

RESEARCH ARTICLE

Open Access



Comparative genome analysis reveals key genetic factors associated with probiotic property in *Enterococcus faecium* strains

Vikas C. Ghattargi^{1,2}, Meghana A. Gaikwad¹, Bharati S. Meti², Yogesh S. Nimonkar¹, Kunal Dixit¹, Om Prakash¹, Yogesh S. Shouche¹, Shrikant P. Pawar^{1*}  and Dhiraj P. Dhotre^{1*}

Abstract

Background: *Enterococcus faecium* though commensal in the human gut, few strains provide a beneficial effect to humans as probiotics while few are responsible for the nosocomial infection. Comparative genomics of *E. faecium* can decipher the genomic differences responsible for probiotic, pathogenic and non-pathogenic properties. In this study, we compared *E. faecium* strain 17OM39 with a marketed probiotic, non-pathogenic non-probiotic (NPNP) and pathogenic strains.

Results: *E. faecium* 17OM39 was found to be closely related with marketed probiotic strain T110 based on core genome analysis. Strain 17OM39 was devoid of known vancomycin, tetracycline resistance and functional virulence genes. Moreover, *E. faecium* 17OM39 genome was found to be more stable due to the absence of frequently found transposable elements. Genes imparting beneficial functional properties were observed to be present in marketed probiotic T110 and 17OM39 strains. Genes associated with colonization and survival within gastrointestinal tract was also detected across all the strains.

Conclusions: Beyond shared genetic features; this study particularly identified genes that are responsible for imparting probiotic, non-pathogenic and pathogenic features to the strains of *E. faecium*. Higher genomic stability, absence of known virulence factors and antibiotic resistance genes and close genomic relatedness with marketed probiotics makes *E. faecium* 17OM39 a potential probiotic candidate. The work presented here demonstrates that comparative genome analyses can be applied to large numbers of genomes, to find potential probiotic candidates.

Keywords: Non-pathogenic, Pathogenic, Indian, Comparative genome analysis, *In-silico* analysis

Background

Probiotic organisms according to World Health Organisation are 'Live microorganisms which when administered in adequate amounts confer a health benefit on the host'. Beneficial effects may include genes/pathways for production of vitamins, essential amino acids, antioxidants, digestion of complex carbohydrates, susceptibility to antibiotics, antagonism against enteric bacteria and modulation of immune system. But along with conferring health benefits the organism should have or lack series of properties. The probiotic strain should have

genes to compete, adhere, persist and survive in the harsh conditions of the gastrointestinal tract (GIT). Moreover, the probiotic strain should show absence of any virulence factors and multi drug resistance. Pathogenic bacteria may contain genes for survival in GIT along with some beneficial properties, but they hold virulence factors which help them to evade the host immune response and eventually cause disease. Additionally, presence of antibiotic resistance genes in pathogens makes the treatment difficult in disease conditions. Apart from these, many bacteria in the gut do not show either probiotic or pathogenic properties and hence can be termed as non-pathogenic non-probiotic (NPNP).

The genus *Enterococcus* is one of the diverse and ecologically significant group, and members of this genus

* Correspondence: shrikantpawar@nccs.res.in; dhiraj@nccs.res.in

¹National Centre for Microbial Resource (NCMR), National Centre for Cell Science (NCCS), Pune, Maharashtra 411021, India

Full list of author information is available at the end of the article



are ubiquitously distributed in nature viz. animals, human gastrointestinal tract (GIT) and plants [1–7]. *Enterococcus* plays an important role in the ripening of cheese products by lipolysis and proteolytic properties leading to the development of aroma and flavour [7]. In the Mediterranean region, *Enterococcus* spp. have been used in the preparation of various meat and fermented milk products for centuries [5]. Further, they also exhibit the beneficial property of bacteriocin production [5, 7] presenting activity against potential pathogens viz. group D streptococci and *Listeria* in various foods and GIT [7].

E. faecium is widely and extensively studied for its leading cause of nosocomial infections in humans [8]. It is a gut commensal and acts as an opportunistic pathogen due to a variety of virulence factors, including lipopolysaccharides and biofilm formation [9]. Their pathogenic nature is evident in urinary tract infections, endocarditis, and surgical wound infection, displaying its capability of causing a wide range of infections [10]. Another remarkable character of *E. faecium* is its tolerance to many antimicrobial drugs [11, 12]. It has also acquired the antibiotic-resistance gene against vancomycin and a multidrug resistance beta-lactamase gene [13]. Besides, it has been shown that *E. faecium* is capable of acquiring resistance to antibiotics by sporadic mutations and infections caused by these are normally difficult to treat [14]. The strains like Aus0004 and V583 are reported as pathogens [1].

Numerous studies in the last decade have validated the safety claim of *Enterococci* in foods and as probiotics [15–18]. The application of *Enterococci* as a starter culture e.g. *E. faecium* SF68 (Switzerland) and as probiotic e.g. *E. faecium* T110 (Japan) has been used widely [19, 20]. Additionally, *E. faecium* T110 is used in many commercially available probiotics, and no cause of illness or death has been reported [9]. *E. faecium* is among one of the directly fed microorganism recognized by the Association of American Feed Control, 2016. It is permitted as a probiotic supplement in the diet for poultry, dogs, piglets and mice [21–24]. Few strains of *E. faecium* (NRRL B-2354) act as surrogate microorganism used in place of pathogens for validation of thermal processing technologies [25] and some are widely used as laboratory strains, e.g. *E. faecium* 64/3 [26]. These two strains are non-pathogenic and are used routinely without any known disease outbreak [27]. Thus, the diversity and genomic plasticity of *E. faecium* are accountable for both probiotic and pathogenic nature [28–30]. In this study, we have carried out comparative genome analysis to identify genes/pathways which can help in distinguishing probiotic, pathogenic and NPNP strains of *E. faecium*. Further, we have also tried to describe the genetic differences between strain 17OM39 with marketed probiotic, non-pathogenic non-probiotic (NPNP), and pathogenic strains.

Results

Strain selection

Whole genome sequences were downloaded from NCBI genome database, and the strains were grouped into probiotic, non-pathogenic non-probiotic (NPNP) and pathogenic based on the literature survey (Table 1). The pathogenic group had six strains: DO, Aus0004, Aus0085, 6E6, E39 and ATCC 700221 [1–4]. The first four were isolated from the human blood and later two from human stool. The NPNP group had two strains: NRRL B-2354 and 64/3 [25, 26]. The probiotic group had the marketed strain T110 [9] and strain 17OM39 isolated from healthy human gut [31].

General genomic features

Genome sizes ranged from approximately 2.57–2.99 Mb with strain DO exhibiting the smallest and 6E6 the largest genome. Average G + C content varied between 37.25 to 38.55%. The genomic features of strains under the study are provided in Table 1. No significant differences (p -value ≤ 0.05 , Kruskal–Wallis test) could be noted between the groups with respect to their genome size, G + C content, average number of genes and coding DNA sequences (CDS).

The RAST annotation has facilitated to determine the features assigned to subsystems that are present in all organisms (Additional file 1: Figure S1). The average numbers of annotated protein-encoding genes were 2570, 2639 and 3093 for probiotic, NPNP and pathogenic groups, respectively. Annotation based on RAST for strains under the study suggested high abundance of subsystems related to carbohydrates and protein metabolism.

Comparisons of 17OM39 with other *E. faecium* strains

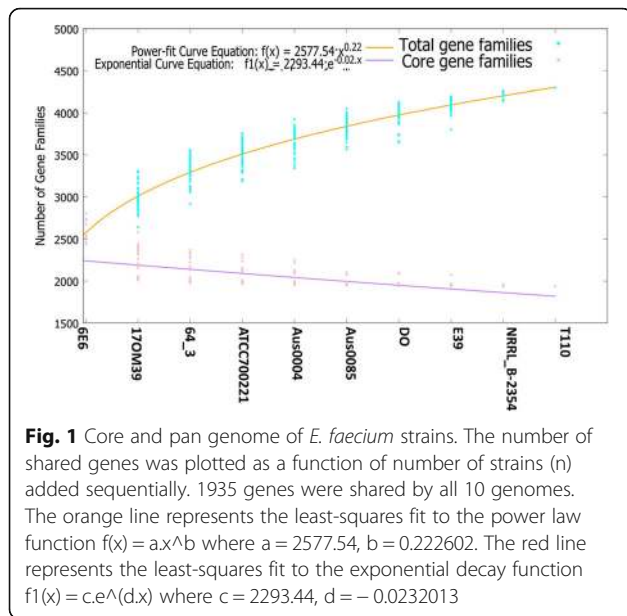
The availability of nearly complete *E. faecium* genomes has helped to define the core, accessory and unique genomic features for all the strains. The comparison of strain 17OM39 with other strains of probiotic, NPNP and pathogenic strains, revealed 1935 (85.53%) core genes, 526 (20.64%) accessory and 87(3.41%) unique genes. The numbers of shared genes were plotted as a function of number of strains (Fig. 1).

