0000-0003-3147-245X

Research article

Keywords: Enterococcus faecium, Bacteriocin, Freshwater fish, Lactic acid bacteria, Whole-genome sequencing.

DOI: https://doi.org/10.21203/rs.3.rs-25205/v2

Abstract

Background: Whole-genome sequencing using high throughput technologies has revolutionized and speeded up the scientific investigation of bacterial genetics, biochemistry, and molecular biology. Lactic acid bacteria (LABs) have been extensively used in fermentation and more recently as probiotics in food products that promote health. Genome sequencing and functional genomics investigations of LABs varieties provide rapid and important information about their diversity and their evolution, revealing a significant molecular basis.

This study investigates the whole genome sequences of the *Enterococcus faecium* genome (HG937697), isolated from the mucus of freshwater fish in Tunisian dams. Genomic DNA was extracted using the Quick-GDNA kit (Zymoresearch) and sequenced using the Illumina HiSeq2500 system. Sequences quality assessment was performed using FastQC software. The complete genome annotation was carried out with the RAST web server then NCBI PGAAP.

Results: The *Enterococcus faecium R.A73* assembled in 28 contigs consisting of 2,935,283 bps. The genome annotation revealed 2,884 genes in total including 2,834 coding sequences and 50 RNAs containing 3 rRNAs (one rRNA 16s, one rRNA 23s and one rRNA 5s) and 47 tRNAs. Twenty-two genes implicated in bacteriocin production are identified within *the Enterococcus faecium R.A73* genome.

Conclusions: Data obtained provide insights to further investigate the effective strategy for testing this *Enterococcus faecium R.A73* strain in the industrial manufacturing process. Studying their metabolism with bioinformatics tools represents the future challenge and contribution to improving the utilization of the multi-purpose bacteria in food.

Background

Nowadays, LABs are widely used for animal and human health diseases and becoming available as probiotics [1]. *Enterococcus* is a LABs large genus, ubiquitous, having the capacity to adapt to harsh environments. These species are isolated from different habitats including water (i.e. waste, freshwater, and seawater), soil, plants, and the digestive tract of warm-blooded animals and/or humans [2]. Due to the nature of some opportunistic pathogen strains, the use of enterococci in food is being challenged. Over recent years, particular efforts have been devoted to the incidence of nosocomial infections caused by enterococci and their multi-resistance to antibiotics [3]. However, several studies have demonstrated the beneficial effects of *Enterococcus faecium* as probiotic in humans, animals, and aquatic culture[4–9].

Recently, the preselected *Enterococcus faecium R. A73* strain, isolated strain from freshwater fish mucus, has proven to have specific probiotic properties [1]. In the current study, we performed whole-genome sequencing of *Enterococcus faecium R.A73* and investigate the genome contents and gene functions through comparison to related species.

Results

E. faecium R.A73 genome annotation

Genome content

The present draft genome includes 2,935,283 bases, with a GC content of 38.0%, and was assembled into 28 scaffolds. The Genomic annotations illustrated a total number of 2,884 genes, corresponding to 2,834 coding sequences (CDSs) and 50 RNAs with single predicted copies of the 16S, 23S, and 5S rRNA genes and 47 predicted tRNAs (Fig. 1). A total of 342 RAST genome sub-systems were identified, with many features of carbohydrates subsystem, including the genes involved in the metabolism of central carbohydrate, amino sugars, di- and oligosaccharides, the carbon metabolism, organic acids, the fermentation metabolism, sugar alcohols, polysaccharides, and monosaccharides. There are also many amino acids and derivative characteristics of the sub-system, including the lysine, threonine, methionine, and cysteine.

Functional annotation

A total of 2,063 protein-coding genes (72.58 % of the total protein-coding genes) were assigned a putative function (by COGs). Genes associated with carbohydrate transport and metabolism (294 ORFs), translation (206 ORFs), and transcription (205 ORFs) were ranked among the most abundant COG functional categories. The genes distribution into COG functional categories is summarized in (Fig. 2).

Bacteriocin and antibacterial peptide production genes

Several genes involved in bacteriocin production as well antibacterial peptides were identified in *Enterococcus faecium R.A73* genome. These genes include colicin V CvpA family protein (DTX73_02515), production bacteriocin pole, antibacterial peptides agents synthesis, bacteriocin-associated protein (DTX73_02925), bacteriocin immunity protein (DTX73_04250, DTX73_06025, DTX73_06505), bacteriocin (DTX73_04255, DTX73_06475, DTX73_06480, DTX73_07350, DTX73_09680, DTX73_09715), EntF family bacteriocin induction factor (DTX73_06500), TmhB bacteriocin enhancer peptide (DTX73_09690), ThmA bacteriocin (DTX73_09695), ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain (DTX73_09710), class IIb bacteriocin, lactobin A/cerein 7B family (DTX73_09720). Other genes possess different roles implicated in amidophosphoribosyl transferase (EC 2.4.2.14) (DTX73_12820), acetyl-coenzyme A chain carboxyl beta transferase (EC 6.4.1.2) (DTX73_04140), a synthase dihydrofolate (EC 6.3.2.12) (DTX73_05510), an rRNA pseudouridine synthase a (EC 4.2.1.70) (DTX73_10445) and the bifunctional folylpolyglutamate synthase/dihydrofolate synthase (EC 6.3.2.17) (DTX73_05510). Furthermore, the genome revealed the presence of a gene encoding for one enterocin (DTX73_06510).

Antibiotics resistance and virulence genes

We used the ResFinder-2.1 server [10] available at cge.cbs.dtu.dk/services/ResFinder/ in combination with PGAAP and RAST server annotation [11] to investigate genes involved in resistance to antibiotics and toxic compounds in the *E. faecium* R.A73 genome.

