MICROBIAL GENOMICS

RESEARCH ARTICLE

Freitas *et al.*, *Microbial Genomics* 2020;6

DOI 10.1099/mgen.0.000350

Comparative genomics of global *optrA*-carrying *Enterococcus faecalis* uncovers a common chromosomal hotspot for *optrA* acquisition within a diversity of core and accessory genomes

Ana R. Freitas[1,*], Ana P. Tedim[2], Carla Novais[1], Val F. Lanza[3] and Luísa Peixe[1,*]

Abstract

Linezolid-resistant *Enterococcus faecalis* (LREfs) carrying *optrA* are increasingly reported globally from multiple sources, but we lack a comprehensive analysis of human and animal *optrA*-LREfs strains. To assess if *optrA* is dispersed in isolates with varied genetic backgrounds or with common genetic features, we investigated the phylogenetic structure, genetic content [antimicrobial resistance (AMR), virulence, prophages, plasmidome] and *optrA*-containing platforms of 27 publicly available *optrA*-positive *E. faecalis* genomes from different hosts in seven countries. At the genome-level analysis, an in-house database with 64 virulence genes was tested for the first time. Our analysis showed a diversity of clones and adaptive gene sequences related to a wide range of genera from *Firmicutes*. Phylogenies of core and accessory genomes were not congruent, and at least PAI-associated and prophage genes contribute to such differences. Epidemiologically unrelated clones (ST21, ST476-like and ST489) obtained from human clinical and animal hosts in different continents over eight years (2010–2017) could be phylogenetically related (3–126 SNPs difference). *optrA* was located on the chromosome within a Tn*6674*-like element (*n*=10) or on medium-size plasmids (30–60 kb; *n*=14) belonging to main plasmid families (RepA_N/Inc18/Rep_3). In most cases, the immediate gene vicinity of *optrA* was generally identical in chromosomal (Tn*6674*) or plasmid (*impB-fexA-optrA*) backbones. Tn*6674* was always inserted into the same Δ*radC* integration site and embedded in a 32 kb chromosomal platform common to strains from different origins (patients, healthy humans, and animals) in Europe, Africa, and Asia during 2012–2017. This platform is conserved among hundreds of *E. faecalis* genomes and proposed as a chromosomal hotspot for *optrA* integration. The finding of *optrA* in strains sharing common adaptive features and genetic backgrounds across different hosts and countries suggests the occurrence of common and independent genetic events occurring in distant regions and might explain the easy *de novo* generation of *optrA*-positive strains. It also anticipates a dramatic increase of *optrA* carriage and spread with a serious impact on the efficacy of linezolid for the treatment of Gram-positive infections.

DATA SUMMARY

The whole genome of all *Enterococcus faecalis* strains was downloaded from NCBI (https://www.ncbi.nlm.nih.gov/Traces/wgs/) and EMBL (https://www.ebi.ac.uk/ena). A list of the accession numbers, related epidemiological data, and references (if available) can be found in Table 1. The authors confirm all supporting data, code and protocols have been provided within the article or through supplementary data files. Seven supplementary tables and four supplementary figures are available with the online version of this article.

INTRODUCTION

Oxazolidinones (linezolid and tedizolid) are increasingly used for the treatment of human infections caused by

Received 04 October 2019; Accepted 15 February 2020; Published 09 March 2020

Author affiliations: [1]UCIBIO/REQUIMTE. Departamento de Ciências Biológicas. Laboratório de Microbiologia. Faculdade de Farmácia. Universidade do Porto. Porto, Porto, Portugal; [2]Grupo de Investigación Biomédica en Sepsis - BioSepsis. Hospital Universitario Río Hortega, Instituto de Investigación Biomédica de Salamanca (IBSAL), Valladolid, Spain; [3]Departamento de Bioinformática. Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid, Spain.

***Correspondence:** Ana R. Freitas, anarpf@gmail.com; Luísa Peixe, lpeixe@ff.up.pt

Keywords: enterococci; linezolid resistance; *optrA*; Tn*6674*; phylogeny; accessory genome.

Abbreviations: ABR, antibiotic resistance; AMR, antimicrobial resistance; CGE, center for genomic epidemiology; COGs, Clusters of Orthologous Groups of proteins; ICEs, integrative conjugative elements; LRE, linezolid resistant enterococci; MGE, mobile genetic element; PHASTER, PHAge Search Tool Enhanced Release; RCR, rolling-circle replicating; SNP, single nucleotide polymorphism; ST, sequence type.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Seven supplementary tables and four supplementary figures are available with the online version of this article.

relevant pathogens such as methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant enterococci [1, 2]. Linezolid resistance rates remain generally low in enterococci causing infections worldwide (<1%) [1, 3], however, acquired linezolid resistance genes (*cfr*, *optrA* and/or *poxtA*) are being increasingly reported in different enterococcal species and across different settings [2–7]. Among the three transferable linezolid resistance genes, *optrA* has been the main one responsible for the recent increase in linezolid-resistant enterococci (LRE) in human isolates [8–13]. According to available studies, *optrA*-carrying LRE are globally circulating in hospitals since at least 2005 [14], while the first description in food-producing animals dates from 2008 [15].

optrA codifies an ABC-F protein targeting the ribosome of Gram-positive bacteria and mediates resistance to both phenicols and oxazolidinones through ribosomal protection [16]. As phenicols are allowed in veterinary medicine to control respiratory infections, the use of florfenicol in food-producing animals might co-select resistance for oxazolidinones and promote the occurrence of an *optrA* animal reservoir [7]. *optrA*-carrying LRE strains were already described in a variety of hosts including hospitalized humans, healthy humans and different food-producing animals (poultry, pigs, cattle, and ducks) worldwide, and embedded in a high diversity of genetic backgrounds in association with different chromosomal or plasmid genetic platforms [1–3, 14, 15, 17–19]. Moreover, *optrA* has been more commonly described in *Enterococcus faecalis* than in *Enterococcus faecium* [2, 20].

Even though *optrA*-carrying LRE strains pose a growing threat to public health, little has been done regarding a global comparative genomics analysis and detailed *optrA* genetic context characterization [5, 9, 21, 22]. Indeed, most *optrA*-positive strains are only analysed with respect to their *optrA*-carrying platform neglecting the deep characterization of their genetic makeup. Although no clear distinction between clinical and non-clinical *E. faecalis* isolates exists in terms of gene repertoire, we aimed to address if *optrA*-carrying strains from different hosts carry similar genetic markers by performing a comprehensive comparison of publicly available *optrA*-carrying *E. faecalis* genomes. Given the exponential increase of *optrA* among *E. faecalis* from different settings and areas, we also aimed to investigate deeper about *optrA* genetic platforms and their mobilizable nature.

Impact Statement

Enterococcus rank as one of the most common causes of nosocomial infections and they are well known for their ability to acquire and transfer different gene adaptive traits including antimicrobial resistance and virulence genes, their genomic plasticity and high numbers of extra-chromosomal elements. In recent years, transferable genes conferring resistance to last-resort antibiotics as linezolid, especially *optrA*, emerged in different regions and, in particular, in *E. faecalis* strains obtained from different hosts. Nevertheless, little is known about the similarity of populations and adaptive genes circulating globally. Here, we compared the genomes of *optrA*-carrying *E. faecalis* strains from disparate origins and regions and identified some epidemiologically unrelated strains from human clinical and animal hosts in different continents as being phylogenetically related. *optrA* was commonly associated with a Tn*6674*-like element (besides diverse plasmids) inserted into the same chromosomal integration site in different strains, here proposed as a genetic hotspot for *optrA* integration. Furthermore, *optrA* was found to be embedded in a common 32 kb chromosomal platform that is conserved among hundreds of *E. faecalis* genomes. The data provide novel insights about *optrA* mobilization across different settings that might help to explain the easy *de novo* generation of *optrA*-positive strains. It also highlights the relevance of further genomic studies exploring the dissemination and reservoirs of *optrA* at a global and One Health scales.

METHODS

Bacterial collection and comparative genomics

Publicly available *optrA*-carrying *E. faecalis* genomes (*n*=27) at the NCBI and EMBL databases (last search March 2019) were retrieved together with all available epidemiological information. An additional genome not carrying *optrA* was also considered since it is included in the same research project ("*Gastrointestinal health of animal handlers with special emphasis on potential zoonotic transmission of bacterial infection*") of other related *optrA*-carrying *E. faecalis* isolated from pigs [23], accounting for a total of 28 genomes. Only complete, assembled and identifiable genomes were included in this analysis (Table 1).

All 28 genomes were mapped against the *E. faecalis* V583 reference strain to infer a phylogeny based on the concatenated alignment of high-quality single nucleotide polymorphisms (SNPs) using CSI Phylogeny standard settings (CGE, https://cge.cbs.dtu.dk/services/CSIPhylogeny). Recombination was removed from the lineage-specific analyses using Gubbins [24]. Accessory genome was determined using PATO (https://github.com/irycisBioinfo/PATO), genes were considered homologous if they had 80% identity and 80% coverage. Both trees were obtained with IQ-tree [25]; in the case of the accessory genome tree a binary model was made by using a presence/absence matrix. Comparison between SNPs-based and accessory genome trees was performed using R package phytools [26].