Pan-genome analysis

The pan-genome analysis revealed presence of 1935 core genes and 5718 accessory genes (Fig. 2). The numbers of strain-specific genes observed were 67, 87, 10, 64, 62, 16, 13, 36, 54 and 14 for strains 17OM39, T110, NRRL B-2354, 64/3, DO, AUS0004, AUS0085, 6E6, E39 and ATCC700221, respectively (Fig. 2). Identification of core, accessory and unique gene families by orthoMCL analysis revealed the proportion of known, hypothetical and uncharacterized proteins in these groups (Additional file 1: Figure S2A). Large percentages (61.19%) of unique

Table 1 General Genome Features

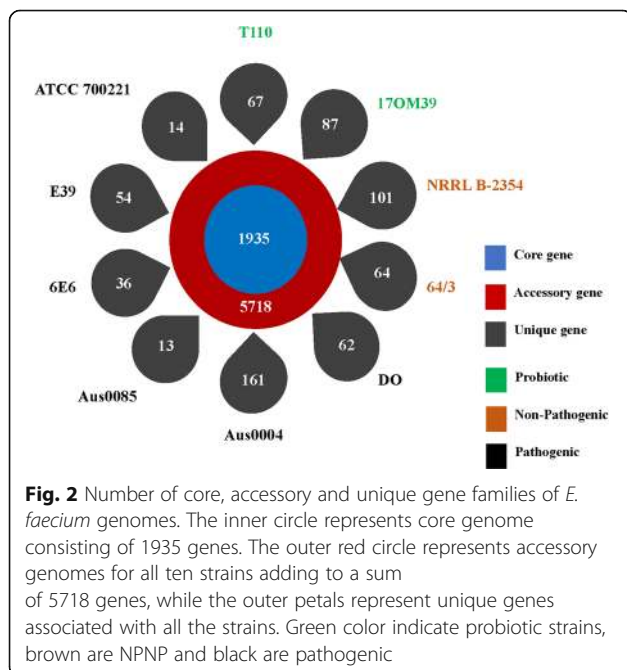
Features	Probiotic		Non-Pathogenic		Pathogenic					
	T110	170M39	NRRLB-2354	64/3	DO	Aus0004	Aus0085	6E6	E39	ATCC 700221
Size (mb)	2.6	2.6	2.6	2.5	2.6	2.9	2.9	2.9	2.7	2.8
GC%	38.4	38.5	37.8	38.2	37.9	38.3	37.9	37.6	37.8	37.8
Genes	2639	2865	2771	2508	2795	2960	3214	3404	3043	3145
CDS	2502	2639	2658	2418	2703	2825	2938	3307	2907	2725
Pseudo Genes	54	148	47	32	32	70	181	73	44	326
rRNAs (5S, 16S, 23S)	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6	6, 6, 6
tRNAs	65	62	48	68	62	47	76	75	70	72
Plasmid	1	-	1	-	3	3	6	2	5	3
Accession code	NZ_CP006030.1	LWHF_000000000.1	NC_020207.1	NZ_CP012522.1	NC_017960.1	NC_017022.1	NC_021994.1	NZ_CP013994.1	NZ_CP011281.1	CP_014449.1
Source	probiotic	feces	cheese	feces	blood	blood	blood	feces	blood	feces
Country	Japan	India	-	Germany	USA	Australia	Australia	USA	USA	USA
Reference	Natarajan and Parani 2015	Ghattargi et al. 2018	Kopit et al. 2014	Bender et al. 2015	Lam et al. 2012	Qin et al. 2012	Qin et al. 2012	Geldart and Kaznessis 2017	Geldart and Kaznessis 2017	McKenney et al. 2016



genes in all genomes were assigned to an uncharacterized group, and further studies will be required to examine the unexplored attributes.

Core genome analysis

orthoMCL analysis of core genes led to the identification of 850 genes present in single copy and 772 genes present in multiple copies in all ten strains. Functional analysis of core genes showed distribution in a varied range of functional categories within Cluster of Orthologous Genes (COG) viz. growth, DNA replication,



transcription, translation, carbohydrate and amino acid metabolism, stress response and transporters. Categories representing transport and metabolism of coenzymes, lipids, amino acids and nucleotides comprised of 16.24% of the core genes, while 11.30% of core genes were ascribed to carbohydrate metabolism.

Accessory genome analysis

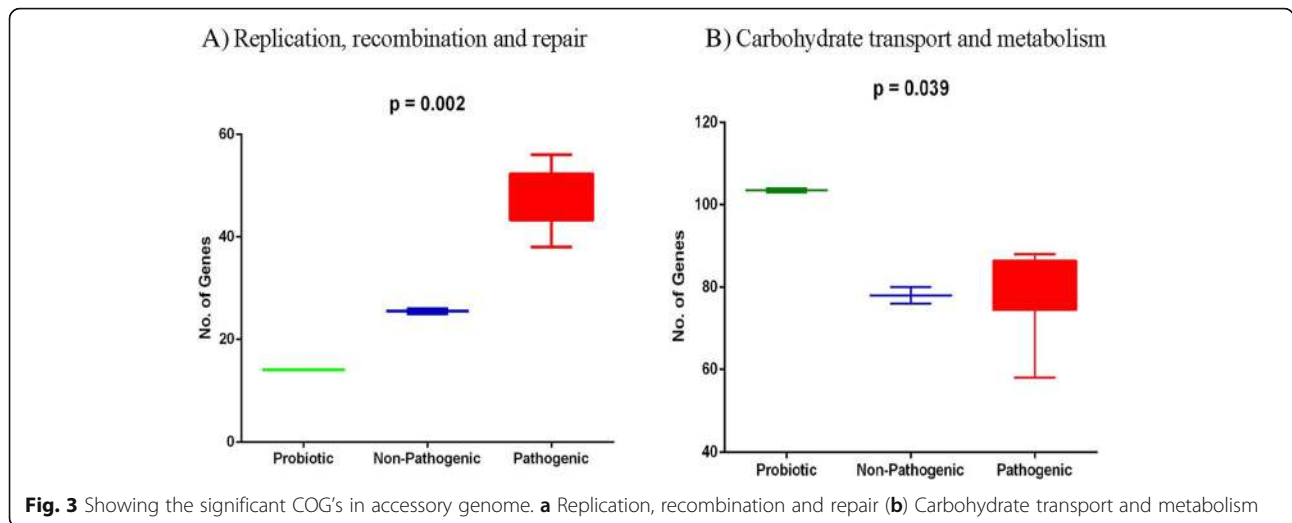
Functional analysis of the accessory genes showed diverse distribution in COG categories as similar to core gene annotations (Additional file 1: Figure S3). Two important subsystems observed in accessory genes were a) carbohydrate metabolism and b) replication, recombination and repair systems. The former was abundant in the probiotic group ($p = 0.002$) while later in the pathogenic group ($p = 0.039$) (Fig. 3). We also attempted to find accessory genes shared between the groups. The probiotic and pathogenic group shared 15 genes; four of them were general transporters, two were manganese-containing catalase gene while others were hypothetical proteins (Additional file 1: Figure S2B).

Unique genome analysis

The important unique genes associated with various strains were as follows: Phosphotransferase (PTS) system for mannose/fructose/sorbose was found in probiotic strain 17OM39. Marketed probiotic strain T110 had macrolide-efflux transmembrane protein which acts as a drug efflux pump. The important unique genes for other strains were hexosyltransferase in strain 64/3, type III restriction-modification system in strain NRRL B- 2354, Cro/CI family transcriptional regulator protein in strain ATCC 700221, streptogramin A acetyltransferase in strain Aus0004, Patatin-like proteins in strain Aus0085, IS1668 transposase in strain DO and plasmid recombination enzyme in strain E39. Additional file gives detailed information on core, accessory and unique genes (see Additional files 2, 3, 4).

Antibiotic resistance determinants

Screening of antibiotic resistance determinants in genomes was necessary to understand if probiotic strains harboured these genes. Genes conferring resistance to kanamycin were found in all the genomes. The NPNP group showed presence of general multidrug transporter. Within pathogenic strains, Aus004 and Aus0085 showed presence of tetracycline, trimethoprim and vancomycin resistance gene. Strains E39 and 6E6 showed presence of genes responsible for trimethoprim and tetracycline resistance. Pathogenic strain E39 presented daptomycin resistance gene and strain ATCC 700221 showed presence of genes responsible for resistance to the antimicrobial activity of cationic antimicrobial peptides and



antibiotics such as polymyxin. Table 2 shows presence of various antibiotic resistance genes found in each strain.

Virulence determinants

Various virulence determinants such as adherence, biofilm formation and exo-enzyme production in probiotic, NPNP and pathogenic groups were identified. Genes (*acm*, *scm*, *EbpA*, *EbpC*) described as adherence factors have been attributed to pathogenic effects. Excluding strain DO all other pathogenic strains showed presence of enterococcus surface protein (*esp*) gene. The *bopD* gene involved in biofilm was intact in all groups, but the operon was absent in strain 17OM39, marketed probiotic strain T110 and NPNP strains. In an exo-enzyme group, hyaluronidase gene was found to be associated with marketed probiotic strain alone, while the gene in strain 17OM39 displayed an alteration in sequence at position 167 (G > T) suggesting this could affect its functionality due to the nonsense mutation. Gene *acm* in the probiotic and the non-pathogenic group was not functional due to the non-sense mutation at position 1060 (G > T). Also, virulence genes *scm*, *efaA*

and *srtC* are not well characterized as virulence determinants in *E. faecium* [9] (Table 3).

Mobile genetic elements

A number of Mobile Genetic Elements (MGEs) have been described in *E. faecium* including transposons, plasmids, and bacteriophage [32]. Based on the screening for Insertion sequences (ISs) (Additional file 1: Table S1), the IS1542 was present only in probiotic strains. The IS element ISEfa12 was present only in NPNP group and IS1216, IS1216E, IS1216V, IS16, IS6770, ISEf1, ISEfa10, ISEfa11, ISEfa5, ISEfa7, ISEfa8, ISEnfa3, ISS1W were present only in the strains belonging to the pathogenic group. The detection of IS16 element was seen only in pathogenic group of *E. faecium* strains. A strong correlation between IS elements and virulence factors was observed in all genomes (Fig. 4) suggesting that these IS elements might have a role in transfer of these virulence factors. Further studies are required to verify these findings.