Two genes involved in resistance to antibiotics and toxic compounds were identified. These genes correspond to an homolog of aac(6')-Ii involved in Aminoglycoside resistance (% identity: 98.36; Query/HSP length: 549/549; Accession number: L12710) and an homolog to msr(C) involved in MLS - Macrolide, Lincosamide and Streptogramin B (% identity: 97.70; Query/HSP length: 1479/1479; Accession number: AF313494). Besides, PGAAP and RAST annotation systems were also able to detect 52 other genes potentially involved in virulence, disease, and defense mechanisms. These genes found in the HG937697 genome are presented in (Table 2).

Phylogeny and classification

Based on rDNA 16S sequences, the phylogenetic tree showed that the *R.A73* is more similar to *E. faecium* LMG 11423 and *E. durans* NBRC 100479 than other *Enterococcus* species (Fig. 3).

The later analysis combined to the *in silico* DNA-DNA hybridization method confirmed its identification as *E. faecium* species. Indeed, DNA-DNA hybridization is considered as the best indicator for distinguishing species. The probabilities of DDH value higher than 70% detected through logistic regression under three formulae indicate that *E. faecium R.A73* is different from other species of the genus excepting *Enterococcus faecium*. A DDH value > 96 % was found following the comparison against *E. faecium* T110 (Supplementary data Table S1).

Comparative genomics

Comparative analysis of genome sequences

The comparative genomics help to understand several aspects related to the pathogenicity, the resistance to antibiotics, and probiotic characteristics.

Enterococcus faecium protein sequences predicted by the PGAAP annotation system, have been retrieved and compared with 14 protein sequences of related organisms corresponding to *Enterococcus* 7L76 uid197170, *Enterococcus casseliflavus* This20 uid55693, *Enterococcus faecalis* 62 159663 uid, *Enterococcus faecalis* D32 171261 uid, *Enterococcus faecalis* og1RF54927 uid, *Enterococcus faecalis Symbioflor* 1 uid183342, *Enterococcus faecalis* V583 uid57669, *Enterococcus faecium* AUS0004 uid87025, *Enterococcus faecium* AUS0085 uid214432, *Enterococcus faecium* do uid55353, *Enterococcus faecium* NRRL B 2354 uid188477, *Enterococcus hirae ATCC* 9790 uid70619, *Enterococcus mundtii* that 25 uid229420 and *Enterococcus faecium* T110.

The comparative proteome analysis against 14 other *Enterococcus* related genomes (Table 1) showed that HG937697 is highly similar to the genome of *E. faecium* T110 with 2,318 common orthologs genes (80.37 %). We also identified 208 protein-coding genes that were specific to *E. faecium R.A73* strain. The BRIG tool confirmed our previous finding with a very high similarity between *Enterococcus faecium* R.A73 and *Enterococcus faecium* T110 genomes (Figure 4).

Comparative analysis of virulence genes

We investigated the presence of genes related to virulence in *Enterococcus faecium R.A73*. Among several *Enterococcus* virulence genes available in the virulence factor database VFDB (http://www.mgc.ac.cn/VFs/), 30 genes, including the gene for enterococcal surface protein (esp), were absent in *Enterococcus faecium* R.A73 while EbpA (DTX73_01685), EbpB (DTX73_01690), EbpC (DTX73_01695), srtC (DTX73_017000), EcbA (DTX73_00685), EfaA (DTX73_03830) were found to be present.

Table 1 Genome size and gene count of 14 pathogens and probiotics *Enterococcus* species used in genome comparative study

Species	Genome size (Mb)	Gene count
Enterococcus 7L76 uid197170		
Enterococcus casseliflavus This20 uid55693		
Enterococcus faecalis 62 159663 uid	3.13	3158
Enterococcus faecalis D32 171261 uid	3.06	3174
Enterococcus faecalis og1RF54927 uid	2.73	2676
Enterococcus faecalis Symbioflor 1 uid183342	2.81	2761
Enterococcus faecalis V583 uid57669	3.35	3412
Enterococcus faecium AUS0004 uid87025	3.01	3118
Enterococcus faecium AUS0085 uid214432	3.23	3318
Enterococcus faecium do uid55353	3.05	3209
Enterococcus faecium NRRL B 2354 uid188477		
Enterococcus hirae ATCC 9790 uid70619	2.85	2752
Enterococcus mundtii that 25 uid229420	3.35	3229
Enterococcus faecium T110		

Table 2 Number of genes associated with the general COG functional *Enterococcus faecium* R.A73 genome categories

Code	Value	% Of total features	Description
J	206	7.14	Translation
A	0	0.0	RNA processing and modification
K	205	7.10	Transcription
L	109	3.77	Replication, recombination, and repair
B	0	0.0	Chromatin structure and dynamics
D	29	1.00	Cell cycle control, mitosis and meiosis
Y	0	0.0	Nuclear structure
V	62	2.14	Defense mechanisms
T	80	2.77	Signal transduction mechanisms
M	120	4.16	Cell wall/membrane biogenesis
N	12	0.41	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	17	0.58	Intracellular trafficking and secretion
O	63	2.18	Posttranslational modification, protein turnover, chaperones
C	73	2.53	Energy production and conversion
G	294	10.19	Carbohydrate transport and metabolism
E	145	5.02	Amino acid transport and metabolism
F	75	2.60	Nucleotide transport and metabolism
H	72	2.49	Coenzyme transport and metabolism
I	73	2.53	Lipid transport and metabolism
P	91	3.15	Inorganic ion

			transport and metabolism
Q	18	0.62	Secondary metabolites biosynthesis transport and catabolism
R	140	4.8	General function prediction only
S	183	6.34	Function unknown
-	821	28.48	Not in COGs

Discussion

A genomics study was performed in a preselected *Enterococcus faecium R.A73* strain, which showed probiotic characteristics and an interesting efficacy in its use as food additives.

The complete genome annotation was performed using the online server RAST (Rapid Annotation using Subsystem Technology) [12].

The plasmid was absent in *Enterococcus faecium R.A73*. It may be due to growing temperature, copies number as well as isolation methods [13].