Genomes were screened for genes encoding antibiotic resistance (ABR), plasmids, virulence and MLST using *in silico* genomic tools (ResFinder 3.0, PlasmidFinder 2.1, VirulenceFinder 1.5 and MLST 1.8 tools, respectively) available at the Center for Genomic Epidemiology (CGE; http://www.genomicepidemiology.org). As the VirulenceFinder

Table 1. Epidemiological background of the 28 *E. faecalis* genomes analysed in this study

Isolate	Source	Sample	Country	Year	ST	No Contigs	Genome size (bp)	GenBank/ENA	Ref
12E	RCk	Meat	COL	2010–2011	59	234	2834796	ERX1670553	[33]
34E	RCk	Meat	COL	2010–2011	59	220	2776507	ERX1670554	[33]
745E	RCk	Meat	COL	2010–2011	489	277	2979622	ERX1670555	[33]
612T	RCk	Meat	TUN	2017	21	55	2944588	SEWV01000000	[64]
728T	RCk	Meat	TUN	2017	859	62	3278897	SEWT01000000	[64]
697T	RCk	Meat	TUN	2017	476	28	2973239	VDMY00000000	[64]
N60443F	RCt	Meat	USA	2014	489	3	2853753	CP028724-6	[5]
N48037F	RP	Meat	USA	2013	141	4	3151142	CP028720-3	[5]
712T	Chicken	Faeces	TUN	2017	476	55	2982138	SEWU01000000	[64]
P.En090	Pig	Faeces	MYS	2012	756	60	3071086	MJBX01000000	[23]
Enfs51	Pig	Faeces	MYS	2012	754	58	3024957	MJEK01000000	[23]
P.En250	Pig	Faeces	MYS	2012	476	42	2834146	MJBZ01000000	[23]
Enfs85	Pig	Faeces	MYS	2012	755	62	2834811	MJBV01000000	[23]
IBUN9046YE	Pig	Faeces	COL	2014	16	113	2945287	NCAD01000000	–
12T	WW	ww	TUN	2014	86	27	3009161	NHNF01000000	[37]
P.En218	H (PF)	Faecal	MYS	2012	765	61	2862267	MJBY01000000	[23]
Enfs94*	H (PF)	Faecal	MYS	2012	227	53	2625583	MJBW01000000	[23]
Efs 599	HH	Unknown	USA	Unknown	16	130	3053367	ALZI01000000	–
11340	HP	Ascites	CHN	2013	476	119	2920750	QNGJ01000000	[30]
13484†	HP	Fester	CHN	2013	21	122	2918217	QNGN01000000	[30]
18026	HP	Ascites	CHN	2014	476	231	2959223	QNGT01000000	[30]
27149†	HP	Ascites	CHN	2014	858	97	2836434	QNHA01000000	[30]
Iso1Bar	HP	Urine	ESP	2016	585	56	2878959	ERX2067873	[12]
Iso2Bar	HP	Urine	ESP	2016	585	72	2902425	ERX2067874	[12]
Iso3Bar	HP	Urine	ESP	2016	585	79	2872904	ERX2067875	[12]
Iso4Bar	HP	Urine	ESP	2016	474	128	2943732	ERX2067876	[12]
Iso5Bar	HP	Urine	ESP	2016	474	204	2949268	ERX2067877	[12]
16–372	HP	Urine	FRA	2016	480	6	2779791	PHLF01000000	[10]

*Enfs94 strain does not carry *optrA*.
†These isolates were obtained from hospitalized patients that work as farmers.
‡These isolates were obtained from patients that work as farmers.
CHN, China; COL, Colombia; ESP, Spain; FRA, France; H, Human; HH, Healthy Human; HP, Hospitalized Patient; MYS, Malaysia; PF, Pig Farmer; RCk, Retail Chicken; RCt, Retail Cattle; RP, Retail Pig; ST, sequence type; TUN, Tunisia; USA, United States of America; WW, Wastewater.

database is not complete for the *E. faecalis* species (n=22 virulence genes typical of this species), we created an in-house database with 42 additional virulence genes and tested it using the MyDbFinder (BLAST) tool available at CGE, making a total of 64 virulence genes (Table S1, available in the online version of this article). The 28 assembled genomes were also analysed by PHASTER (PHAge Search Tool Enhanced Release) to identify the presence of prophage sequences [27]. Genomic data for 23S rDNA mutations and for ribosomal protein genes *rplC, rplD* and *rplV* were compared to reference gene loci of the linezolid-susceptible *E. faecalis* V583 (GenBank acc. no. AE016830) using Geneious Prime BLASTN. The presence of *cfr* variants, *poxtA* and plasmid replicons was also performed using Geneious Prime BLASTN and reference *cfr*/*poxtA* sequences [GenBank accession numbers AM408573 (*cfr*1), AJ879565

(*cfr*2), NG_055643 (*cfr*C), PHLC01000011 (*cfr*D) and MF095097 (*poxtA*)].

PATO was used for the analysis of the accessory genome allowing us to infer the accessory genome from the proteomes and cluster them based on 80% protein similarity and 80% of coverage. The set of representative proteins was used to build a binary matrix (presence/absence of proteins in the accessory genome) in the R-environment (http://www.r-project.org) and a tree to classify the strains according to the accessory genomes. The tree was build using IQTree and binary data with model GTR2+FO+ASC+R2 (selected by ModelFinder).

Characterization of *optrA* genetic platforms

Prokka annotation of the 612T *optrA*-containing contig was confirmed using BLASTN, BLASTP and ISFinder (https://isfinder.biotoul.fr/). The same strains used for comparative genomics analysis were mapped against the *optrA*-containing contig of 612T using Geneious Prime in order to determine the presence of the same *optrA* platform and the genome region where it was inserted. Vector NTI advance v11 and Easyfig v2.2.2 [28] were used for drawing *optrA* platform schemes and contig comparisons. Gene functions of the 612T *optrA*-containing contig and the different *optrA* platforms were classified according to the database of Clusters of Orthologous Groups of proteins (COGs) using eggNOG 4.5.1 (http://eggnogdb.embl.de/#/app/home). The identification of contigs belonging to the *optrA*-containing plasmid in 728T strain and the corresponding annotation was determined by comparison with the GenBank database using BLASTN, BLASTP and ISFinder. Vector NTI advance v11 and EggNOG 4.5.1 were used to determine gene function and represent the different plasmid contigs.

RESULTS AND DISCUSSION

Phylogenetic relatedness of *E. faecalis* strains

Fig. 1 represents the phylogenetic tree obtained with SNPs, not afected by recombination, of the 28 *E. faecalis* genomes using as reference the genome of *E. faecalis* V583. The healthy humans included correspond to a human volunteer (Efs 599) of the NIH Human Microbiome Project [29] and two swine farmers (Enfs94/P.En218) working in pig farms from Malaysia [23]. In the four *E. faecalis* genomes from hospitalized patients in China, linezolid resistance was not associated with the consumption of linezolid and an asymptomatic gut colonization was speculated by the authors [30]. Moreover, two of them (13484/27149) were recovered from Chinese farmers [30].

Based on SNP analysis, we observed an overall diverse population of *optrA*-positive *E. faecalis* (range 3–29 258 SNPs) with a correspondence between the sequence types identified (17 STs, 15 nodes) (Fig. 1, Table S2). Strains originating from the same source or country were dispersed across the different tree nodes, suggesting no significant correlation between strain origin and phylogeny as commonly observed in *E. faecalis* [31]. These results further confirm the easy promiscuity of *optrA* that has been identified in enterococci from highly variable epidemiological and genetic backgrounds [2, 32]. The exceptions corresponded to specific strains originating both temporally and locally from the same scenario namely the clinical isolates Iso1Bar/Iso2Bar/Iso3Bar (15–35 SNPs) or Iso4Bar/Iso5Bar (69 SNPs) from an hospital in Barcelona, and the chicken meat isolates 12E/34E (39 SNPs) from Colombia, with close clonal relationships as described [12, 33]. Still, strains from different origins and lacking evident epidemiological relationships were identified as possibly related (<126 SNPs) given their distant geographical and temporal relationships: (i) five ST476 and one ST858 (single locus variant of ST476) isolates from chicken faeces and chicken meat (712T and 697T, respectively) in Tunisia, pigs in Malaysia (P.En250) and hospitalized patients in China (11340, 18026, 27149) differing between 3–121 SNPs; (ii) two ST21 isolates from chicken meat in Tunisia (612T) and a hospitalized patient in China (13484) differing in 113 SNPs; and (iii) two ST489 isolates from retail chicken in Colombia (745E) and cattle in USA (N60443F) differing in 126 SNPs (Table S2).

The finding of the related *optrA E. faecalis* strains (ST21 and ST476) in human clinical and animal hosts from different continents over >5 years (2012–2017) was remarkable and demonstrates that they might originate from an animal reservoir and are able to easily colonize humans. An inter-hospital spread of ST476 *optrA*-carrying *E. faecalis* in different Chinese cities has been previously reported [22]. Noteworthy, most of the food animal strains included in this phylogenetic analysis (ST21, ST59, ST86, ST141, ST227, ST474, ST476, ST489, ST585) have been previously associated with *E. faecalis* from human infections in different continents, linked to linezolid resistance or not [www.pubmlst.org] [9, 12, 13, 34].