We could trace at-least one intact prophage in all ten genomes (Additional file 1: Table S2). In total, we could

Table 2 Antibiotic Resistance genes found in *Enterococcus* genomes as performed by CARD analysis, where + Present and - Absent

Property	Probiotic		Non-pathogenic Non-Probiotic		Pathogenic					
	T110	17OM39	NRRL B-2354	64_3	DO	Aus0004	Aus0085	6_E6	E39	ATCC 700221
Daptomycin	-	-	-	-	-	-	-	-	+	-
Trimethoprim	-	-	-	-	-	+	+	+	+	-
Multidrug	-	-	-	-	-	+	-	-	-	-
Macrolide	-	-	-	-	-	+	-	-	-	-
Polymyxin	-	-	-	-	-	-	-	-	-	+
Tetracycline	-	-	-	-	-	+	+	+	+	-
Vancomycin	-	-	-	-	-	+	+	-	-	-

Table 3 Virulence factors found in *Enterococcus* genomes, where + Present; – Absent; * Non-functional due to presence of stop codon

Property	GENES	Probiotic		Non-pathogenic Non-probiotic		Pathogenic					
		T110	17OM39	NRRL B-2354	64/3	DO	Aus004	Aus0085	6E6	E39	ATCC 700221
Adherence	<i>acm</i>	*	*	*	*	+	+	+	+	+	+
	<i>EbpA</i>	-	-	+	+	+	+	+	+	+	+
	<i>EbpC</i>	-	-	+	+	+	+	+	+	+	+
	<i>srtC</i>	+	+	+	+	+	+	+	+	+	+
	<i>EcbA</i>	-	-	-	-	+	+	+	+	+	+
	<i>efaA</i>	+	+	+	+	+	+	+	+	+	+
	<i>Esp</i>	-	-	-	-	-	+	+	+	+	+
	<i>Scm</i>	*	*	*	*	+	+	+	+	+	+
	<i>SgrA</i>	-	-	-	-	+	+	+	+	+	+
Biofilm	<i>bopD</i>	*	*	*	*	+	+	+	+	+	+
Exoenzymes	<i>EF0818</i>	+	*	-	-	-	-	-	-	-	-

identify three incomplete (PHASTER score < 70) and twenty-three intact phages (PHASTER score < 70–90). Nine phages whose completeness status was doubtful (PHASTER score < 90) were also identified. The NPNP strains had two intact prophages, while the probiotic strain 17OM39 had one intact and strain T110 had two intact prophages. *E. faecium* ATCC 700221 had the highest number of intact phages (Additional file 1: Table S2). Further, Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) associated (Cas) system were found to be absent within the all genomes.

Genomic islands of strain 17OM39 were compared with the other strain of *E. faecium* to find out genes transferred by Horizontal Gene Transfer (HGT)

(Additional file 1: Table S3). We identified a total of 11 genomic islands in strain 17OM39 amounting to 3.5% of the total genome. The genomic island GI1 was common across all groups except for the strain 6E6. The choloylglycine hydrolase gene was found to be present in the genomic island of probiotic strain T110, pathogenic strain 6E6 and NPNP strain 64/3. Pathogenic group showed a large number of IS elements, transposons and antibiotic resistance genes within genomic island. All pathogenic strains showed presence of tetracycline resistance gene and cell adhesion protein within the genomic island. Only two pathogenic strains Aus0004 and Aus0085 showed presence of *esp* (enterococcal surface protein) virulence gene, and vancomycin resistance genes.

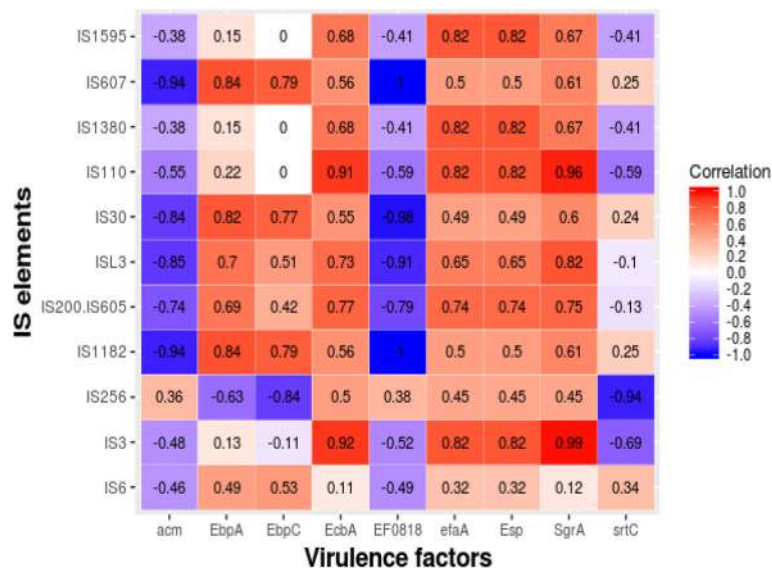


Fig. 4 Heatmap showing correlation between IS elements and virulence factors found across the genomes. Red color indicated strong positive correlation while blue indicated negative correlation

A higher similarity was observed between genomic islands of probiotic and NPNP strains as compared to pathogenic group (Fig. 5). Pathogenic strain AUS0004 had the highest mobile genome with almost 25% of its genome falling within the MGEs (Fig. 6). Additional file gives detailed information on genes present in genomic islands (see Additional file 5).

Survival in gastrointestinal tract

Biologically active microorganisms are usually required at the target site to induce a health benefits or pathogenic effect. To induce such effects it is necessary for the organism to survive and persist in the GIT. Thus, a list of genes encoding for survival and growth were first identified in the strain 17OM39 and then compared with other strains. We found Permease IIC component gene only in the probiotic group. All the groups showed presence of genes that impart resistance to acid, bile and could hydrolyse bile salt. Moreover, these strains were also able to adhere and grow in the GIT (Table 4).

Probiotic properties

As stated earlier the strain 17OM39 and marketed probiotic strains T110 were devoid of any clinically relevant antibiotic resistance gene while all the strains were able to survive in GIT. Strains 17OM39 and T110 (marketed probiotic) showed presence of complete pathways for essential amino acid synthesis viz. valine, lysine, and methionine and vitamins such as folate and thiamine (Table 5). Genes responsible for antibacterial activity (bacteriocin) specific against *Listeria* were found. Genes for exopolysaccharide (EPS) and anti-oxidant production (hydro-peroxidases) were noted which in-turn help the probiotic strains to establish themselves in the gut. The

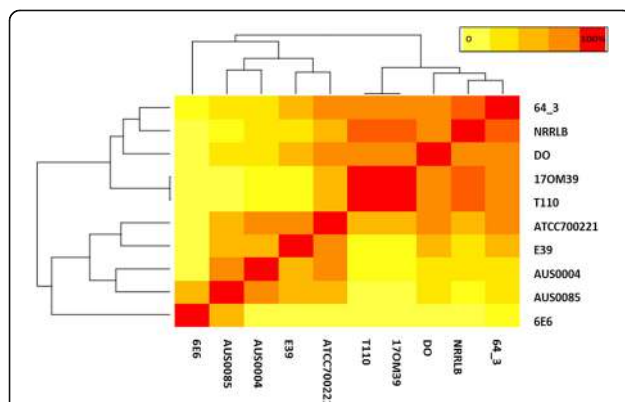


Fig. 5 Heatmap showing similarities between genomic islands of the strains considered in this study. The light-yellow shows least percent similarity while the red indicate 100% similarity in all genomes. Pathogenic strains (E39, Aus0004, Aus0085 and 6E6) showed clustering while probiotic strains (T110 & 17OM39), NPNP strains (64_3 & NRRLB) and pathogenic strain DO show separate cluster. The color scheme is as shown in the legend

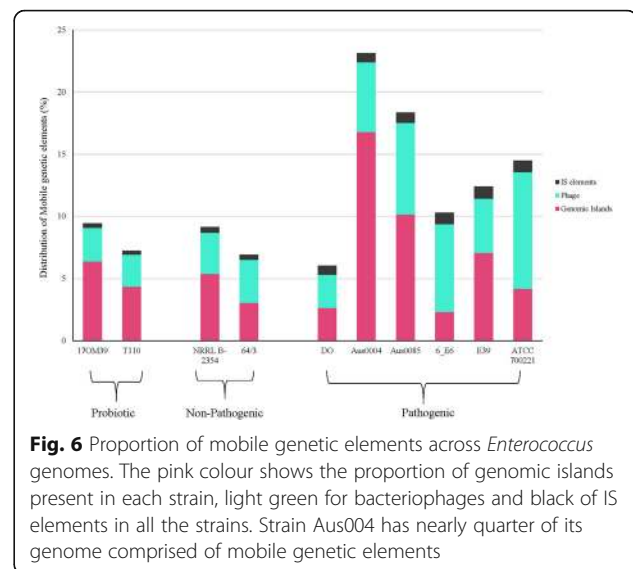


Fig. 6 Proportion of mobile genetic elements across *Enterococcus* genomes. The pink colour shows the proportion of genomic islands present in each strain, light green for bacteriophages and black of IS elements in all the strains. Strain Aus004 has nearly quarter of its genome comprised of mobile genetic elements

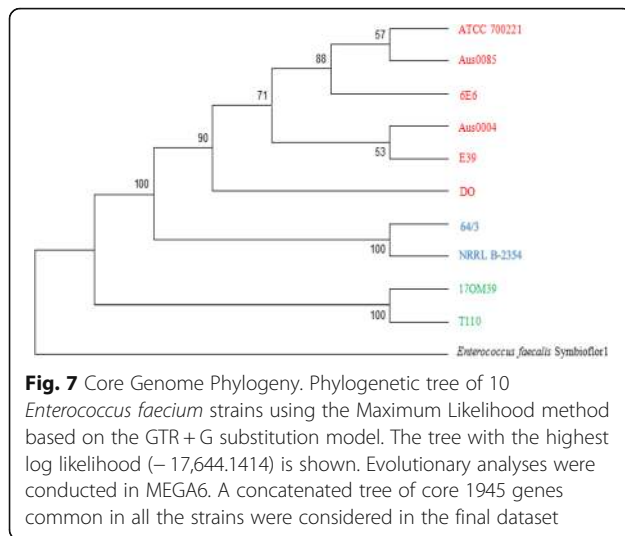
NPNP group only had EPS gene cluster. On the other hand, complete pathways for amino acid and vitamin synthesis were absent in NPNP and pathogenic group. Thus, probiotic strains have pathways/genes imparting beneficial effects to human host unlike NPNP and pathogenic group.