There are many features of the carbohydrate subsystem in *Enterococcus faecium R.A73*. The degradation of carbohydrates and their related compounds are mainly responsible for the primary metabolic activity of lactic acid bacteria, resulting in energy and carbon source molecules [14, 15]. Further metabolic activities such as proteins, lipids, and other compounds decomposition are important for normal growth. LABs metabolic activities include carbohydrate metabolism, protein metabolism, lipid metabolism, and other metabolic activities.

LAB needs amino acids and peptides to respond to their nitrogen complex [16]. Amino acids and peptides may be obtained through proteases or proteolysis actions. In such actions, peptides are metabolized to free amino acids and other compounds for further use. Due to the requirements of peptide differences, peptides can either be essential growth promoters or stimulating factors, some strains can grow up independently. LAB amino acid requirements are strain-dependent with a large range of species differences [17, 18]. Interestingly, many amino acids and derivatives characteristic of the subsystem, including lysine, threonine, methionine, and cysteine, exist in *Enterococcus faecium R.A73*. Bacteriocins and protein-coding for ABC transporters have been detected as well. These latter are known to have an antibacterial activity that may contribute to probiotic potential in such strains [19].

Enterococcus faecium genome identified 22 genes involved in bacteriocin production as well as antimicrobial peptides. There are several genes implicated in colicin V (ColV). Colicin V is an antibiotic,

which has a naturally occurring peptide to kill sensitive cells, disrupting potential membrane structures. It is produced by certain members of *Enterobacteriaceae* to kill bacterial cells and reduce the competition for vital nutrients [20, 21]. Furthermore, the protein-coding enterocin gene was detected in the *R.A73* genome. A notable property of enterocins is their activity against pathogenic bacteria, including members of the genera *Listeria*, *Clostridium*, and *Staphylococcus* [22–24]. This and many other enterocins, therefore, have great utility in the food industry [25].

The comparative proteomes analysis showed that *E. faecium R.A73* presented 208 specific genes including five bacteriocins (bacteriocin (DTX73_07350, DTX73_09680), ThmB bacteriocin enhancer peptide (DTX73_09690), ThmA bacteriocin (DTX73_09695), ABC-type bacteriocin/lantibiotic exporters, contain a N-terminal double-glycine peptidase domain (DTX73_09710), class IIb bacteriocin, lactobin A/cerein 7B family (DTX73_09720)). Lantibiotics that constitute a group of bacteriocins were shown to have several pharmaceutical applications including Blood pressure treatment, inflammations and allergies treatment, Skin, mastitis, herpes infections treatment, dental caries treatment, and peptic ulcer treatment. ThmA/ThmB known as termophilin 13 produced by *S. thermophiles* SPi13 possesses natural antimicrobial activities [21, 26, 27]. Furthermore, fifty-one genes out of 208 were assigned to COG functional categories associated with carbohydrate transport and metabolism (6 genes), amino acid transport and metabolism (6 genes), and cell wall/membrane/envelope biogenesis (5 genes).

Comparative proteome analysis using BLASTp best reciprocal hits with an E-value <= 1E-05 between *Enterococcus faecium R.A73* and related Enterococcus strains was performed. The comparison showed that *R.A73* is more close to *E. faecium* T110 with 2,318 common orthologs genes (80.37 %). Moreover, the probabilities of DDH value > 70 % accessed *via* logistic regression under three formulae indicate that *E. faecium R.A73* is different from other species of the genus excepting *Enterococcus faecium*. The comparison against *E. faecium* T110 corresponded to the highest probability of DDH value (> 96 %) and confirmed the previous result. Similarly, the BRIG platform showed a similarity between *Enterococcus faecium R.A73* and *Enterococcus faecium* T110 genomes (Fig. 4). This latter strain is a commercially probiotic widely prescribed for humans, animals, and aquaculture [7].

The presence of virulence genes in the probiotic strain *Enterococcus faecium* R. A73 was studied. Some virulent genes were found highly homologs to EbpA (DTX73_01685), EbpB (DTX73_01690), EbpC (DTX73_01695), srtC (DTX73_017000), EcbA (DTX73_00685), EfaA (DTX73_03830).

The *ebp* locus consists of an operon of three genes, *ebpA*, *ebpB*, and *ebpC*, which encode the pilus subunits, or pilins, and *srtC*, encoding a class C sortase [28]. *Enterococcus pili* have been demonstrated to be directly involved in biofilm development. Mutations in these genes showed a strong defect in biofilm formation and the initial adherence to the host tissue [29]. *Ebp* pili may play a role during colonization of the mammalian host, adherence to abiotic surfaces, or bacterial surface components [30] rather than the pathogenesis, although their exact biological functions remain to be determined.

Conclusions

Such results respond to potential probiotic properties. This bacterial strain can be safely used as bio-ingredients in conservation and fish processing consumed by humans and animals. Marine microbiology fields are still evolving and significant progress can be expected on marine pollution issues including bacterial oil degradation, which is under investigation at present.

Methods

Bacterial strain

In total, 177 LABs have been isolated from different organs in freshwater fish (*Mugil cephalis* and *Oreochromis niloticus*). The purified bacterial isolates have been tested for morphological and biochemical characterization and their antimicrobial activity against twenty-three fish pathogenic indicators causing human food toxicity.

The genetic identification was performed through 16S rDNA bacterial gene amplification.

The novel R.A73, isolated from Tilapia *Oreochromis niloticus* mucus, was identified as *Enterococcus faecium*.

The *in vitro* probiotic potential was carried out for the most active lactic acid bacteria, 4 strains chosen from the initial collection among which *Enterococcus faecium* R. A73 exhibited high inhibitory activities against food-borne pathogens and spoilage microbial species and has significant probiotic profiles, since it survived at pH 3.0 and in the presence of bile salts, pancreatin, and pepsin, without any detectable hemolytic activity. Further, moderate heat resistance, adhesion ability to steel surfaces, and sensitivity to clinically relevant antimicrobial agents were revealed for these novel isolate [1].