Comparative pan-genomic analysis of *optrA*-carrying *E. faecalis*: antibiotic resistance, virulence, prophages and plasmid replicons

We constructed an additional tree using the accessory genome to segregate the strains based on their unique accessory gene repertoire. If we consider two main clades in the core genome tree, the lower clade including the bottom cluster with two main nodes, in addition to strains Enfs51 and Enfs94, corresponded to the strains that remained stable in both trees (Fig. 2). The node placement in both trees was not consistent with the two methods used with 19 out of the 28 genomes changing their tree positions and, in some cases, also their cluster's relatives. This confirms that even though some strains are quite close to one another from a phylogenetic (SNPs) point-of-view, they were not necessarily close in their overall genetic repertoires. Isolates changing their nodes within a given cluster were: 16–372 (ST480) and P.En218 (ST765) co-clustering with ST476-like strains in the accessory genome tree; the phylogenetically distant P.En90 (ST756) and N48037F (ST141) joining and co-clustering into the V583 strain group that moved to the upper clade of the accessory tree; and the ST21 612T and 13484 strains moving into the lower clade of the accessory tree (Fig. 2).

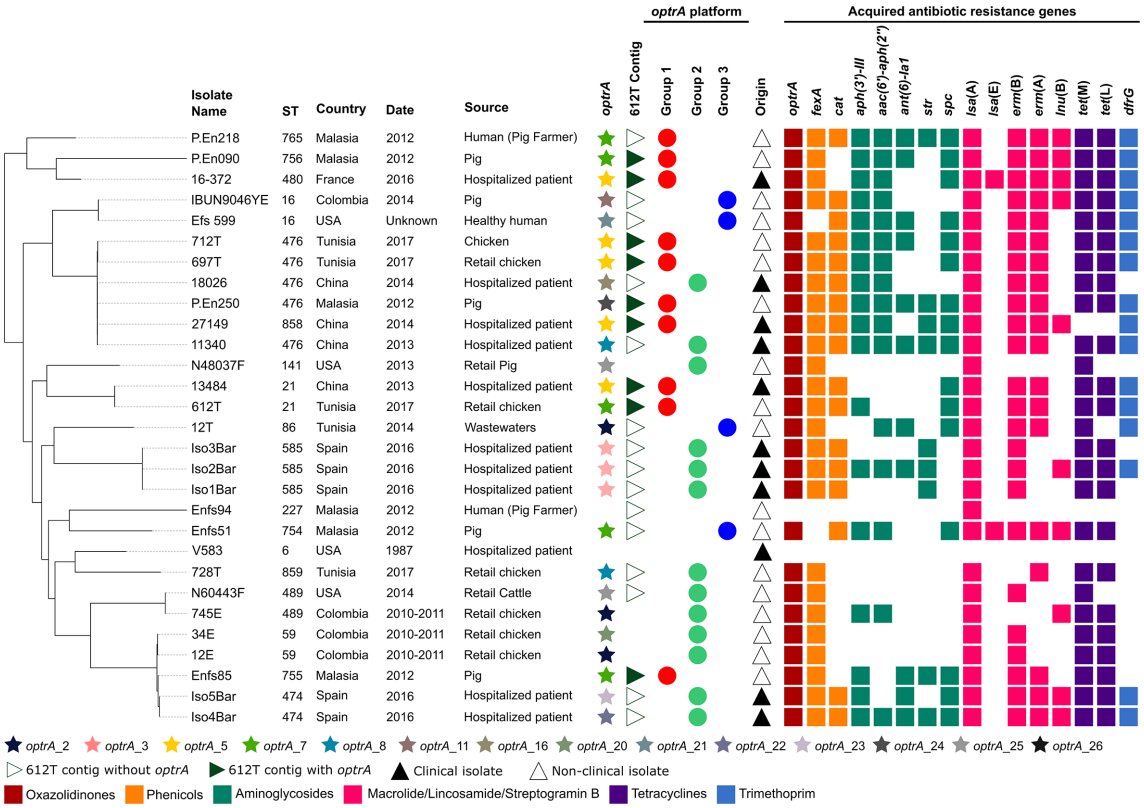

Fig. 1. Phylogenetic tree of the *E. faecalis* strains (*n*=28) analysed in this study. SNP alignment was obtained using standard conditions of CSI Phylogeny and *E. faecalis* strains V583 as reference strain. Recombination regions were eliminated using Gubbins and maximum likelihood tree was obtained using IQ-tree. Stars indicate the presence of the *optr*A gene. The colour of the star indicates the different *optrA* variants as indicated in the key. Triangle filled or not filled with dark-green respectively indicates the presence or absence of the chromosomal 612T contig containing *optr*A used as a reference for comparison of *optrA* chromosomal platforms (please refer to this section below). The different *optrA* genetic platforms are represented with circles filled with red (group 1), green (group 2) and blue (group 3) as in Fig. 6. Group 1 includes isolates carrying *optrA* in the chromosomal Tn*6674*, group 2 includes isolates carrying *optrA-fexA* on medium-size plasmids, and group 3 includes isolates carrying *optrA-araC* on plasmids/chromosome. Coloured cells represent the presence of acquired AMR genes with each colour indicating the correspondent family as indicated in the key (oxazolidinones in red; phenicols in yellow; aminoglycosides in green; macrolides, lincosamides or streptogramins in pink; tetracyclines in purple; and trimethoprim in blue). AMR genes present in only one isolate are not represented in the figure: *cfr* and *poxtA* (Efs599), *tet*(O) (728T) and *tet*(K) (12E). Abbreviations: ST, sequence type. White cells represent the absence of genes.

The analysis of the 28 genomes using the AcCNET tool [35] yielded a set of 4792 protein families representing the accessory genome (Fig. 3). Again, the accessory genome network did not reveal a trend for host-specific proteins (Fig. 3a) but strains grouped according to their ST (Fig. 3b) with protein sharing in specific cases. A total of 1570 unique proteins were identified across all strains when grouping them by ST, with more than half (*n*=921 proteins) corresponding to hypothetical proteins. The greatest number of unique proteins by ST (*n*=388 genes) belonged to a single strain from a retail chicken in Tunisia (ST859). When analysing the unique proteins by origin, all of them (*n*=139) were identified from wastewaters (one ST86, Tunisia), with all other origins sharing protein content. Unique proteins were diversely annotated mainly as surface proteins, ABC transporters, conjugative and mobile proteins, phages, carbohydrate metabolism and transporters (Table S3).

When we compared the isolation source and the genetic content of the 28 strains by clustering them based on their accessory genome (1Fig. S1 and Fig. S2), no clear patterns emerged among related strains. The only exceptions corresponded to those strains carrying *esp* and *cyl* genes as they clustered in the upper part of the accessory genome tree while being more dispersed in the core genome-based tree. Esp (enterococcal surface protein) is located within a pathogenicity island (PAI) and the *cyl* operon encodes a cytolysin with haemolytic activity located on pheromone-responsive plasmids or within a PAI [36], which explains their contribution to the accessory genome of the strains carrying them. Most of these *esp*- and/or *cyl*-positive strains (N48037F, P.En090, Iso1Bar, Iso2Bar, Iso3Bar, Efs599, IBUN9046YE) clustered together in a pan-matrix showing the sharing of accessory proteins, but other pan-proteins may contribute to their close relationship (Fig. S3). This heatmap (Fig. S3) shows