Plasmids

Plasmids comprise a substantial portion of the accessory genome and are accountable for antibiotic and virulence properties. Thus, an attempt was made to compare the plasmids of strains considered in this study with respect to their virulence factors, antibiotic resistance, phage regions and IS elements. Plasmid-encoded gene in marketed probiotic strain T110 showed 66% similarity to the cytolysin (*cyl*) gene. Strains in the pathogenic group (6E6, ATCC700221, Aus0085, and E39) showed presence of plasmid-encoded genes for vancomycin, streptothricin, erythromycin, gentamicin and kanamycin resistance (Additional file 1: Table S4). Surprisingly, no phage elements were associated with the probiotic strain T110, while the non-pathogenic strain (NRRL B-2354) and pathogenic strains (ATCC 700221, Aus0085, DO, and E39) harboured incomplete or complete prophages in the plasmids. The list of IS elements found in the plasmids is summarised in Additional file 1: Table S5.

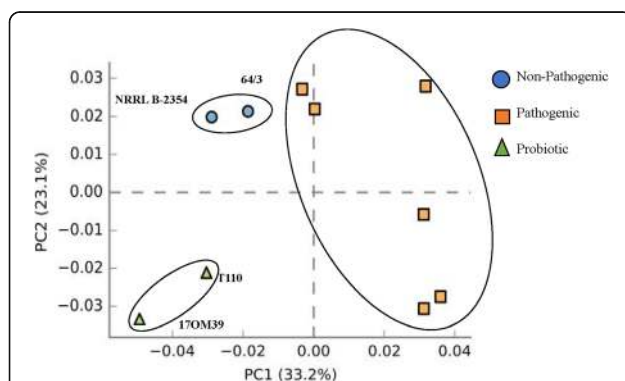
Comparison of probiotic, NPNP and pathogenic strains

The core genes (1935) were used to construct a phylogenetic tree of the 10 strains along with *Enterococcus faecalis* symbioflor as an out-group. Phylogenetic reconstruction by using Maximum Likelihood method separated 10 strains in 3 distinct clusters with high bootstrap support (bootstrap > 90) (Fig. 7). We found no clustering based on the source of isolation, while strain 17OM39



Discussion

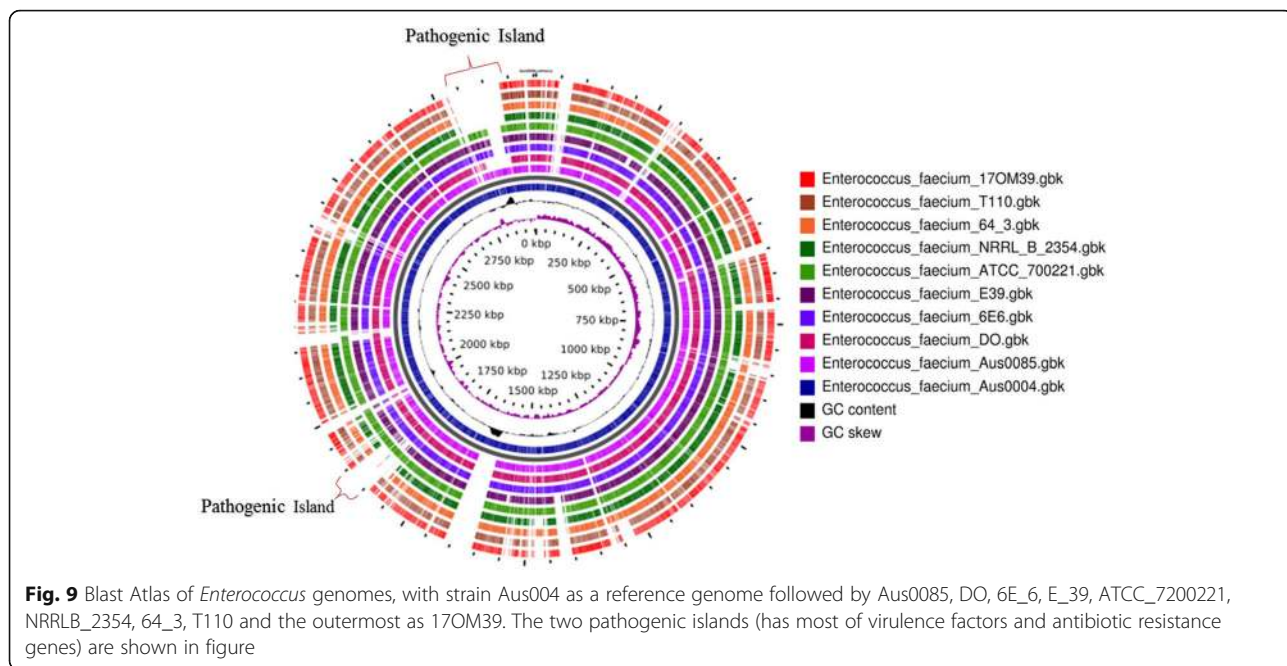
The genus *Enterococcus* is one of the diverse and ecologically significant group, and members of this genus are ubiquitously distributed in nature [1–7]. Numerous studies in the last decade have validated the safety claim of enterococci in foods and as probiotics [15–18]. The diversity and genomic plasticity of *E. faecium* is accountable for both probiotic and pathogenic nature [28–30]. In this study, we have carried out comparative genome analysis to identify genes/pathways which can help in distinguishing probiotic, pathogenic and NPNP strains of *E. faecium*. Further, we have also tried to describe the genetic differences between strain 17OM39 with marketed probiotic, non-pathogenic non-probiotic (NPNP), and pathogenic strains.



The study was carried out on 10 strains, and these strains were grouped into probiotic, non-pathogenic non-probiotic (NPNP) and pathogenic based on the literature survey (Table 1). We could not observe any correlation between higher G + C content with higher number of coding sequences as described in the earlier study [33]. On further annotation by RAST we identify enriched carbohydrates metabolism in all the strains and this is in agreement with the *E. faecium* ability to utilize a wide range of mono-, di-, oligo-saccharides [34, 35]. The comparative genome analysis revealed proportion of core (23.95%), accessory (70.80%) and unique (5.25%) genes. Also, the pan-genome size grew continuously with addition of strains indicating an open pan-genome while size of the core-genome gradually stabilized. These results are in accordance with the previous study for *E. faecium* genome [36]. The small size of core genome and huge number of accessory genes support the observation of the genomic fluidity in *E. faecium* [37].

Functional analysis of core genes have shown that 11.30% of genes were ascribed to carbohydrate metabolism which is in agreement with an earlier report [38] and the distribution of genes in categories of secondary metabolism and motility as contrast to earlier reports [38, 39]. Functional analysis of the accessory genes showed two important subsystems viz. a. carbohydrate metabolism and replication, and b. recombination and repair system. The carbohydrate metabolism was abundant in probiotic group while later in the pathogenic group (Fig. 3), this can be attributed to the properties of probiotic strains to utilize various carbohydrates [31], while the pathogenic group had higher abundance of replication and recombination genes known to be associated with a large number of mobile elements [1, 2]. Comparison of accessory genes between probiotic and pathogenic group helped in identifying two manganese-containing catalase gene, which provide resistance to hydrogen peroxide present in human GIT [40–42]. Among unique genes, Phosphotransferase (PTS) system for mannose/fructose/sorbose was present in probiotic strain 17OM39 which is involved in sugar uptake [43, 44] while marketed probiotic strain T110 had macrolide-efflux transmembrane protein responsible for drug efflux pump which plays a key role in drug resistance [45, 46]. Moreover, large percentages (61.19%) of unique genes were assigned to an uncharacterized group. Further studies will be required to examine the unexplored attributes.

Enterococci can exhibit resistance to a number of antibiotics, which have been attributed to their innate resistance and ability to successfully acquire resistance through horizontal gene transfer (HGT) [47, 48]. Multiple-drug-resistant strains of *E. faecium* have been increasingly associated with nosocomial infections particularly the vancomycin resistance [49]. Genes conferring



resistance to kanamycin were found across all the genomes as this has been attributed to intrinsic property within *E. faecium* [50]. In our study, genes imparting resistance to one or more antibiotics were seen in different strains of *E. faecium* (Table 2). Overall, the pathogenic group of *E. faecium* was found to have higher prevalence of antibiotic resistance genes; a factor that contributes to

the challenge of selecting therapeutic measures. The probiotic group was devoid of any major clinically relevant antibiotic resistance [9, 31].