The *in vivo* probiotic effect on Nile tilapia (*Oreochromis niloticus*), mainly for this selected probiotic strain *Enterococcus faecium R.A73*, was evaluated. Fish diet maintained improved for a period of 20 days, in aquariums and supplement or not with probiotic. The results obtained, showed increase in protein contents and in lysozyme activity for fish fed with a regime fortified by probiotic. There is no obvious effect of probiotic on water quality including the temperature, pH and dissolved oxygen in treatment using the probiotic. Tilapia survival rate was 100% (data currently submitted).

Growth conditions and DNA preparation/isolation

Enterococcus faecium R.A73 was inoculated in Man-Rogosa-Sharpe (MRS) broth for 48 h at 20 °C. Pure genomic DNA was then extracted using the Quick-GDNA kit (Zymo Research) and subsequently sent it to the platform service “BaseClear” in Netherlands, for whole genome sequencing.

Genome sequencing

Enterococcus faecium R.A73 genome has been sequenced using the Illumina HiSeq 2500 system. FASTQ paired-end sequence data files have been generated using the Illumina CASAVA pipeline version 1.8.3.

Initial quality assessment was based on the data that passed Illumina chastity filtering. Readings with adapters and/or the PhiX control signal were then deleted. The second assessment of quality based on the remaining reads was performed using the FASTQC quality control tool version 0.10.0. FASTQ sequence quality has been enhanced by removing the low-quality bases, with the ”Trim Sequences” options from CLC Genomics Version 7.0.4.

De novo assembly

The quality-filtered sequence reads were assembled in some contig sequences. The analysis was carried out by using the option "*De novo* Assembly" in the genomics workshop CLC version 7.0.4. The optimal k-mer size was automatically determined using KmerGenie [31]. Contigs were then linked to each other’s and put into scaffolds or supercontigs. The orientation, order, or distance between the contigs was estimated by using the insert size between the paired-end.

The scaffolding has been performed using the SSPACE Premium scaffolder version 2.3 [32]. Gapped regions within the scaffolds were partially closed in an automated manner using GapFiller version 1.10 [33]. The method takes advantage of the insert size between the paired-end reads.

Genome annotation

We used RAST (Rapid Annotation using Subsystem Technology) webserver [12] to perform genome annotation. Briefly, protein-coding genes were predicted using the Classic RAST annotation scheme [12]. RNAmmer tool [34] was used to predict ribosomal RNAs, while tRNAs can-SE [35] was used to detect transfer RNAs. The NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/) was used to perform a final annotation.

Functional annotation

Clusters of Orthologous Group were assigned based on comparative proteomes analysis against the COG database [36] using protein sequences that have previously been predicted by PGAAP. Briefly, using the best reciprocal hits approach with an e-value <= 1E-05, protein sequences were retrieved and compared against the protein sequences available in the COG database.

Phylogenetic analysis and genome-to-genome distance calculation

Identification of closely related strains to *E. faecium R.A73* was performed based on Basic Local Alignment Search Tool (BLAST) searches and pairwise global sequence alignments through the well-curated EzTaxon database; which covers not only type strains of prokaryotic species with validly published names but also phylotypes that may represent species in nature. The 16S rDNA gene sequences with pairwise similarity higher than 96% to *E. faecium* R.A73 (locus_tag="DTX73_13310") were chosen for phylogenetic tree construction. 16S rDNA sequences were downloaded from the National Center for Biotechnology Information (NCBI) database. They were aligned using Muscle [37] as part of the MEGA7 [38] software to generate 1000 bootstrap replicates followed by a search for the best-scoring

Maximum Likelihood (ML) tree. This latter was displayed, manipulated, and annotated using iTOL 3 [39]. Digital DDH similarities between *the E. faecium R.A73* genome and those of other *Enterococcus* species were calculated using the GGDC web server version 2.0 under the recommended setting [40].

Comparative genomics

Genome comparison of *E. faecium* HG937697 with related species was performed using BRIG (Blast Ring Image Generator), an open-source multi-platform software application, which displays multi-genome comparisons and similarity between the reference genome at the center of one image compared to other related genomes listed in (**Table 1)**, in the form of a concentric colored ring set according to BLAST identity [41].

Furthermore, protein sequences of *E. faecium R.A73* that were predicted by RAST and PGAAP annotation system were extracted and compared to protein sequences of the proteomes of related *Enterococcus* cited in (**Table 1)**. The comparison was computed using Inparanoid (http://InParanoid.sbc.su.se) [42] then MultiParanoid (http://multiparanoid.cgb.ki.se/) [43] Perl programs to identify the cluster of orthologous genes between pairs of species than between all the species, respectively.

Genbank submission

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession QOVC00000000. The version described in this paper is version QOVC01000000.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests

Funding

This work was jointly supported by the ENPI BIOVecQ P.S.1.3_08 project coordinated by Institut National des Sciences et Technologies de la Mer (INSTM), Tunisia, B3Aqua laboratory, and the FP7 ERAWIDE,

GM_NCD_InCo project coordinated by Pasteur Institute of Tunis, (IPT), LGBMO laboratory.

Authors' contributions

RELJ carried out the molecular experiments and generated the data; REJ, KG carried out bioinformatic data analysis and interpretation and drafted the manuscript; MEB contributed to the microbiological analysis; SA, ABK facilitated the NGS designed experiments; BBZ designed the experiments and contributed to the data analysis/interpretation and final manuscript preparation. All authors read and approved the final manuscript.

Acknowledgments

We thank Dr. Saloua Sadok, director of Laboratoire Biotechnology Bleu et Bioproduits Aquatiques (B3Aqua), INSTM, for her constant encouragement.