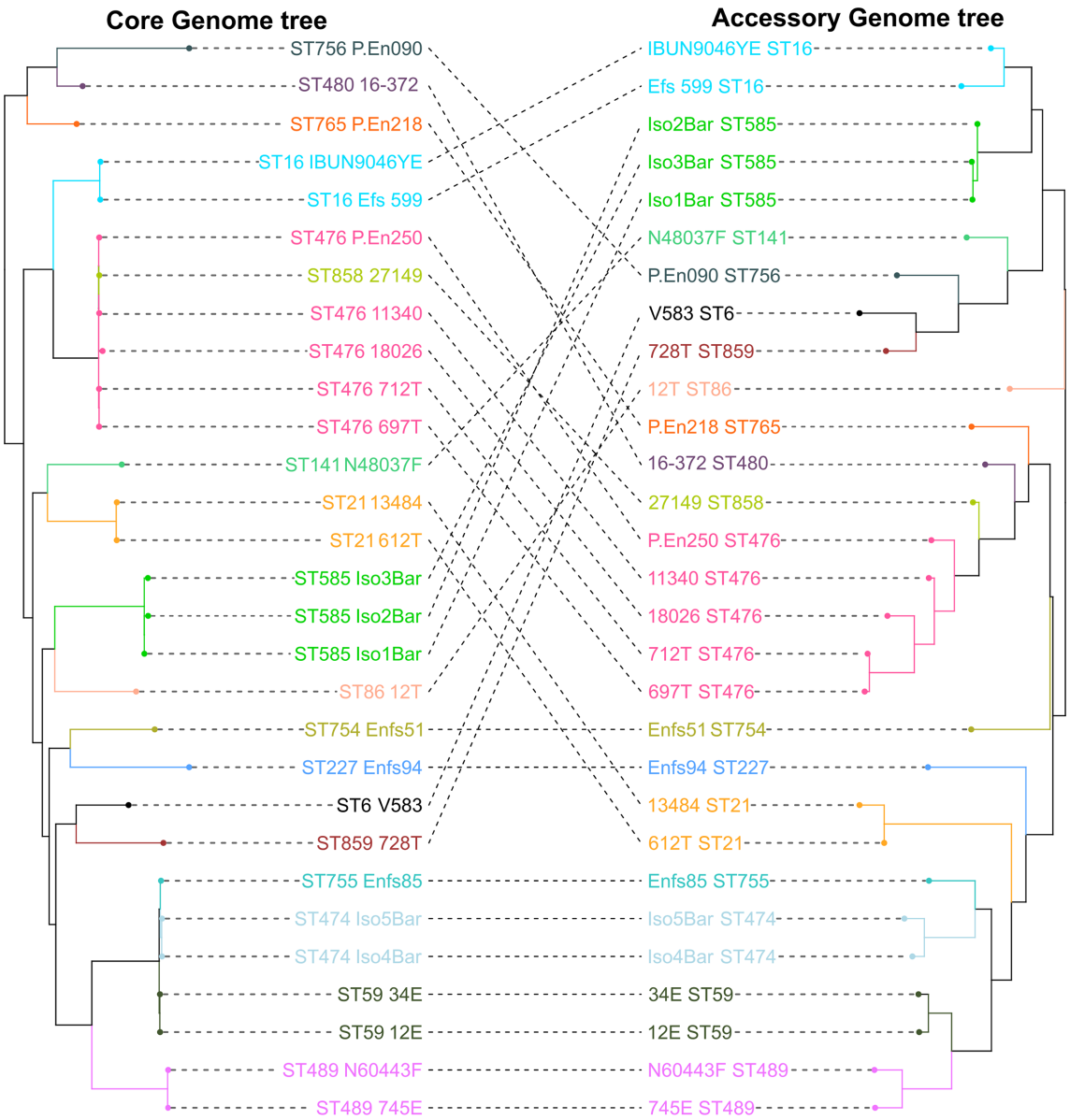

Fig. 2. Comparison of core and accessory genome trees. SNP alignment was obtained using standard conditions of CSI Phylogeny and *E. faecalis* strains V583 as reference strain. Recombination regions were eliminated using Gubbins. Analysis of the accessory genome was performed using PATO and the resulting presence/absence matrix was used to obtain a maximum likelihood tree. Both trees were obtained using IQ-tree. Tree comparison was performed using R package phytools. Abbreviations: ST, sequence type.

the highest number of shared proteins for related isolates (same ST) and according to the accessory strain clusters.

Specific genes codifying for antibiotic resistance, virulence, phages and plasmids are analysed in detail below (Figs 1 and 4).

Antibiotic Resistance

A high amount of antimicrobial resistance genes was identified in most of the genomes, with no clear distinct patterns being observed among related strains. The overall acquired ABR gene profiles identified are the same as that commonly identified in *optrA*-carrying *E. faecalis* strains from non-clinical sources in different regions [15, 18], with no significant differences between clinical (8–15 genes, 11 in average) and non-clinical (1–15 genes, 10 in average) isolates (Fig. 1). They included acquired genes putatively conferring resistance to oxazolidinones, phenicols, aminoglycosides, macrolides, lincosamides or streptogramins, tetracyclines, and trimethoprim, and varying in their homology against reference strains (Table S4). Chromosomal point mutations putatively conferring resistance to linezolid (in 23S rRNA and L3/L4/L22 ribosomal proteins) and ciprofloxacin [*gyrA* (S83I, E87G) and/or

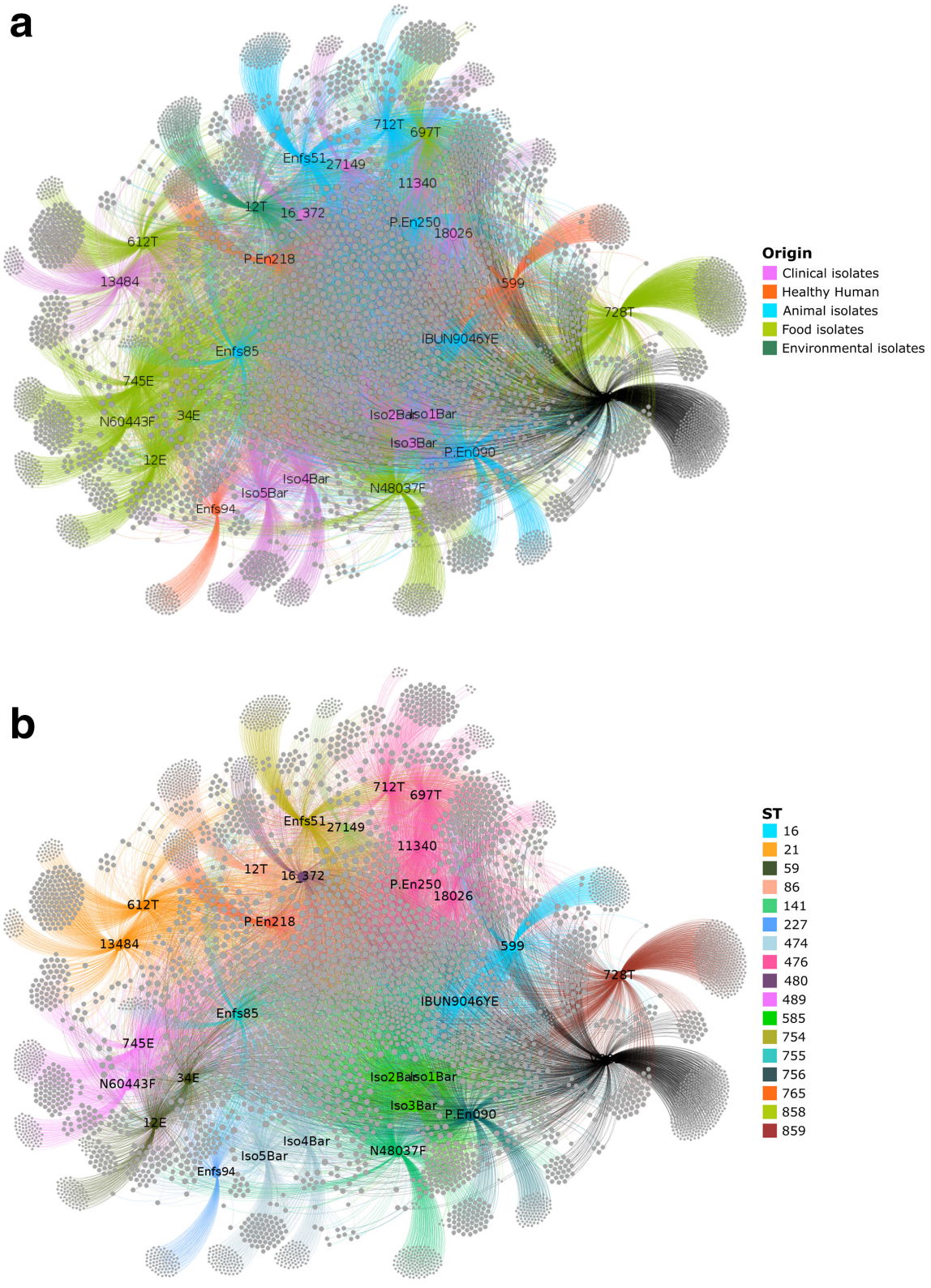

Fig. 3. AcCNET analysis of the 28 *E. faecalis* strains analysed in this study. (a) Representation of the accessory genome network by the origin of isolates. Each colour indicates a different origin as indicated in the figure key. (b) Representation of the accessory genome network by the ST of each isolate. Each colour indicates a ST origin as indicated in the figure key. Abbreviations: ST, Sequence type. Clusters of isolates (larger nodes) as well as proteins based on 80% similarity can be visualized according the colours of the nodes.

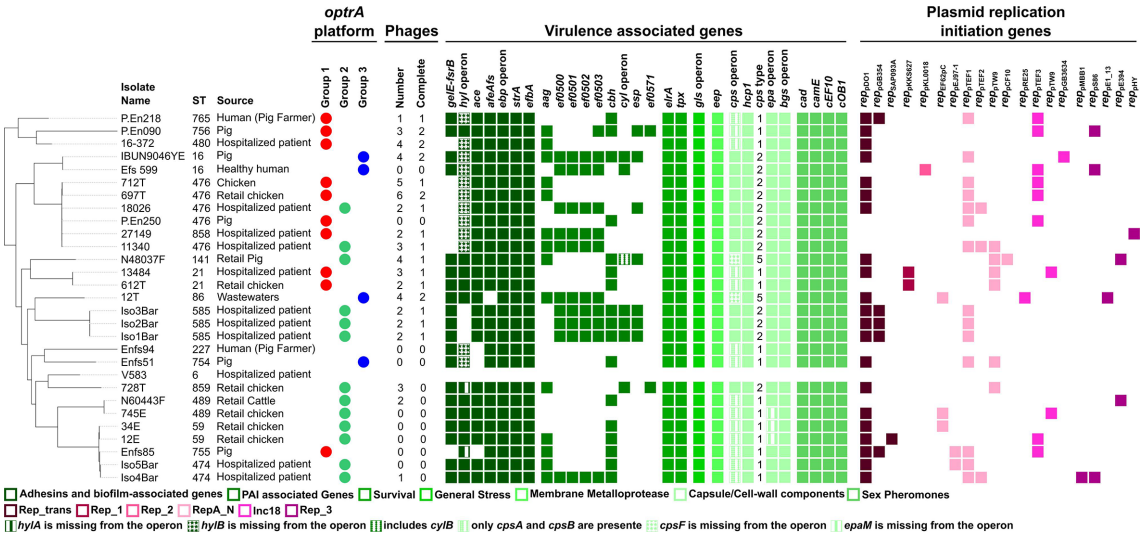

Fig. 4. Profiles of virulence, prophages and replication initiation genes present in the *E. faecalis* (*n*=28) strains analysed in this study. SNP alignment was obtained using standard conditions of CSI Phylogeny and *E. faecalis* strains V583 as reference strain. Recombination regions were eliminated using Gubbins and maximum likelihood tree was obtained using IQ-tree. The presence of the virulence genes, prophages and plasmid replication initiation genes is indicated by coloured squares as indicated in the key (virulence genes in green; prophages in blue; replication initiation genes in pink) and white cells indicate the absence of genes. Group 1 includes isolates carrying *optrA* in the chromosomal Tn*6674*, group 2 includes isolates carrying *optrA-fexA* on medium-size plasmids, and group 3 includes isolates carrying *optrA-araC* on plasmids/chromosome. Abbreviations: ST, sequence type.