Virulence genes contribute to the pathogenicity of an organism [51]. Despite the increasing knowledge of *E. faecium* as an opportunistic pathogen, the distribution of virulence factors is still poorly understood [51]. Knowledge

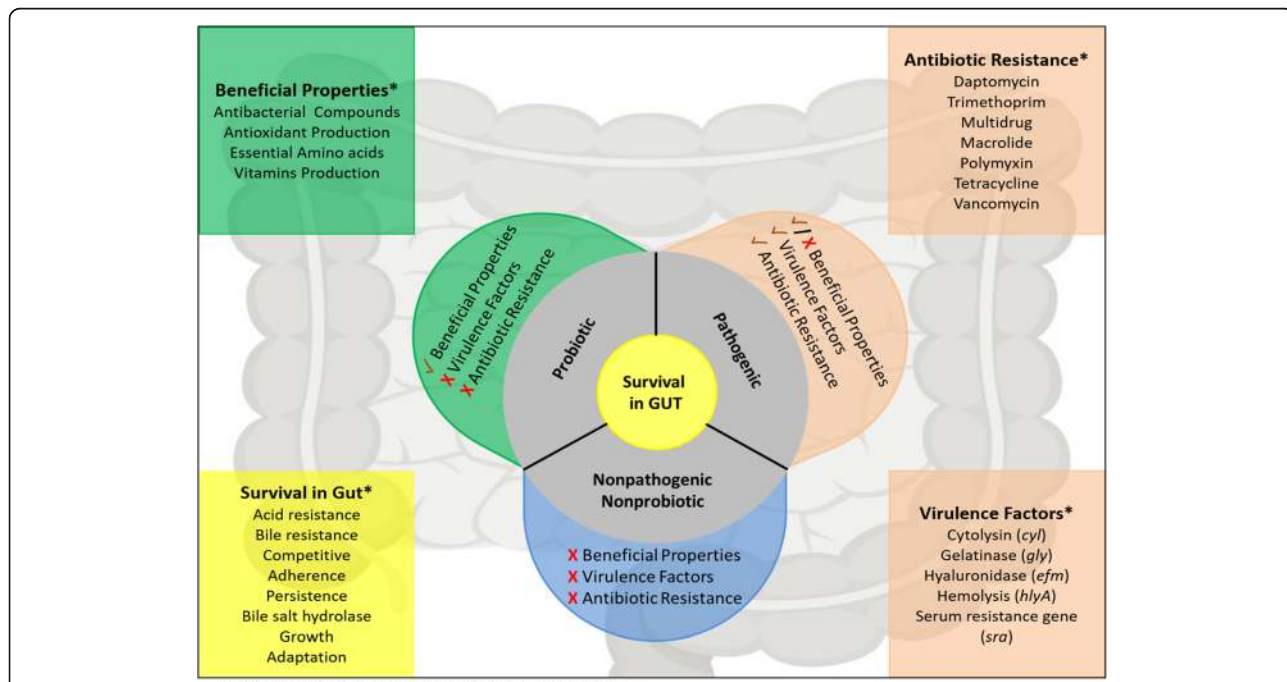


Fig. 10 Summarizes properties which can help in delineating probiotic, pathogenic and NPNP strains of *E. faecium*. ✓ indicates presence of a property, X indicates absence of property and ✓ / X indicates either presence or absence of a property

of the virulence characteristics helps to understand the complex pathogenic process of the pathogenic strains. The ability to adhere to the GIT is reflected to be one of the main selection criteria for potential probiotics as it extends their persistence in the intestine [52] and thus allows the bacterium to exert its probiotic effects for an extended time. However, adhesion is also considered a potential virulence factor for pathogenic bacteria [53]. The intestinal mucus is an important site for bacterial adhesion and colonization [54], and thus adherence property is beneficial to humans in case of probiotics, and it possesses adverse effects in pathogenic strains. The genes described as adherence factors (*acm*, *scm*, *EbpA*, *EbpC*) in *Enterococcus* have been attributed to pathogenic group. All pathogenic strains showed presence of enterococcus surface protein (*esp*) gene which contributes as a major virulence factor, except strain DO [55–58]. The operon for *bopD* gene involved in biofilm was intact in pathogenic group, but was absent in probiotic and NPNP groups [9]. Nonsense mutation was seen in hyaluronidase and *acm* gene in both probiotic and NPNP strains suggesting its non-functionality. Also, virulence genes *scm*, *efaA* and *srtC* are not well characterized as virulence determinants in *E. faecium* [9] (Table 3). Although, one could expect a virulence trait depending on the source of isolation, our study did not find any such traits and differed from the earlier reports [59]. Also, the strains showed significantly different patterns of virulence determinants, which underlines the findings of another author [59]. The strain 17OM39 within the probiotic group was devoid of any clinically relevant functional virulence determinants.

MGEs play an important role in the HGT [38, 60–62]. A number of MGEs have been described in *E. faecium* including transposons, plasmids, and bacteriophages [32]. ISs are possibly the smallest and most independent transposable elements which play an important role in shaping the bacterial genomes [63]. IS1542 was present only in probiotic group, and earlier studies on IS1542 have shown its presence in just 2 out of 65 human pathogenic strains, suggesting no direct relation with the strains pathogenicity [64]. The presence of the insertion sequence families in all groups imply that these elements may spread by HGT [36]. However, particular IS elements were distributed in only one group suggesting that these IS elements might have evolved over the time [36, 65]. Notably, the presence of IS16 has been used as a marker within the hospital strains of *E. faecium* with 98% sensitivity and 100% specificity [36]. This observation was further supported by the detection of IS16 in only pathogenic group of *E. faecium* strains. Moreover, ISEfa11 and ISEfa5 were found to be associated with vancomycin resistance genes viz. *VanS*, *VanX*, and *VanY* [36, 66]. A strong correlation between the IS elements and virulence factors was observed in all

genomes (Fig. 4) suggesting that these IS elements might transfer virulence factors. Further studies are required to verify these findings.

E. faecium are known to harbour bacteriophages, hence the presence of prophage was predicated in all the ten genomes [36, 67]. Bacteriophages contribute actively to bacterial evolution by integrating and excising from the genome [67]. In certain conditions, they provide new genetic properties to the bacterial host leading to the development of new pathogens, as shown for *Escherichia coli*, *Vibrio cholera* and *Corynebacterium diphtheriae* [68–70]. Various prophages were identified within the strains under study (Additional file 1: Table S2). We could not find any known functional virulence factors or genes associated with probiotic or pathogenic properties within these bacteriophage regions. Further, CRISPRs system was found to be absent within the genomes as oppose to its closest neighbour, *E. faecalis* [71].

Genomic islands are distinct DNA fragments differing between closely related strains, which usually are associated with mobility [72, 73]. The choloylglycine hydrolase gene was found to be present in the genomic island of probiotic strain T110, pathogenic strain 6E6 and NPNP strain 64/3. This gene imparts resistance to bile salts and thus help in survival within the gut environment [74, 75]. A higher similarity was observed between genomic islands of probiotic and NPNP strains as compared to pathogenic group (Fig. 5). The underlying mechanisms for probiotic, NPNP and pathogenic properties by *E. faecium* may be intrinsic or acquired by horizontal exchange of genetic material. Genes found within genomic islands can be considered as acquired properties [14, 72, 73, 76, 77] while others as intrinsic. As all the genes/pathways associated with probiotic properties were not found within MGEs for strain 17OM39, they can be considered as intrinsic. From this study, it is evident that MGEs play an important role in driving the evolution of *E. faecium* strains by adding new genetic features [74, 75]. Certainly, studies like one carried out here will be helpful to understand the evolution of predominant strains.

Biologically active microorganisms are usually required at the target site to induce health benefits or pathogenic effect. To induce such effects it is necessary for the organism to survive and persist in the GIT [78–81]. All the groups showed presence of genes that impart resistance to acid, bile, hydrolyse bile salt and were also able to adhere and grow in the GIT. This finding correlates with the fact that *E. faecium* are normal inhabitants of the gut [82] (Table 4). We found Permease IIC component gene accountable for catalysing the phosphorylation of incoming sugar substrates which helps in competence and survival [83] only in the probiotic group.

Along with the survival ability in GIT, a probiotic strain should be capable of producing antimicrobial substances but on the other hand, should be devoid of acquired antibiotic resistance [84–87]. Furthermore, they must give beneficial effects to the host by producing essential amino acids and vitamins. For strains 17OM39 and T110 (marketed probiotic), we could trace complete pathways for amino acid synthesis viz. valine, lysine, and methionine (Table 5). These are among the essential amino acids and need to be supplied exogenously to humans [88]. Vitamins such as folate and thiamine are the components of Vitamin-B and are considered as essential nutrients for humans. Folate (folic acid) cannot be synthesized by human cells and hence is necessary to be supplemented exogenously as it plays an important role in nucleic acid synthesis and amino acid metabolism [89–91]. Thus, strains (T110 and 17OM39) producing such amino acids and vitamins can be considered beneficial for humans [92, 93]. Antibacterial activity (bacteriocin gene) specific against *Listeria* were found were found in probiotic strains. Genes for exopolysaccharide (EPS) and anti-oxidant production (hydro-peroxidases) were also found in the probiotic strains, thus helping them to establish in the GIT. In summary, probiotic strains have pathways/genes imparting beneficial effects to human host unlike NPNP and pathogenic group.

Earlier studies within *E. faecium* isolates have shown the abundance of plasmids by finding 1–7 number of plasmids in 88 out of 93 isolates [94]. Plasmids comprise a substantial portion of the accessory genome and are accountable for antibiotic and virulence properties which are usually acquired by the HGT [94]. Plasmid from marketed probiotic strain T110 showed 66% similarity to the cytolysin (*cyl*) gene, an important determinant in lethality of endocarditis [94]. Various antibiotics resistance genes viz. vancomycin, streptothricin, erythromycin, gentamicin and kanamycin resistance were identified in plasmids of pathogenic group (Additional file 1: Table S4).