References

1. El-Jeni R, El Bour M, Calo-Mata P, Böhme K, Fernández-No IC, Barros-Velázquez J, et al. In vitro probiotic profiling of novel Enterococcus faecium and Leuconostoc mesenteroides from Tunisian freshwater fishes. Can J Microbiol. 2015;62:60–71.
2. Manero A, Blanch AR. Identification of Enterococcus spp. with a biochemical key. Appl Environ Microbiol. 1999;65:4425–30.
3. Allerberger F, Küenburg B. Organization of control of nosocomial infections in Central Eastern European countries. Wiener Medizinische Wochenschrift. 2019;169:1–2.
4. Bednorz C, Oelgeschläger K, Kinnemann B, Hartmann S, Neumann K, Pieper R, et al. The broader context of antibiotic resistance: Zinc feed supplementation of piglets increases the proportion of multi-resistant Escherichia coli in vivo. Int J Med Microbiol. 2013;303:396–403. doi:10.1016/j.ijmm.2013.06.004.
5. Liu W, Zhang L, Shi J, Yi H, Zhang Y, Zhang S, et al. Assessment of the safety and applications of bacteriocinogenic Enterococcus faecium Y31 as an adjunct culture in North-eastern Chinese traditional fermentation paocai. Food Control. 2015;50:637–44. doi:10.1016/j.foodcont.2014.10.004.
6. Barbosa J, Borges S, Teixeira P. Selection of potential probiotic Enterococcus faecium isolated from Portuguese fermented food. Int J Food Microbiol. 2014;191:144–8. doi:10.1016/j.ijfoodmicro.2014.09.009.
7. Natarajan P, Parani M. First complete genome sequence of a probiotic Enterococcus faecium strain T-110 and its comparative genome analysis with pathogenic and non-pathogenic Enterococcus faecium genomes. J Genet Genomics. 2015;42:43–6. doi:10.1016/j.jgg.2014.07.002.
8. Lodemann U, Strahlendorf J, Schierack P, Klingspor S, Aschenbach JR, Martens H. Effects of the Probiotic Enterococcus faecium and Pathogenic Escherichia coli Strains in a Pig and Human

Epithelial Intestinal Cell Model . Scientifica (Cairo). 2015;2015:1–10.

9. Klingspor S, Martens H, Çaushi D, Twardziok S, Aschenbach JR, Lodemann U. Characterization of the effects of Enterococcus faecium on intestinal epithelial transport properties in piglets. J Anim Sci. 2013;91:1707–18.
10. Kleinheinz KA, Joensen KG, Larsen MV. Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and E. coli virulence genes in bacteriophage and prophage nucleotide sequences . Bacteriophage. 2014;4:e27943.
11. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014;42:206–14.
12. Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: Rapid annotations using subsystems technology. BMC Genomics. 2008;9:1–15.
13. Casey J, Daly C, Fitzgerald GF. Chromosomal integration of plasmid DNA by homologous recombination in Enterococcus faecalis and Lactococcus lactis subsp. lactis hosts harboring Tn919. Appl Environ Microbiol. 1991;57:2677–82.
14. Sánchez B, Chaignepain S, Schmitter JM, Urdaci MC. A method for the identification of proteins secreted by lactic acid bacteria grown in complex media. FEMS Microbiol Lett. 2009;295:226–9.
15. Hoefnagel MHN, Starrenburg MJC, Martens DE, Hugenholtz J, Kleerebezem M, Van Swam II, et al. Metabolic engineering of lactic acid bacteria, the combined approach: Kinetic modelling, metabolic control and experimental analysis. Microbiology. 2002;148:1003–13.
16. Savijoki K, Ingmer H, Varmanen P. Proteolytic systems of lactic acid bacteria. Appl Microbiol Biotechnol. 2006;71:394–406.
17. Letort C, Juillard V. Development of a minimal chemically-defined medium for the exponential growth of Streptococcus thermophilus. J Appl Microbiol. 2001;91:1023–9.
18. Tomita S, Saito K, Nakamura T, Sekiyama Y, Kikuchi J. Rapid discrimination of strain-dependent fermentation characteristics among Lactobacillus strains by NMR-based metabolomics of fermented vegetable juice. PLoS One. 2017;12.
19. Lebeer S, Vanderleyden J, De Keersmaecker SCJ. Genes and Molecules of Lactobacilli Supporting Probiotic Action. Microbiol Mol Biol Rev. 2008;72:728–64.
20. Gérard F, Pradel N, Wu LF. Bactericidal activity of colicin V is mediated by an inner membrane protein, SdaC, of Escherichia coli. J Bacteriol. 2005;187:1945–50.
21. Marciset O, Jeronimus-Stratingh MC, Mollet B, Poolman B. Thermophilin 13, a nontypical antilisterial poration complex bacteriocin, that functions without a receptor. J Biol Chem. 1997;272:14277–84.
22. Aymerich T, Garriga M, Ylla J, Vallier J, Monfort JM, Hugas M. Application of enterocins as biopreservatives against Listeria innocua in meat products. J Food Prot. 2000;63:721–6.
23. Giraffa G, Carminati D, Tarelli GT. Inhibition of listeria innocua in milk by bacteriocin-producing enterococcus faecium 7C5. J Food Prot. 1995;58:621–3.