parC (S80I)] were also identified in most strains. Regarding linezolid resistance mutations, only synonymous ones were identified in *rplC*/L3 and *rplV*/L22 ribosomal proteins while in the case of the 23S rRNA some strains (*n*=10/28) exhibited mutations not previously described (Table S4). In common to previous studies, all *E. faecalis* strains exhibited an amino acid substitution at position 101 of the *rplD*/L4 protein, which has been linked to both linezolid-resistant and susceptible isolates [7, 9, 37]. Interestingly, the *optrA*-negative Enfs94 is described as linezolid-resistant (MIC not available) according to Tan *et al.* [23], and so its 23S rRNA (G388A/D130N) and *rplD*/L4 protein (T301C/F101L) mutations might contribute to linezolid resistance expression as this strain does not have any other known linezolid resistance mechanism. Curiously, strains from healthy humans presented very different ABR profiles even from the same local/country: ST227 (Enfs94, strain without *optrA*) and ST765 (P.En218) obtained from pig farmers in Malaysia, respectively, carried *lsa*(A) only and 15 antibiotic resistance different genes. The two related clusters of ST476- and ST21-like strains from foodborne animals and hospitalized patients exhibited similar profiles of both acquired ABR genes and point mutations, with genes conferring resistance to aminoglycosides being the more variably present among the different strains. Enterococci displaying the referred ciprofloxacin point mutations in *gyrA/parC* have been commonly identified in samples obtained from animals and foodstuffs, but also from human infections [38, 39]. Interestingly, some of the strains clustering into the same node in the phylogenetic tree shared similar combinations of *gyrA/parC* mutations but the same was not observed for linezolid chromosomal mutations or *optrA* variants which greatly varied even between closely related strains (Fig. 1, Table S4).

Virulence gene profiles

We herein firstly provide an *in silico* analysis of an extended number of putative virulence genes (*n*=64), as the few available studies analysing virulence generally include a small set of 22 genes available in the VirulenceFinder database [12, 33, 37]. The genomes here analysed contained a quite high number of virulence genes (41–61, 49 in average), though with variable gene identity in comparison to reference strains and some in truncated or increased size versions (Table S5). Virulence profiles were generally identical beweteen clinical (41–59 genes) and non-clinical (41–61 genes) isolates (Fig. 4). The related clusters comprising ST476-like, ST21 and ST489 strains exhibited identical virulence gene profiles including 51–55, 44 and 42–44 genes, respectively.

Thirty-six out of the 61 virulence genes detected were present in all strains and they were related with adhesion, survival, general stress, cell surface behaviour, capsule production and sex pheromones, some of which associated with *E. faecalis* PAI. Interestingly, all strains carried four major virulence genes or operons relevant for *E. faecalis* endocarditis, including: *bgsA/bgsB* (glucosyltransferases with a key role in the production of the major cell wall glycolipids, biofilm production and presumably in endocarditis) [40], *gls24/glsB* (stress genes important for survival *in vivo*, and *gls24* in particular to bile salts resistance and hypothetically for virulence in peritonitis and endocarditis) [41], *efbA* (a PavA-like fibronectin adhesin

previously shown to be important in experimental urinary tract infection and endocarditis) [42], and *eep* (a membrane metalloprotease crucial in biofilm formation and required for endocarditis) [43], all exhibiting nucleotide identities varying between 97 and 100% (Table S5). Except for *epaM* missing in three strains, all genes comprising the epa locus (*epaA-epaR*) were found in all strains. The *epa* locus is required for the biosynthesis of an antigenic rhamnopolysaccharide that plays a role in *E. faecalis* urinary tract infections and peritonitis, in intestinal colonization and in protecting against human host defenses [44]. Teng *et al.* have also described that *epa* genes are widespread among *E. faecalis* despite showing differences in sequence and cluster composition [45], but the impact of such differences on their expression remains unknown. Survival [*elrA* (leucine-rich protein A associated with macrophage persistence), *tpx* (thiol peroxidase for oxidative stress resistance)] and sex pheromone-associated (*cad, camE, cCF10, coB1*) genes were also common to all strains.

Among the six virulence gene operons identified (*hylA/hylB, ebpA/ebpB/ebpC, gls24/glsB, cpsC-cpsK, epaA-epaR and bgsA/bgsB*), only the *bgsA/bgsB* was complete in all strains. The six *cps* genes essential for capsule biosynthesis *(cpsC, cpsD, cpsE, cpsG, cpsI*, and *cpsK*) were identified in 14 of the 28 strains obtained from human, environmental and animal (including the ST476 cluster) sources. The identity of these genes varied from those carried by the V583 strain (97–100% in nucleotides), and some corresponded to pseudogenes *(cpsF* and *cpsG*; Table S5). Isolates were classified based on capsule operon polymorphisms in type 1, 2 or 5 as described: (i) CPS type 1 do not express the CPS as they possess only *cpsA* and *cpsB* lacking all remaining genes (*cpsC-cpsK*) essential for capsule production; (ii) CPS type 2 possess all genes (*cpsA-cpsK*) of the *cps* operon (e.g. V583); (iii) and the *cpsF* gene is absent in the CPS 5 serotype [46]. Within each cluster of strains, of the core genome tree, the CPS type was invariant. In common to other studies, the most common CPS type was type 1 (*n*=14), followed by type 2 (*n*=12) and type 5 (*n*=2) [46]. All isolates with the CPS type 2 carried *cpsF* with different mutations in comparison to V583 so the expression of CPS cannot be ruled out only by capsule genotyping. Although isolates with CPS type 1 should not express a capsule, non-encapsulated strains have been already linked to *E. faecalis* endocarditis [47].

Even though we cannot infer by genomics the true opportunistic infection potential of the 28 strains, several of them were enriched in different virulence genes that are known to play key roles in *E. faecalis* pathogenicity. Strains recovered from human and animal sources carried most virulence genes tested including some (*ace*-89%; *gelE-fsrB*-71%; *agg*-50%; *elrA*-100%) that were previously described in *E. faecalis* hospital/epidemic strains [48, 49]. Four genes from the module A of *E. faecalis* PAI (EF0500, EF0501, EF0502 and EF0503) have been considered the most discriminatory within the PAI to identify putative pathogenic strains [50]. In that study, those genes were present in all strains characterized as potential pathogens and not in commensals. In our 28 genomes, they were identified in eleven strains dispersed in the phylogenetic tree, seven of which were obtained from hospitalized patients (ST474, ST476, ST585, ST858), two were obtained from pig faeces (ST16) and the remaining two were obtained from a wastewater isolate (ST86) and from a healthy human (ST16). In general, no obvious differences arose from the profiles and sequences of virulence genes between strains obtained from foodstuffs/livestock and those causing human infections and only the occurrence of *cyl* and *cpsC-K* seemed lineage (ST)-specific (Table S5). These data support previous works indicating the generalist lifestyle of *E. faecalis* and the opportunist nature of the human infections they cause [31, 51].

Prophages

Prophages-associated sequences are commonly found in clinical strains such as the well-known V583 harbouring seven prophage-like elements, of which only two correspond to intact ones [52]. They are also commonly found in human and animal sources but the identification and distribution of phages is scarcely known in well-characterized *E. faecalis* populations [53].