Multi-Locus Sequence Analysis (MLSA) based phylogeny using 6 housekeeping genes (*adk*, *atpA*, *gyd*, *gdh*, *ddl*, *purK* and *pstS*) could not distinguish between pathogenic and non-pathogenic strains of *E. faecium* [95], but this could be achieved by the core genome SNP based phylogeny [96–99]. Thus the phylogenetic reconstruction by using Maximum likelihood method on core genome separated 10 strains in 3 distinct clusters with high bootstrap support (bootstrap > 90) (Fig. 7). Additionally, the PCA plot based on euclidean distances showed a distinct grouping of strains based on probiotic, pathogenic and NPNP groups (Fig. 8). Pathogenic Island (2,812,458–2,878,042 and 1,860,143–1,894,650 bp) were identified in the genome atlas and mainly consist of virulence-associated genes, IS elements, transposons, integrases and antibiotic resistance-related genes. It also

has vancomycin resistance gene cluster and presence of *esp* gene which correlates with the previous studies [100, 101].

Conclusions

This study provides valuable insights based on the genomic differences of probiotic, NPNP and pathogenic strains of *E. faecium*. Analyses of core, accessory and unique genes present in the genomes have helped in differentiating strains with different properties. We observed a strong correlation between insertion sequence elements and virulence factors in pathogenic *E. faecium* strains which needs to be investigated further. Moreover, the analysis of intrinsic and acquired properties helped us to know the inherent probiotic properties of strain 17OM39. The work presented here demonstrates that comparative genomic analyses can be applied to large numbers of genomes, to find potential probiotic candidates.

Methods

Bacterial sequences and strains

Whole Genome Sequence of *E. faecium* was retrieved from NCBI genomes, and a total of ten strains were used in this study. The genome for 17OM39 was sequenced using the Illumina MiSeq platform using 2 × 300 paired-end libraries. *De-novo* assembly method was employed to carry out the assembly of quality-filtered reads using MIRA assembler version 4.9.3 [102]. All the genomes were RAST annotated [103].

Comparative analysis

Comparative analysis of ten whole genome sequences of *Enterococcus faecium* was done by an ultra-fast bacterial pan-genome analysis pipeline (BPGA) [104] which performs GC content analysis, pan-genome profile analysis along with sequence extraction and phylogenetic analysis. Furthermore, the genome was investigated for the presence of putative virulence genes using Virulence Factor of Bacterial Pathogens Database (VFDB) [105]. Screening of probiotic genes was done by performing a BLAST of probiotic genes to the genome by online NCBI's BLASTX tool [106]. Comprehensive Antibiotic resistance Database (CARD) was used for analysis of antibiotic resistance [49]. Presence of CRISPR repeats was predicted using the CRISPRFinder tools [107]. PHASTER: rapid identification and annotation of prophage sequences within bacterial genomes were used for identification of prophages within the genome [108]. Bacterial insertion elements (ISs) were identified by ISfinder [109]. Horizontal gene transfer was detected by genomic island tool: Islandviewer [110, 111]. The clustering and annotation of protein sequences were done with the help of orthoMCL [112]. COG analysis was done

with the help of webMGA server [113]. STAMP software was used to generate a PCA plot [114]. A blast atlas was generated with the help of GVIEW Server (<https://server.gview.ca/>) [115].

Additional files

Additional file 1: Figure S1 Features assigned to subsystems from RAST present in all ten *Enterococcus* strains. **Figure S2. (A)** Proportion of known, hypothetical and unknown proteins in the group of core, accessory and unique genes **(B)** Venn Diagram for accessory genome between probiotic, non-pathogenic and pathogenic group. **Figure S3.** Functional analysis of the accessory genes in COG categories. **Table S1.** IS elements found in *Enterococcus* genomes by ISfinder tool. + Present, – Absent. **Table S2.** Number of Phage elements present in *Enterococcus* genomes as intact, questionable and incomplete. **Table S3.** Number of Genomic Islands in *Enterococcus* genomes. **Table S4.** Antibiotic Resistance genes found in *Enterococcus* plasmids as performed by CARD analysis, where + Present and – Absent. **Table S5.** IS elements found in *Enterococcus* plasmids by ISfinder tool. + Present, – Absent. **Table S6.** Table showing the various genes used in the study for generating PCA plot. (DOCX 351 kb)

Additional file 2: Detailed information on the core genes present in the genomes under study (XLSX 57 kb)

Additional file 3: Detailed information on the accessory genes present in the strains under study. (XLSX 115 kb)

Additional file 4: Detailed information on the unique genes present in the strains under study. (XLSX 29 kb)

Additional file 5: Detailed information on genes present in genomic islands in all the strains. (XLSX 137 kb)

Abbreviations

BPGA: Ultra-fast bacterial pan-genome analysis pipeline; CARD: Comprehensive antibiotic resistance database; CDS: Coding DNA sequence; EPS: Exopolysaccharides; GI: Genomic Island; GIT: Gastrointestinal tract; HGT: Horizontal gene transfer; ISs: Insertion sequences; MGEs: Mobile genetic elements; PTS: Phosphotransferase system; VFDB: Virulence Factor of Bacterial Pathogens Database

Acknowledgements

Additional financial support was provided by the Department of Biotechnology, Basaveshwar Engineering College, Bagalkot, under TEQUIP Grant to Bharati S. Meti.

Funding

This work was supported by the Department of Biotechnology, Government of India, under Grant National Centre for Microbial Resource (NCMR) (BT/Coord.II/01/03/2016). The funding body has helped in providing the infrastructure and computing facilities and it did not involve in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The datasets generated and/or analysed during the current study are available at NCBI genome repository, [<https://www.ncbi.nlm.nih.gov/genome/browse>]. The accession numbers are *E. faecium* T110 (NZ_CP006030.); *E. faecium* 17OM39 (LWHF 00000000.1); *E. faecium* NRRLB-2354 (NC 020207.1); *E. faecium* 64/3 (NZCP012522.1); *E. faecium* DO (NC 017960.1); *E. faecium* Aus0004 (NC 017022.1); *E. faecium* Aus0085 (NC 021994.1); *E. faecium* 6E6 (NZ_CP013994.1); *E. faecium* E39 (NZ_CP011281.1) and *E. faecium* ATCC 700221(CP 014449.1).

Authors' contributions

VG, SP and DD conceived and designed the study; VG, MG, DD, YN, KD carried out data analyses; BM, OP, YS gave suggestions and intellectual inputs to the paper. VG, MG, SP and DD wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹National Centre for Microbial Resource (NCMR), National Centre for Cell Science (NCCS), Pune, Maharashtra 411021, India. ²Department of Biotechnology, Basaveshwar Engineering College, Bagalkot, Karnataka 587102, India.

Received: 19 April 2018 Accepted: 27 August 2018

Published online: 04 September 2018

References

- Lam MMC, Seemann T, Bulach DM, Gladman SL, Chen H, Haring V, et al. Comparative analysis of the first complete enterococcus faecium genome. *J Bacteriol.* 2012;194:2334–41.
- Qin X, Galloway-Peña JR, Sillanpaa J, Roh JH, Nallapareddy SR, Chowdhury S, et al. Complete genome sequence of enterococcus faecium strain TX16 and comparative genomic analysis of enterococcus faecium genomes. *BMC Microbiol.* 2012;12:135.
- Geldart K, Kaznessis YN. Characterization of class IIa Bacteriocin resistance in enterococcus faecium. *Antimicrob Agents Chemother.* 2017;61:1–17.
- McKenney PT, Ling L, Wang G, Mane S, Pamer EG. Complete genome sequence of enterococcus faecium ATCC 700221. *Genome Announc.* 2016; 4:e00386–16.
- dos Santos KMO, Vieira ADS, Salles HO, Oliveira JD, Rocha CRC, Borges MD, et al. Safety, beneficial and technological properties of *Enterococcus faecium* isolated from Brazilian cheeses. *Brazilian J Microbiol.* 2015;46: 237–49.
- Byappanahalli MN, Nevers MB, Korajkic A, Staley ZR, Harwood VJ. Enterococci in the environment. *Microbiol Mol Biol Rev.* 2012;76:685–706.
- Rasouli Pirouzian H, Hesari J, Farajnia S, Moghaddam M, Ghiassifar S. Effects of *Enterococcus faecalis* and *Enterococcus faecium*, isolated from traditional Lighvan cheese, on physicochemical and sensory characteristics of Iranian UF white cheese. *J Agric Sci Technol.* 2012;14:1023–34.
- Guggenbichler JP, Assadian O, Boeswald M, Kramer A. Incidence and clinical implication of nosocomial infections associated with implantable biomaterials - catheters, ventilator-associated pneumonia, urinary tract infections. *GMS Krankenhhyg Interdisziplin.* 2011;6:Doc18.
- 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.
- Kajihara T, Nakamura S, Iwanaga N, Oshima K, Takazono T, Miyazaki T, et al. Clinical characteristics and risk factors of enterococcal infections in Nagasaki, Japan: a retrospective study. *BMC Infect Dis.* 2015;15:426.
- Coque TM, Willems RJL, Fortún J, Diz S, Loza E, Cantón R. Population structure of enterococcus faecium causing bacteremia in a Spanish University Hospital : setting the scene for a future increase in Vancomycin Resistance ? Population structure of enterococcus faecium causing bacteremia in a Spanish University H. *Antimicrob Agents Chemother.* 2005; 49:2693–700.
- O'Driscoll T, Crank CW. Vancomycin-resistant enterococcal infections: epidemiology, clinical manifestations, and optimal management. *Infect Drug Resist.* 2015;8:217–30.
- Miller WR, Munita JM, Arias CA. Mechanisms of antibiotic resistance in enterococci. *Expert Rev Anti-Infect Ther.* 2014;12:1221–36.
- Hollenbeck BL, Rice LB. Intrinsic and acquired resistance mechanisms in enterococcus. *Virulence.* 2012;3:421–569.
- Huys G, Botteldoorn N, Delvigne F, De Vuyst L, Heyndrickx M, Pot B, et al. Microbial characterization of probiotics-advisory report of the working