24. Moreno MRF, Leisner JJ, Tee LK, Ley C, Radu S, Rusul G, et al. Microbial analysis of Malaysian tempeh, and characterization of two bacteriocins produced by isolates of Enterococcus faecium. J Appl Microbiol. 2002;92:147–57.
25. Javed I, Ahmed S, Manam S, Riaz M, Ahmad B, Ishtiaq Ali M, et al. Production, characterization, and antimicrobial activity of a bacteriocin from newly isolated enterococcus faecium IJ-31. J Food Prot. 2010;73:44–52.
26. Gilbreth SE, Somkuti GA. Thermophilin 110: A bacteriocin of Streptococcus thermophilus ST110. Curr Microbiol. 2005;51:175–82.
27. Rossi F, Marzotto M, Cremonese S, Rizzotti L, Torriani S. Diversity of Streptococcus thermophilus in bacteriocin production; inhibitory spectrum and occurrence of thermophilin genes. Food Microbiol. 2013;35:27–33. doi:10.1016/j.fm.2013.02.006.
28. Sillanpää J, Chang C, Singh K V., Montealegre MC, Nallapareddy SR, Harvey BR, et al. Contribution of Individual Ebp Pilus Subunits of Enterococcus faecalis OG1RF to Pilus Biogenesis, Biofilm Formation and Urinary Tract Infection. PLoS One. 2013;8:1–20.
29. Freitas AR, Tedim AP, Novais C, Coque TM, Peixe L. Distribution of putative virulence markers in Enterococcus faecium: towards a safety profile review. J Antimicrob Chemother. 2018;73:306–19.
30. Börner RA, Kandasamy V, Axelsen AM, Nielsen AT, Bosma EF. Genome editing of lactic acid bacteria: Opportunities for food, feed, pharma and biotech. FEMS Microbiol Lett. 2019;366.
31. Chikhi R, Medvedev P. Informed and automated k-mer size selection for genome assembly. Bioinformatics. 2014;30:31–7.
32. Boetzer M, Henkel C V., Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics. 2011;27:578–9.
33. Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. Genome Biol. 2012;13.
34. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 2007;35:3100–8.
35. Lowe TM, Eddy SR. TRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1996;25:955–64.
36. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin E V., et al. The COG database: An updated vesion includes eukaryotes. BMC Bioinformatics. 2003;4:1–14.
37. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
38. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33:1870–4.
39. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44:W242–5.
40. Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci. 2010;2:117–

34.
41. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. BMC Genomics. 2011;12.
42. O'Brien KP, Remm M, Sonnhammer ELL. Inparanoid: A comprehensive database of eukaryotic orthologs. Nucleic Acids Res. 2005;33 DATABASE ISS.:476–80.
43. Alexeyenko A, Tamas I, Liu G, Sonnhammer ELL. Automatic clustering of orthologs and inparalogs shared by multiple proteomes. Bioinformatics. 2006;22:e9-15. doi:10.1093/bioinformatics/btl213.

Supplementary Information

Table S1 The DDH probabilities to distinguish between R.A73 strain and reference strains that belonged to the similar Enterococcus genus

Figures

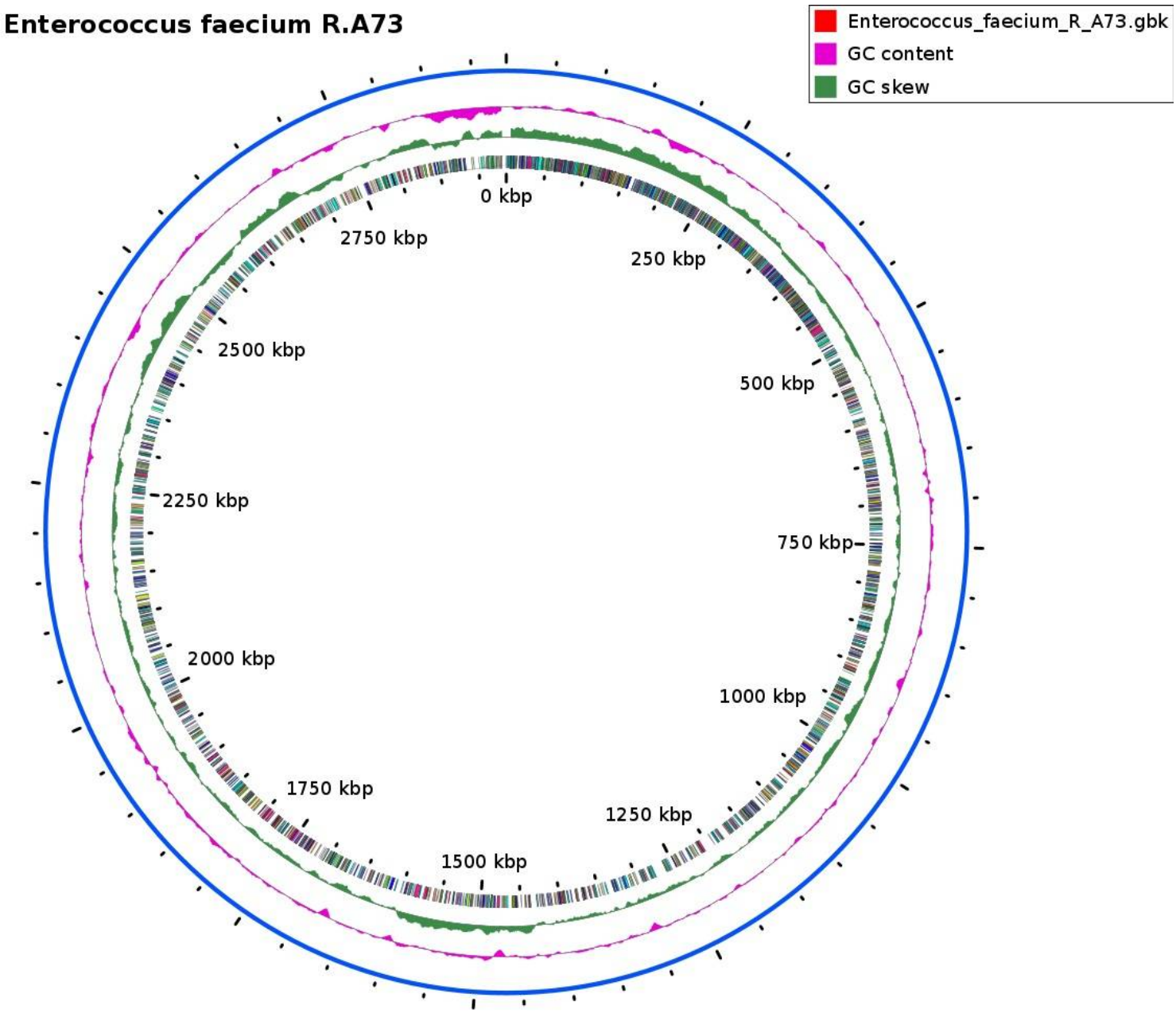

Figure 1

Enterococcus faecium R.A73 genomic annotations Genes total number of illustrations. The green color shows the GC skew and the pink color shows the GC content.