Even though its difficult to assess the precise content of phages for each genome given the similarity of different phage proteins in several phages, a remarkable variety of putative prophages and number of intact phages were identified in most strains of this study (Table S6, Fig. 4). PHASTER identified the presence of intact phages in 16 out of 28 isolates from both clinical and non-clinical sources. Four isolates from varied sources (farm pig, chicken meat, clinical, wastewaters) carried two intact phages each. Isolates carrying *optrA*::Tn*6674* contained the highest number of prophages (two carrying 5–6 per genome). We did not identify any obvious relationship between the presence of prophages and isolation source or phylogeny. Nevertheless, and in common to the aforementioned PAI-related genes, the lower clade of the pangenome tree included genomes without intact phages and with a generally poorer number of prophages. Intact or relics of prophages highly identical to sequences from common enterococcal prophages such as phiL1A and vB_IME197 were identified in most strains, but the exact phage(s) present in each genome cannot be ruled out with accuracy. The phiL1A and vB_IME197 phages are widespread in *E. faecalis* from variable sources [31] and were originally described in clinical settings in UK and China, respectively [54, 55]. Putative prophages associated with other bacterial genera (*Listeria, Bacillus, Lactobacillus*) from *Firmicutes,* all classified as temperate *Siphoviridae* bacteriophages and originally obtained from variable sources (humans, farms, hospital sewage) [55–57], were also identified. A comparative genomic approach performed by La Rosa *et al.* demonstrated that phage03-like elements (phage03 from V583 similar to the *Listeria monocytogenes* phage B054; GenBank acc. no. NC_009813) have an increased *in vivo* virulence and pathogenicity and are particularly enriched in clinical isolates [58]. In this study, this phage was putatively identified in three related ST476 strains from hospitalized patients, farm and retail chicken in China and Tunisia (Table S6), once more enphasizing the lack of a demarcation between clinical and

non-clinical genomes in relevant mobile genetic elements (MGEs).

We cannot discard the possibility of an erroneous output of the 'defective' prophages if these are distributed in different contigs, but the finding of several intact prophages might indicate recent integrations. In any case, the number of prophage sequences found in such a small set of genomes evidences the contribution of prophages to the diversity of *E. faecalis* populations by acting as vectors of different gene traits during intra- and inter-species genetic exchange, and that their spread is independent from the isolation source [53].

Plasmidome

The extraction and analysis of plasmid replicon sequences was performed using the *Firmicutes* PlasmidFinder 2.1 (CGE) which in turn is based on the recent and comprehensive review of Lanza *et al.* [59]. All but the ST227 *optrA*-negative strain from the healthy farmer in Malaysia (with no *rep*) carried at least one *rep* gene (1–5 *rep* genes) (Fig. 4, Table S7). The strains analysed carried plasmids belonging to all the six families known to date among enterococci, either replicating by a theta mechanism (Rep_3, Inc18, RepA_N) or by rolling circle replication (RCR: Rep_trans, Rep_1 and Rep_2) [60]. When it was possible to associate *optrA* to a specific plasmid, different plasmid types were identified (Rep_3, Inc18, RepA_N) (Table S7).

Within rolling-circle replicating (RCR) plasmids, most strains (21 out of 28) carried one or two replicons of the Rep_trans family, namely rep_{pDO1} from *E. faecium* (*n*=21), and additionally rep_{pGB354} from *Streptococcus agalactiae* (*n*=5) linked to chloramphenicol resistance (*catA7*) or $rep_{SAP093A}$ from *S. aureus* (*n*=1). Rep_trans plasmids are of small size and widespread among diverse enterococcal populations, being often mobilized by other ABR conjugative theta-replicating plasmids or integrated in the chromosome [59, 61].

The content in the theta-replicating plasmids was highly variable and included replicons from pheromone-responsive RepA_N plasmids, typical of *E. faecalis*, the broad-host range Inc18, and Rep_3 which includes small non-conjugative and cryptic plasmids obtained from a variety of environments but often associated with bacteriocin production. Only the pHTβ-like plasmids were not identified as they are usually exclusive of *E. faecium* species [61]. Notably, all but two strains carried at least one *rep* from Inc18 or RepA_N plasmids that correspond to the conjugative plasmids usually involved in the acquisition and spread of ABR among different *Firmicutes* genera. Within RepA_N, pheromone-responsive plasmids were present in most strains. They are unique and highly important in the diversification and evolution of *E. faecalis* species and the finding of truncated pTEF1 rep genes in most strains might indicate events of chromosomal integration across similar IS elements with the transfer of plasmid and large genomic DNA portions as proposed by Manson *et al.* [62]. In fact, also several PAI-related genes were identified as truncated (Table S7), which can be related to the movement of the *E. faecalis* PAI by pheromone-like plasmid integration, as described [62].

The high variability of plasmid content found in *optrA*-carrying *E. faecalis* (Table S7) illustrates the diverse nature of MGE that can be found in *E. faecalis* in general and in particular in *optrA*-carrying strains.

optrA variants and genetic context

The *optrA* nucleotide sequences of the 28 genomes were compared against the original *optrA* gene (*optrA_1* or $optrA_{WT}$; GenBank acc. no. KP399637). Fourteen different variants (seven new), with point mutations not including the original one, were identified (Fig. 1). As the nomenclature of *optrA* variants is not uniformed across different studies [9, 63] and available databases (CGE), we compiled the different designations and gene mutations in Fig. 5 to facilitate results' interpretation. Three different *optrA* variants were identified among the four epidemiological related chicken strains from Tunisia (*optrA_5, optrA_7 and optrA_8*) which have been also identified in different *E. faecalis* clones (e.g. ST476) from humans and pigs in China [21]. *optrA* location was inferred, when possible, from the genomic sequences. This gene was either located on the chromosomally-borne Tn*6674* (*n*=10) or on medium-size plasmids (30–60 Kb; *n*=14) belonging to the plasmid families (RepA_N, Inc18 and Rep_3) usually involved in the acquisition and spread of AMR among different genera of *Firmicutes* [59]. Independently of *optrA* localization, the analysis of the *optrA* genetic platforms carried out in the 27 strains enabled their grouping into three main groups (Fig. 6). These groups, dispersed through the nodes of the phylogenetic tree, showed no obvious correlation with their strain genetic background, neither with the respective *optrA* gene variant (Fig. 1). They corresponded to:

(i) Group I (*n*=10) included three chicken strains from Tunisia (one ST21 and two ST476), all the pig strains from Malaysia (ST476, ST755, ST756 and ST765) and three hospital strains from France (ST480) and China (ST21 and ST858) carrying five different *optrA* variants (Fig. 6). Three additional genomic regions were included for comparison (Tn*6674*, TZ2 and A101). Whole genome sequences of these strains are not available so they were not included in the genomic analysis. These strains shared a chromosomal segment containing *fexA-optrA* adjacent to the entire Tn*554* from *S. aureus*, varying between 11.7–12.9 kb, which is identical to the recently identified Tn*6674* [64]. Tn*6674* is a chromosomal 12.9 kb Tn*554*-related transposon obtained from the *E. faecalis* strain E1731 (GenBank acc. no. MK737778.1) that contains 7 ORFs: *tnpABC* and resistance genes [*spc-erm(A)-fexA-optrA*] conferring resistance to spectinomycin, macrolides, phenicols and oxazolidinones, respectively; so *fexA* and *optrA* are additionally carried in Tn*6674* in comparison to Tn*554*. As illustrated in Fig. 6, that includes not only some of the genomes of this study but also *optrA* platforms available in GenBank, Tn*6674* has been driving *optrA* spread amongst hospitalized patients

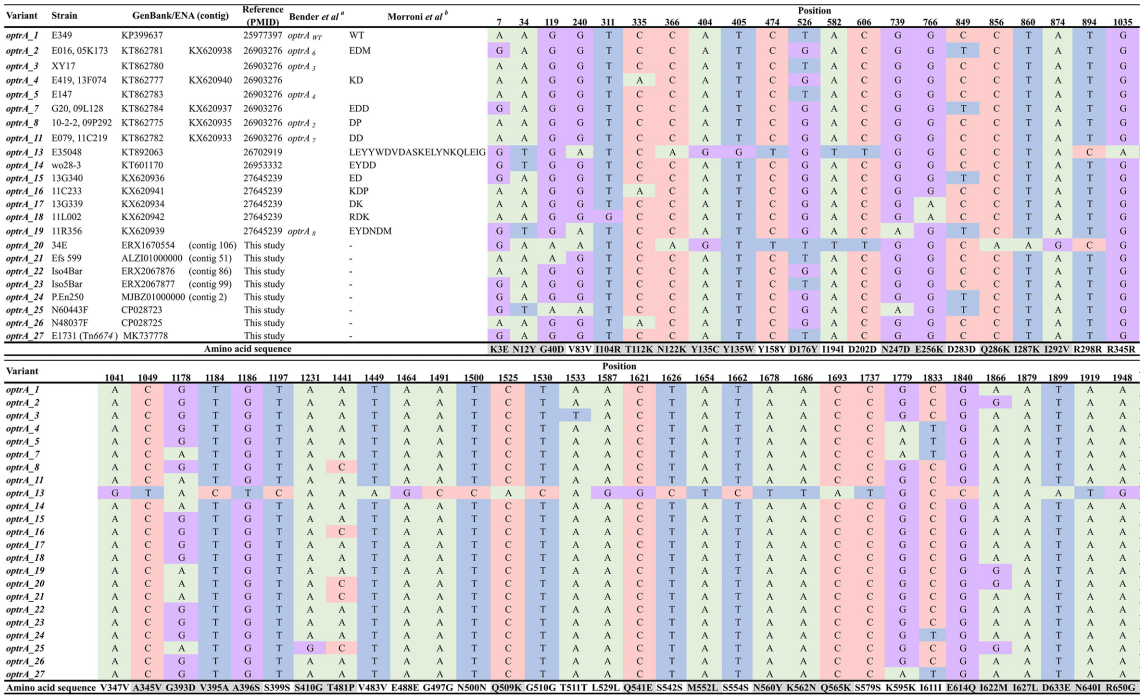

Fig. 5. Representation of the alignment of relevant SNPs in the different *optrA* variants. Different colours indicate the different nucleotides: green – adenine; blue – thymine; red – cytosine; purple – guanine. Squares in grey indicate non-synonymous amino acid mutations. The variant names used are the same as those included in the CGE database until *optrA_19* and new variants were attributed from *optrA_20* to *optrA_27*. Variant *optrA_4* is also labelled as *optrA_6* in CGE database; variant *optrA_8* is also labelled as *optrA_9* and *optrA_10* in CGE database; variant *optrA_11* is also labelled as *optrA_12* in CGE database. [a]reference [9]. [b]reference [63] . Abbreviations: NA, European Nucleotide Archive.