- group "8651 probiotics" of the Belgian superior health council (SHC). *Mol Nutr Food Res*. 2013;57:1479–504.
16. Araújo TF, Ferreira CL de LF. The genus enterococcus as probiotic: safety concerns. *Brazilian Arch Biol Technol*. 2013;56:457–66.
 17. Giraffa G. Enterococci from foods. *FEMS Microbiol Rev*. 2002;26:163–71.
 18. Franz CMAP, Stiles ME, Schleifer KH, Holzapfel WH. Enterococci in foods - a conundrum for food safety. *Int J Food Microbiol*. 2003;88:105–22.
 19. Benyacoub J, Czarnecki-Maulden GL, Cavadini C, Sauthier T, Anderson RE, Schiffrin EJ, et al. Supplementation of food with enterococcus faecium (SF68) stimulates immune functions in young dogs. *J Nutr*. 2003;133:1158–62.
 20. Lee JH, Shin D, Lee B, Lee H, Lee I, Jeong DW. Genetic Diversity and Antibiotic Resistance of Enterococcus faecalis Isolates from Traditional Korean Fermented Soybean Foods. *J Microbiol Biotechnol*. 2017;27:916–24.
 21. Kačániová M, Kmet' V, Ččuboň J. Effect of enterococcus faecium on the digestive tract of poultry as a probiotic. *Turkish J Vet Anim Sci*. 2006;30:291–8.
 22. Kreuzer S, Machnowska P, Amus J, Sieber M, Pieper R, Schmidt MF, et al. Feeding of the probiotic bacterium enterococcus faecium NCIMB 10415 differentially affects shedding of enteric viruses in pigs. *Vet Res*. 2012;43:1–12.
 23. Vahjen W, Männer K. The effect of a probiotic enterococcus faecium product in diets of healthy dogs on bacteriological counts of Salmonella spp., campylobacter spp. and Clostridium spp. in faeces. *Arch Anim Nutr*. 2003;57:229–33.
 24. Benyacoub J, Perez PF, Rochat F, Saudan KY, Reuteler G, Antille N, et al. Enterococcus faecium SF68 enhances the immune response to Giardia intestinalis in mice. *J Nutr*. 2005;135:1171–6.
 25. Kopit LM, Kim EB, Siezen RJ, Harris LJ, Marco ML. Safety of the surrogate microorganism enterococcus faecium NRRL B-2354 for use in thermal process validation. *Appl Environ Microbiol*. 2014;80:1899–909.
 26. Bender JK, Fiedler S, Klare I, Werner G. Complete genome sequence of the gut commensal and laboratory strain enterococcus faecium 64/3. *Genome Announc*. 2015;3:53–60.
 27. VanRenterghem B. 126th AOAC Annual Meeting & Exposition. *Food Saf Mag*. 2012:38–45.
 28. Abejón MC, Medina RB, Katz MB, González SN. Technological properties of enterococcus faecium isolated from ewe's milk and cheese with importance for flavour development. *Can J Microbiol*. 2006;52:237–45.
 29. Hassanzadazar H, Ehsani A, Mardani K. Antibacterial activity of enterococcus faecium derived from Koopeh cheese against Listeria monocytogenes in probiotic ultra-filtrated cheese. *Vet Res Forum*. 2014;5:169–75.
 30. Satish Kumar R, Kanmani P, Yuvaraj N, Paari KA, Pattukumar V, Arul V. Purification and characterization of enterocin MC13 produced by a potential aquaculture probiotic *Enterococcus faecium* MC13 isolated from the gut of *Mugil cephalus*. *Can J Microbiol*. 2011;57:993–1001.
 31. Ghattargi VC, Nimonkar YS, Burse SA, Davray D, Kumbhare SV, Shetty SA, et al. Genomic and physiological analyses of an indigenous strain, enterococcus faecium 17OM39. *Funct Integr Genomics*. 2018;18:385–99.
 32. Mikalsen T, Pedersen T, Willems R, Coque T, Werner G, Sadowy E, van Schaik W, Jensen L, Sundsfjord A, Hegstad K. Investigating the mobilome in clinically important lineages of Enterococcus faecium and Enterococcus faecalis. 2015;16:282–98.
 33. Bonacina J, Suárez N, Hornigo R, Fadda S, Lechner M, Saavedra L. A genomic view of food-related and probiotic enterococcus strains. *DNA Res*. 2017;24:11–24.
 34. Devriese LA, Van De Kerckhove A, Kilpper-Balz R, Schleifer KH. Characterization and Identification of Enterococcus Species Isolated from the Intestines of Animals. *Int J Syst Bacteriol*. 1987;37:257–9.
 35. Manero A, Blanch AR. Identification of Enterococcus spp. with a Biochemical Key. *Appl Environ Microbiol*. 1999;65:4425–30.
 36. Mikalsen T, Pedersen T, Willems R, Coque TM, Werner G, Sadowy E, et al. Investigating the mobilome of clinically important lineages of enterococcus faecium and enterococcus faecalis. *BMC Genomics*. 2015;16:282.
 37. Bakshi U, Sarkar M, Paul S, Dutta C. Assessment of virulence potential of uncharacterized Enterococcus faecalis strains using pan genomic approach – Identification of pathogen-specific and habitat-specific genes. *Sci Rep*. 2016;6:38648.
 38. Beukers AG, Zaheer R, Goji N, Amoako KK, Chaves AV, Ward MP, et al. Comparative genomics of Enterococcus spp isolated from bovine feces. *BMC Microbiol*. 2017;17:52.
 39. Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, et al. Comparative genomics of Enterococci : comparative genomics of enterococci : variation in. *MBio*. 2012;3:1–11.
 40. Wang X, Tong H, Dong X. PerR-regulated manganese ion uptake contributes to oxidative stress defense in an oral streptococcus. *Appl Environ Microbiol*. 2014;80:2351–9.
 41. King KY, Horenstein JA, Caparon MG. Aerotolerance and peroxide resistance in peroxidase and PerR mutants of Streptococcus pyogenes. *J Bacteriol*. 2000;182:5290–9.
 42. Xu L, He SS, Li DY, Mei C, Hou XL, Jiang LS, et al. Hydrogen peroxide induces oxidative stress and the mitochondrial pathway of apoptosis in RAT intestinal epithelial cells (IEC-6). *Mol Biol*. 2016;50:270–7.
 43. Postma PW, Lengeler JW, Jacobson GR. Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev*. 1993;57:543–94.
 44. Monot M, Boursaux-Eude C, Thibonnier M, Vallenet D, Moszer I, Medigue C, et al. Reannotation of the genome sequence of Clostridium difficile strain 630. *J Med Microbiol*. 2011;60:1193–9.
 45. Sangvik M, Littauer P, Simonsen GS, Sundsfjord A, Dahl KH. Mef(a), mef(E) and a new mef allele in macrolide-resistant Streptococcus spp. isolates from Norway. *J Antimicrob Chemother*. 2005;56:841–6.
 46. Stadler C, Teuber M. The macrolide efflux genetic assembly of Streptococcus pneumoniae is present in erythromycin-resistant Streptococcus salivarius [5]. *Antimicrob Agents Chemother*. 2002;46:3690–1.
 47. Tong J, Lu XT, Zhang JY, Sui Q, Wang R, Chen M, et al. Occurrence of antibiotic resistance genes and mobile genetic elements in enterococci and genomic DNA during anaerobic digestion of pharmaceutical waste sludge with different pretreatments. *Bioresour Technol*. 2017;235:316–24.
 48. Hegstad K, Mikalsen T, Coque TM, Werner G, Sundsfjord A. Mobile genetic elements and their contribution to the emergence of antimicrobial resistant enterococcus faecalis and enterococcus faecium. *Clin Microbiol Infect*. 2010;16:541–54.
 49. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother*. 2013;57:3348–57.
 50. Galimand M, Schmitt E, Panvert M, Desmolaize B, Douthwaite S, Mechulam Y, et al. Intrinsic resistance to aminoglycosides in enterococcus faecium is conferred by the 16S rRNA m5C1404-specific methyltransferase EfmM. *RNA*. 2011;17:251–62.
 51. Comerlato CB, de Resende MCC, Caierão J, d'Azevedo PA. Presence of virulence factors in enterococcus faecalis and enterococcus faecium susceptible and resistant to vancomycin. *Mem Inst Oswaldo Cruz*. 2013;108:590–5.
 52. Ouwehand AC, Kirjavainen PV, Shortt C, Salminen S. Probiotics : mechanisms and established effects. *Int Dairy J*. 1999;9:43–52.
 53. Kirjavainen PV, Saxelin M, Rautelin H, Valtonen V, Salminen SJ, Ouwehand AC. Good adhesion properties of probiotics : a potential risk for bacteremia ? *FEMS Immunol Med Microbiol*. 2001;31:35–9.
 54. Finlay BB, Falkow S. Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev*. 1997;61:136–69.
 55. Heikens E, Bonten MJM, Willems RJL. Enterococcal surface protein esp is important for biofilm formation of enterococcus faecium E1162. *J Bacteriol*. 2007;189:8233–40.
 56. Ramadhan AA, Hegedus E. Biofilm formation and esp gene carriage in enterococci. *J Clin Pathol*. 2005;58:685–6.
 57. Toledo-Arana A, Valle J, Solano C, Arrizubieta MJ, Cucarella C, Lamata M, et al. The Enterococcal surface protein, Esp, is involved in *Enterococcus faecalis* biofilm formation. *Appl Environ Microbiol*. 2001;67:4538–45.
 58. Baldassarri L, Bertuccini L, Ammendolia MG, Gherardi G, Creti R. Variant esp gene in Enterococcus faecium. *Lancet*. 2001;357:1802.
 59. Dahlén G, Blomqvist S, Almståhl A, Carlén A. Virulence factors and antibiotic susceptibility in enterococci isolated from oral mucosal and deep infections. *J Oral Microbiol*. 2012;4:1–7.
 60. Jiang X, Ellabaan MMH, Charusanti P, Munck C, Blin K, Tong Y, et al. Dissemination of antibiotic resistance genes from antibiotic producers to pathogens. *Nat Commun*. 2017;8:15784.
 61. Kaplan T. The Role of Horizontal Gene Transfer in Antibiotic Resistance. *Eukaryon*. 2014;10:80–1.
 62. Von Wintersdorff CJH, Penders J, Van Niekerk JM, Mills ND, Majumder S, Van Alphen LB, et al. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol*. 2016;7:1–10.

63. Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev.* 2014;38:865–91.
64. Huh JY, Lee WG, Lee K, Shin WS. Distribution of insertion sequences associated with Tn 1546-like elements among enterococcus faecium isolates from patients in Korea distribution of insertion sequences associated with Tn 1546-like elements among enterococcus faecium isolates from Patie. *J Clin Microbiol.* 2004;42:1897–902.
65. Werner G, Fleige C, Geringer U, van Schaik W, Klare I, Witte W. IS element IS16 as a molecular screening tool to identify hospital-associated strains of enterococcus faecium. *BMC Infect Dis.* 2011;11:80.
66. López M, Sáenz Y, Álvarez-Martínez MJ, Marco F, Robredo B, Rojo-Bezares B, et al. Tn1546 structures and multilocus sequence typing of vanA-containing enterococci of animal, human and food origin. *J Antimicrob Chemother.* 2010;65:1570–5.
67. Matos RC, Lapaque N, Rigottier-Gois L, Debarbieux L, Meylheuc T, Gonzalez-Zorn B, et al. Enterococcus faecalis Prophage dynamics and contributions to pathogenic traits. *PLoS Genet.* 2013;9:1–16.
68. Davis BM, Waldor MK. Filamentous phages linked to virulence of *Vibrio cholerae*. *Curr Opin Microbiol.* 2003;6:35–42.
69. Backman J, Shackleton J, Kent V, Prell L, Hays JD, Denton N, et al. Shiga-like toxin-converting phages from *Escherichia coli*. Shiga-Like Toxin-Converting Phage from *Escherichia coli* Strains That Cause Hemorrhagic Colitis or Infant Diarrhea 1983;226:694–696.
70. FREEMAN VJ. Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *J Bacteriol.* 1951;61:675–88.
71. Palmer KL, Gilmore MS. Multidrug-resistant enterococci lack CRISPR-cas. *MBio.* 2010;1:1–10.
72. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol.* 2004;2:414–24.
73. Juhas M, Van Der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev.* 2009;33:376–93.
74. Jones BV, Begley M, Hill C, Gahan CGM, Marchesi JR. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc Natl Acad Sci U S A.* 2008;105:13580–5.
75. Begley M, Hill C, Gahan CGM. Bile salt hydrolase activity in probiotics bile salt hydrolase activity in probiotics. *Appl Environ Microbiol.* 2006; 72:1729–38.
76. Gold HS. Vancomycin-resistant enterococci: mechanisms and clinical observations. *Clin Infect Dis.* 2001;33:210–9.
77. Marothi YA, Agnihotri H, Dubey D. Enterococcal resistance—an overview. *Indian J Med Microbiol.* 2005;23:214–9.
78. Ispirli H, Demirbaş F, Dertli E. Characterization of functional properties of enterococcus faecium strains isolated from human gut. *Can J Microbiol.* 2015;61:1–10.
79. Banwo K, Sanni A, Tan H. Technological properties and probiotic potential of enterococcus faecium strains isolated from cow milk. *J Appl Microbiol.* 2013;114:229–41.
80. Ahmadova A, Todorov SD, Choiset Y, Rabesona H, Mirhadi Zadi T, Kuliyyev A, et al. Evaluation of antimicrobial activity, probiotic properties and safety of wild strain enterococcus faecium AQ71 isolated from Azerbaijani Motal cheese. *Food Control.* 2013;30:631–41.
81. Rao S, Ihara Y, Sukegawa S, Arakawa F, Fujimura T, Murakami H, et al. Characterization of probiotic properties of enterococcus faecium NHRD IHARA isolated from porcine feces. *Biosci Biotechnol Biochem.* 2013;77: 1967–9.
82. Campos E, Campos E, Baldoma L, Baldoma L, Aguilar J, Aguilar J, et al. Regulation of expression of the divergent. *Society.* 2004;186:1720–8.
83. Minelli EB, Benini A. Relationship between number of bacteria and their probiotic effects. *Microb Ecol Health Dis.* 2008;20:180–3.
84. Charteris WP, Kelly PM, Morelli L, Collins JK. Development and application of an in vitro methodology to determine the transit tolerance of potentially probiotic *Lactobacillus* and *Bifidobacterium* species in the upper human gastrointestinal tract. *J Appl Microbiol.* 1998;84:759–68. Fuller 1991
85. Rehaïem A, Belgacem ZB, Edalati MR, Martínez B, Rodríguez A, Manai M, et al. Assessment of potential probiotic properties and multiple bacteriocin encoding-genes of the technological performing strain enterococcus faecium MMRA. *Food Control.* 2014;37:343–50.
86. Fao J, Working WHO, Report G, Guidelines D, London F. Guidelines for the Evaluation of Probiotics in Food. 2002;1–11.
87. Ganguly NK, Bhattacharya SK, Sesikeran B, Nair GB, Ramakrishna BS, Sachdev HPS, et al. ICMR-DBT guidelines for evaluation of probiotics in food. *Indian J Med Res.* 2011;134:22–5.
88. Mahmood L. The metabolic processes of folic acid and vitamin B12 deficiency. *J Heal Res Rev.* United States: The National Academies Press; 2014;1:5.
89. Mahmood L. The metabolic processes of folic acid and vitamin B12 deficiency. *J Heal Res Rev.* 2014;1:5.
90. Ohrvik VE, Wittthoft CM. Human folate bioavailability. *Nutrients.* 2011; 3:475–90.
91. Tuszyńska M. Folic acid - the occurrence and the role in human nutrition. *Veg Crop Res Bull.* 2012;76(1):43–54.
92. Fijan S. Microorganisms with claimed probiotic properties: an overview of recent literature. *Int J Environ Res Public Heal.* 2014;11:4745–67.
93. Eck P, Friel J. Vitamins & Minerals Editorial. *Vitam Miner.* 2013;3:1–2.
94. Rosvoll TCS, Pedersen T, Sletvold H, Johnsen PJ, Sollid JE, Simonsen GS, et al. PCR-based plasmid typing in enterococcus faecium strains reveals widely distributed pRE25-, pRUM-, pIP501- and pHTbeta-related replicons associated with glycopeptide resistance and stabilizing toxin-antitoxin systems. *FEMS Immunol Med Microbiol.* 2010;58:254–68.
95. Homan WL, Tribe D, Poznanski S, Li M, Hogg G, Spalburg E, et al. Multilocus sequence typing scheme for enterococcus faecium multilocus sequence typing scheme for enterococcus faecium. *J Clin Microbiol.* 2002;40:1963–71.
96. Sankarasubramanian J, Vishnu US, Gunasekaran P, Rajendhran J. A genome-wide SNP-based phylogenetic analysis distinguishes different biovars of *Brucella suis*. *Infect Genet Evol.* 2016;41:213–7.
97. Heydari H, Wee WY, Lokanathan N, Hari R, Mohamed Yusoff A, Beh CY, et al. MabsBase: a *Mycobacterium abscessus* genome and annotation database. *PLoS One.* 2013;8:e62443.
98. Chen C, Zhang W, Zheng H, Lan R, Wang H, Du P, et al. Minimum core genome sequence typing of bacterial pathogens: a unified approach for clinical and public health microbiology. *J Clin Microbiol.* 2013;51:2582–91.
99. van Vliet AHM, Kusters JG. Use of alignment-free Phylogenetics for rapid genome sequence-based typing of helicobacter pylori virulence markers and antibiotic susceptibility. *J Clin Microbiol.* 2015;53:2877–88.
100. Shankar N, Baghdayan AS, Gilmore MS. Modulation of virulence within a pathogenicity island in vancomycin-resistant enterococcus faecalis. *Nature.* 2002;417:746–50.
101. McBride SM, Coburn PS, Baghdayan AS, Willems RJJ, Grande MJ, Shankar N, et al. Genetic variation and evolution of the pathogenicity island of enterococcus faecalis. *J Bacteriol.* 2009;191:3392–402.
102. Kisanand V, Lettieri T. Genome sequencing of bacteria: sequencing, de novo assembly and rapid analysis using open source tools. *BMC Genomics.* 2013;14:1.
103. 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.
104. Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep.* 2016;6:24373.
105. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* 2005;33(Database issue):D325–8.
106. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403–10.
107. Grissa I, Vergnaud G, Pourcel C. CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 2008;36(Web Server issue):52–7.
108. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 2016;44:W16–21.
109. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference Centre for bacterial insertion sequences. *Nucleic Acids Res.* 2006; 34(Database issue):D32–6.
110. Langille MGI, Brinkman FSL. IslandViewer: an integrated interface for computational identification and visualization of genomic islands. *Bioinformatics.* 2009;25:664–5.
111. Dhillion BK, Laird MR, Shay JA, Winsor GL, Lo R, Nizam F, et al. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Res.* 2015;43:W104–8.
112. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 2003;13:2178–89.

113. Wu S, Zhu Z, Fu L, Niu B, Li W, Yooseph S, et al. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*. 2011;12:444.
114. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics*. 2014;30:3123–4.
115. Petkau A, Stuart-Edwards M, Stothard P, Van Domselaar G. Interactive microbial genome visualization with GView. *Bioinformatics*. 2010;26:3125–6.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