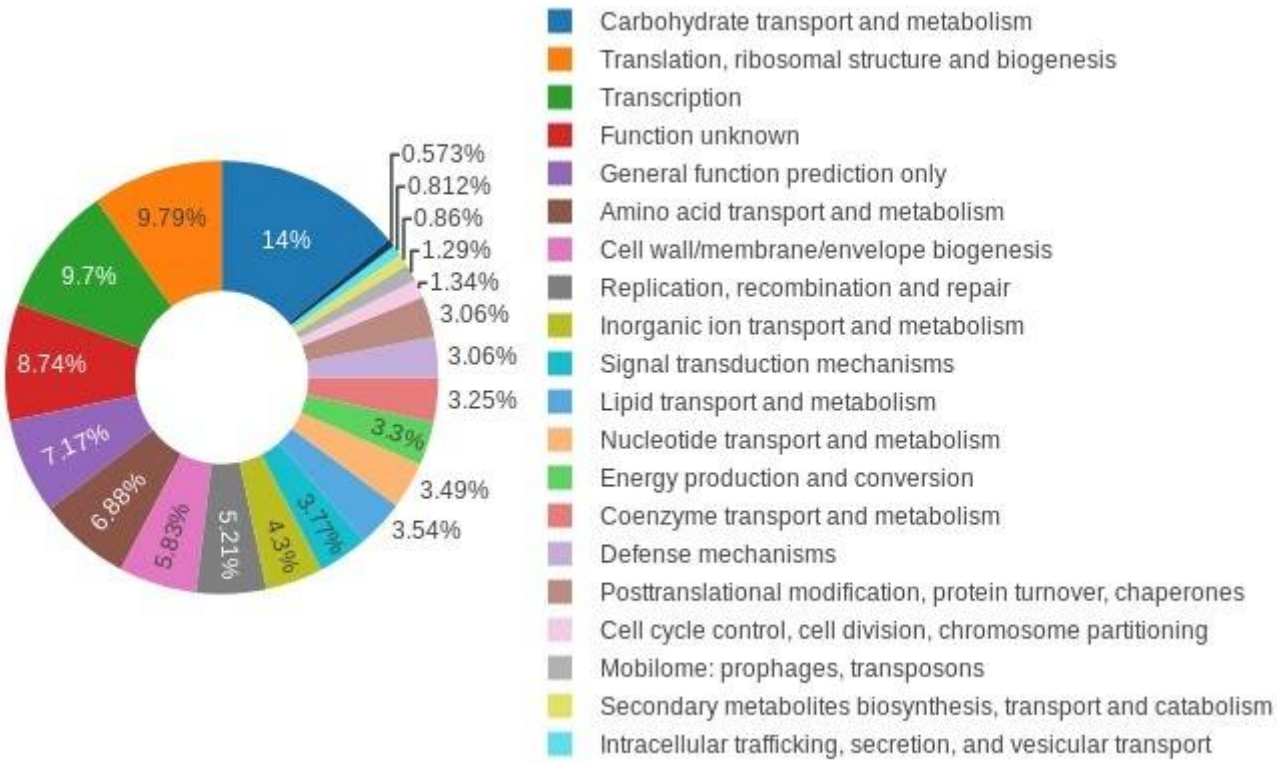

Figure 2

COG Functional classification of Enterococcus faecium R.A73 genes

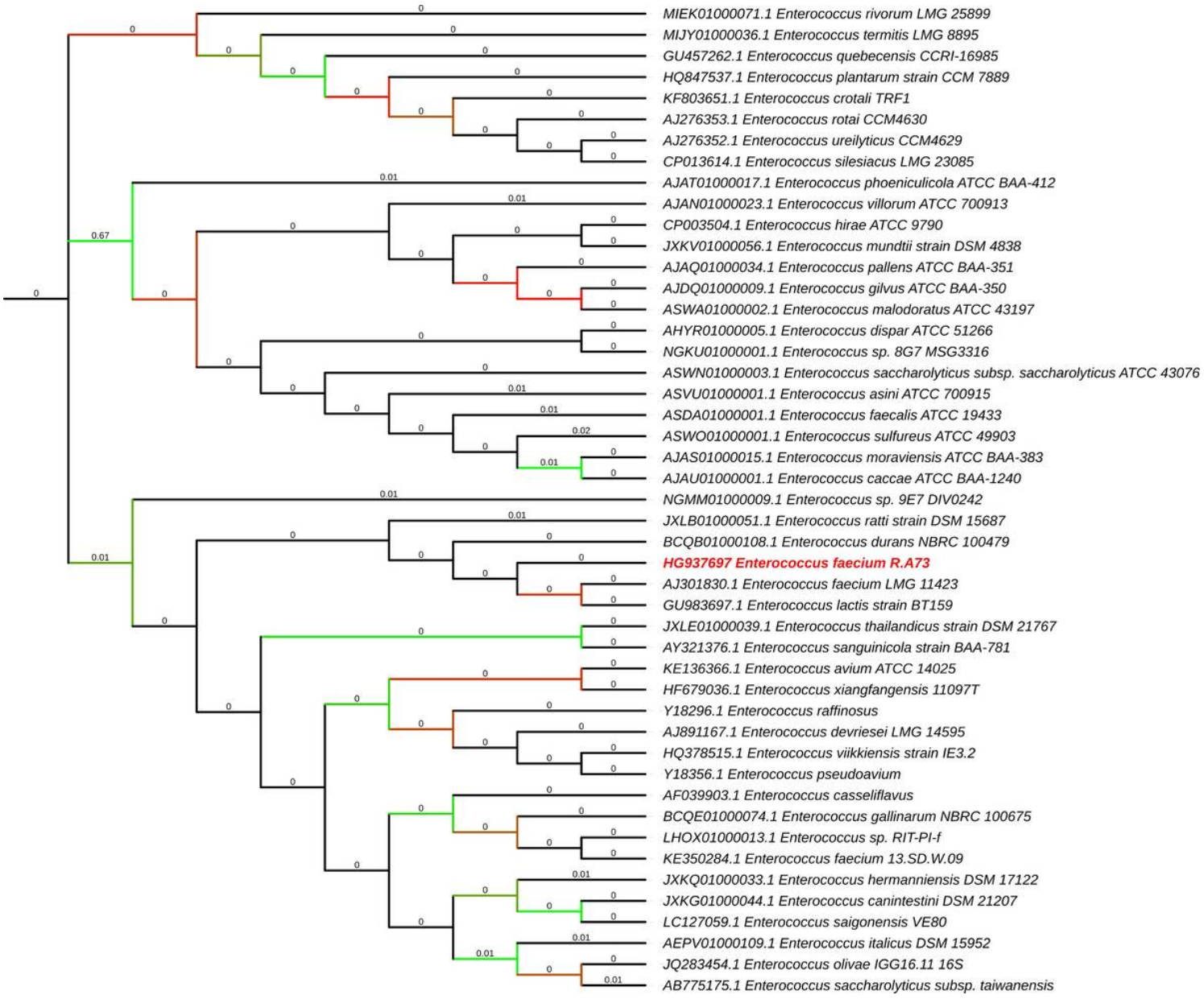

Figure 3