(France, China), healthy humans (China, Malaysia) and food-producing animals (China, Malaysia, Tunisia) isolates across different continents during 2012–2018. A similar structure was also recently described in an *optrA*-carrying ST16 *E. faecalis* from a patient in Greece (sequence not available) [11]. For the isolates and corresponding studies possible to access, this genetic platform could be positively transferred by filter mating [25].

(ii) Group II (*n*=11) included chicken strains from Colombia (two ST59 and one ST489) and Tunisia (ST859), cattle and pig samples from USA (ST141 and ST489), and the clinical strains from Spain (three ST585 and one ST474) and China (two ST476) carrying eight *optrA* variants. *optrA*-carrying regions of pE349 and p10-2-2 plasmids were also included. Whole genome sequences for these strains are not available so they were not included in the genomic analysis. Most of them (*n*=7) contained a common *impB-fexA-optrA* segment that showed high similarity to the originally described *optrA*-positive pE349 plasmid (human, ST116; GenBank acc. no. KP399637) as well as other plasmids obtained from human and pig strains in China (ST27, ST59, ST480, ST585) [21]. This plasmid backbone has been also recently described in *E. faecalis* isolates from Sweden and Ireland [3]. *impB* showed the highest homology with a type VI secretion system from *E. faecium* (GenBank acc. no. ERK33158) and has been suggested to be a plasmid-associated DNA repair gene eventually playing a role in the survival under stressful environments [65]. The ST859 chicken strain (728T) additionally carried *erm*(A) in the same contig of 12.9 kb in common to pig ST59 (p10-2-2 plasmid) and hospitalized patient ST476 strains from China [21]. *optrA* was located on plasmids ranging from ca. 30 to 60 kb in all cases [strains from Group 2 of this study (Fig. 6) and those referred from China, Colombia and the USA] [5, 12, 14, 21] belonging to RepA_N, Inc18 or Rep_3 families whenever known. Some *optrA*-carrying plasmid segments (*n*=5) of this group were flanked by a transposase of the IS*L3* family from *Streptococcus suis* (GenBank acc. no. WP_029876135.1), a species previously recognized as a reservoir of *optrA* [32].

(iii) Group III comprised four miscellaneous strains from pigs (ST16 and ST754), humans (ST16) and wastewaters (ST86). ST16 from a healthy human in USA (Efs 599) and ST86 from wastewaters (12T) in Tunisia commonly carried *optrA* adjacent to an *araC-NADH-spoVG-helicase* region highly similar to that of a blood *E. faecalis* E016 strain from China, despite the fact that in these cases *optrA* was variably located on a plasmid (ST86) or on chromosome (E016) [21, 37]. The two remaining strains (IBUN9046YE and Enfs51) only shared the *araC-optrA* segment with those aforemen-

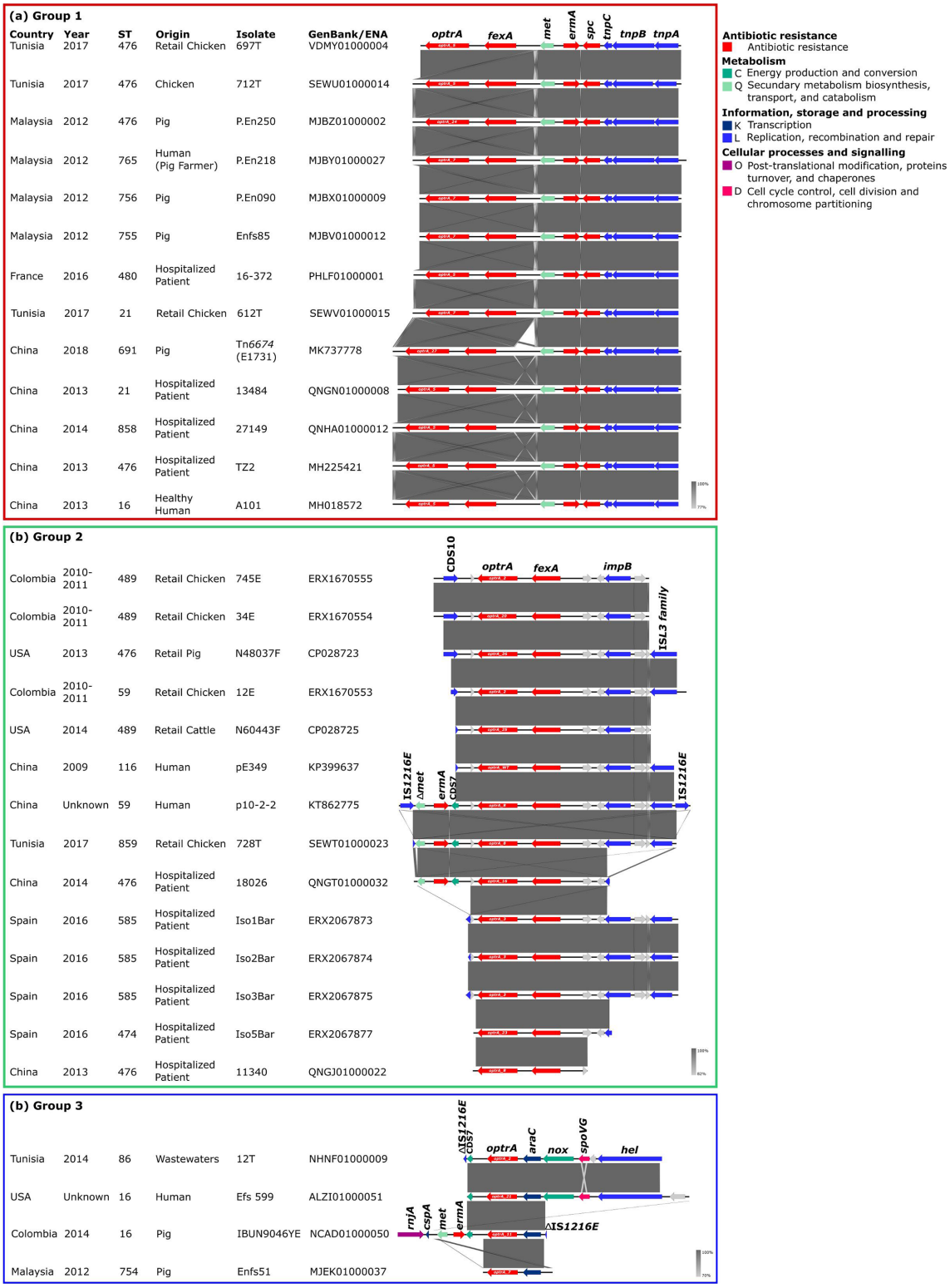

Fig. 6. Different platforms carrying the *optr*A gene. Panel (a) group 1; (b) group 2; (c) group 3. Group 1 includes isolates carrying *optrA* in the chromosomal Tn*6674*, group 2 includes isolates carrying *optrA-fexA* on medium-size plasmids, and group 3 includes isolates carrying *optrA-araC* on plasmids/chromosome. Coloured arrows represent genes and the arrows indicate the orientation. Grey arrows represent hypothetical proteins of unknown function; the remaining colours represent the gene function according to the COGs database classification indicated in the key. Abbreviations: ST, sequence type; ENA, European Nucleotide Archive.

tioned, which has been commonly described flanked by transposases and Tn*558* relics in different enterococcal species [12, 21].