Phylogenetic tree based on 16S rDNA sequences

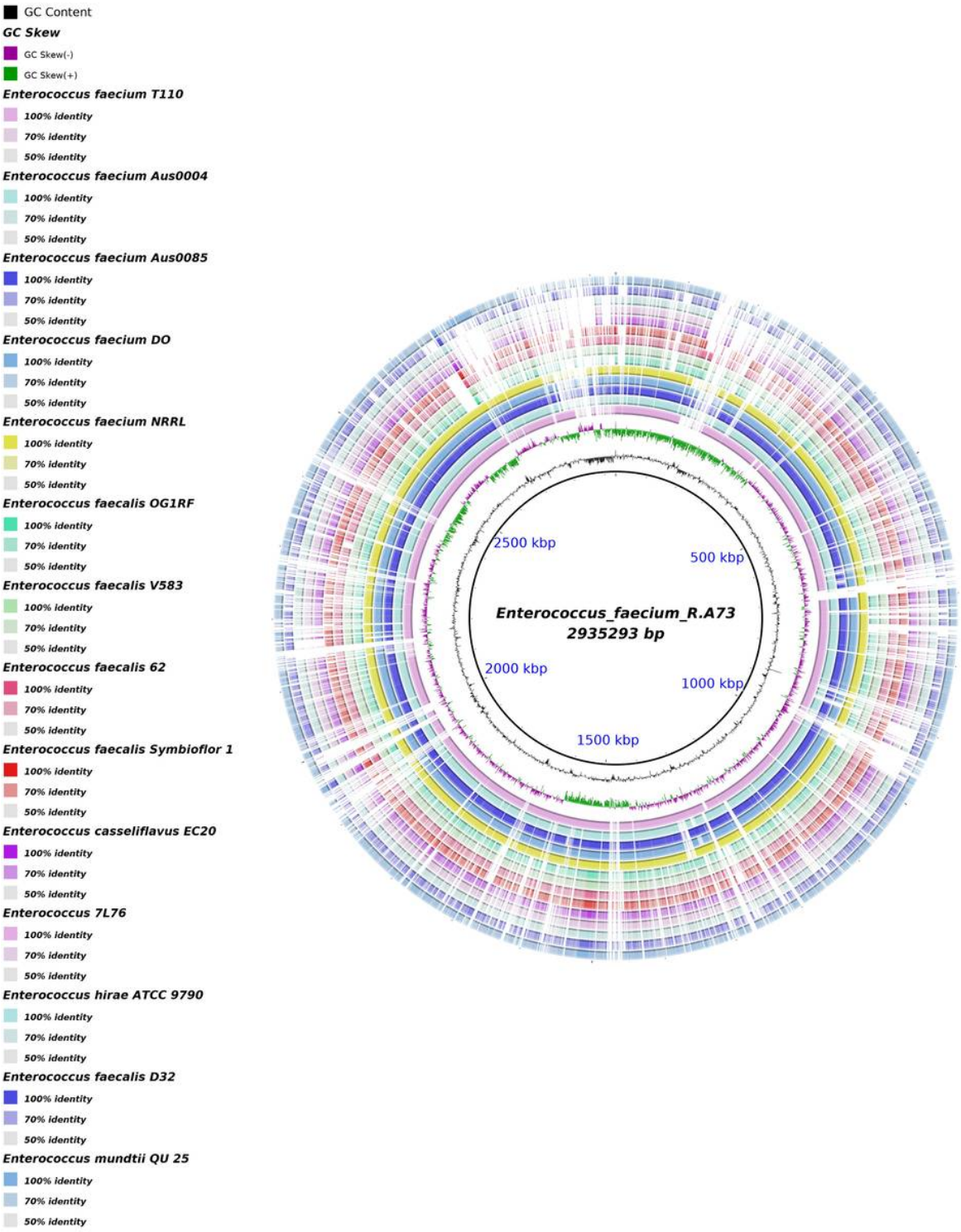

Figure 4

Comparative proteome analysis using the BRIG platform

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table3.tsv