Features of a novel chimeric pheromone-responsive-like *optrA*-carrying plasmid

We were able to presumptively identify the contigs belonging to the *optrA*-carrying plasmid of the Tunisian 728T strain from this group (Fig. S4). The sum of such contigs was 39173 bp, which is compatible with the size inferred from the S1-PFGE gel (~40 kb) [64]. Using EggNOG it was possible to identify the function of some of the proteins codified by the putative ORFs encoded in the plasmid, including ABR [*optrA*, *fexA*, *erm*(A), *tet*(L)], plasmid replication and stability (*repA*, *parA, prgN*), but not of conjugative transfer (compatible with the negative conjugation results), a toxin-antitoxin (TA) system (*mazE, mazF*), class II bacteriocin (*bacA, bacB*), cell fillamentation (*fic*), and a putative ABC-type multidrug transporter (CDS_2 and CDS_3 in the figure). The entire contig 23 of 12971 bp (*ermA-optrA-fexA-impB*; Figs 6 and 7) was identical to that of p10-2-2 plasmid from a pig ST59 *E. faecalis* in China [21]. RepA had the highest homology (92% in nucleotides) with that from the pheromone-responsive plasmids pTW9 (GenBank acc. no. NC_014726), pPD1 (GenBank acc. no. D78016) and pCF10 (GenBank acc. no. AY855841.2) previously identified in MDR *E. faecalis* strains obtained from different hosts and areas. Upstream of *repA* is *parA* and *prgN*, previously described in the pheromone-responsive pTEF2 (GenBank acc. no. AE016831) and pPD1 *E. faecalis* plasmids as putatively involved in plasmid partitioning and replication control, respectively. The region spanning from *repA* to *prgN* had the highest homology with the same *repA-parA-prgN* region of the VanB-type pMG2200 plasmid (GenBank acc. no. AB374546), still of only 73% in nucleotides, meaning this corresponds to a putative new plasmid replication region. Bacteriocin and TA gene clusters were similar to that from the pheromone-responsive pAMS1 plasmid (GenBank acc. no. EU047916) and that described in an *optrA*-positive pEF123 plasmid from chicken *E. faecalis* in China (GenBank acc. no. KX579977), respectively. The ABC transport system was described in pVEF3 (GenBank acc. no. NC_010980) and pVEF4 (GenBank acc. no. FN424376) plasmids from chicken *E. faecium* strains, though it was not associated with a reduced susceptibility to several substances tested at that time by Sletvold *et al.* [66]. Taking all together, our data suggests that this *optrA*-carrying plasmid from chicken meat corresponds to a putative novel pheromone-responsive-like plasmid containing genes of different origins including the first *optrA* pE349 plasmid described, as has been commonly observed worldwide and particularly in the non-clinical setting [2, 14, 21, 37].

Comparison of chromosomal *optrA*-carrying contigs

We further compared the entire chromosomal *optrA*-carrying contig (~44 kb) of a strain from group 1 (612T) by mapping it against the remaining genomes (Fig. 7).

In addition to the Tn*6674*-like element, this contig included genes involved in different cellular functions like oxidative stress response (*tpx*), redox-sensing transcriptional regulation (*rex*), carbohydrate transport and metabolism [*rbsUDKR* operon (ribose transport and regulation), *ulaA* (PTS system ascorbate-specific transporter)], coenzyme transport and metabolism (*folC* and *thiI* – folate and thyamine biosynthesis respectively), protein translation (*valS*, valyl-tRNA synthetase), phosphate transport and regulation (*phoB*), and DNA mismatch repair (*mutS2*). It was interesting to identify the same conserved chromosomal region (32199 bp from Δ*radC* to *mutS2*) in all but the animal/food strains from Colombia and one from the USA (*n*=24/28), independently if strains carried *optrA* on this or other chromosomal contigs, on plasmids or even lacked it. *optrA* adjacent to Tn*554* or Tn*554* relics has been described either embedded on chromosomal platforms or on plasmids, in a plethora of enterococci strains from different host and regions [19, 21, 63].

As the transposons of Tn*554* family, Tn*6674* does not fall into any standard transposon category since it lacks inverted or direct terminal repeats, and it does not generate a duplication of a target sequence upon transposition. It has been suggested that Tn*554* is a prophage-like element characterized by a highly efficient site-specific integration-excision mechanism [67]. Transfer of Tn*554* between Gram-positive species, especially known between *S. aureus*, typically involves integration into another mobile element and hitch-hiking across via conjugation, transformation or transduction, so it can be found in multiple copies in the cell. Li *et al.* [68] have experimentally confirmed the ability of Tn*6674* to form circular forms and suggested its functional activity. Tn*554* has an exceptional preference for integration into a chromosomal insertion site called att*554*, at a lower frequency into secondary insertion sites (mostly in the chromosome) or more rarely on plasmids if that site is occupied or deleted, following transfer [67, 69]. The transposition of Tn*554* generally interrupts the *radC* gene encoding a putative DNA repair protein and we herein confirm the same with Tn*6674* (Δ*radC* in Fig. 7) [70]. The identification of the *optrA*-carrying Tn*6674* adjacent to a common chromosomal 32 kb platform in most of the strains analysed suggests that this specific chromosomal region may be a hotspot for Tn*6674* integration and mobilization as observed for other bacterial species in which hotspots tend to be flanked by recombination and repair core genes [71]. A GenBank search of this 32 kb nucleotide sequence revealed its presence in >500 other *E. faecalis* genomes (562 out of 2701 subject genomes) from different sources (data not shown), thus demonstrating its conserved nature among this species. Remarkably, the Enfs94 strain from the healthy farmer lacking *optrA*, whom was in contact with pigs colonized with *optrA*-positive *E. faecalis*, also contained this conserved chromosomal region and an *optrA* acquisition might occur.

According to available studies, Tn*554* seems to transfer primarily by transducing phages independently of flanking regions [69]. Although the extent of horizontal gene transfer due to transduction with bacteriophages among enterococci remains largely unknown, the transfer of ABR genes between

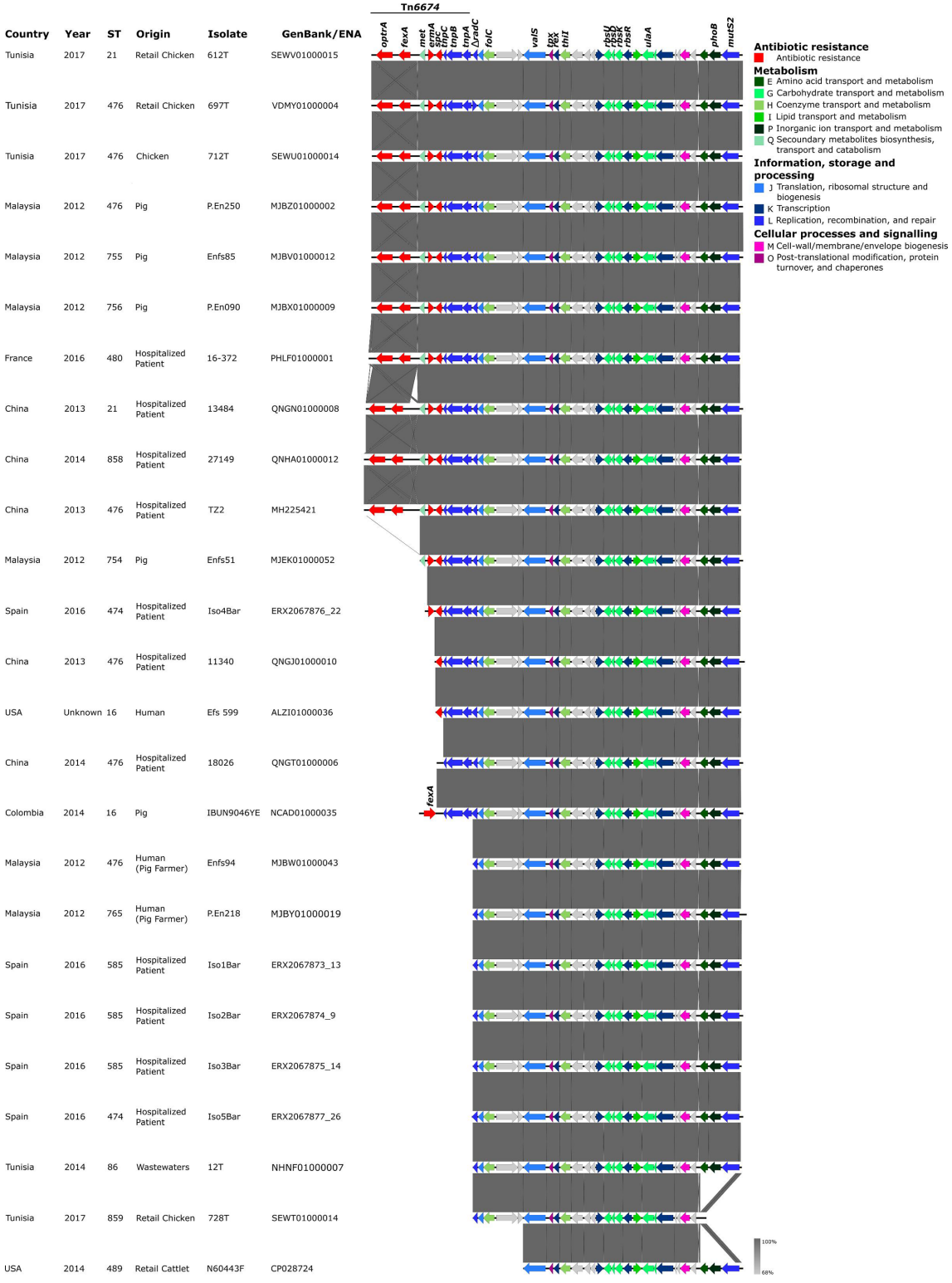

Fig. 7. Comparison of the chromosomal 612T *optr*A-containing contig with the remaining strains. Coloured arrows represent genes and the arrows indicate the orientation. Grey arrows represent hypothetical proteins of unknown function; the remaining colours represent the gene function according to the COGs database classification indicated in the key. Abbreviations: ST, sequence type; ENA, European Nucleotide Archive.

different enterococcal species mediated by phages has been already demonstrated [54, 60], as well as the effect of specific antibiotics (e.g. fluoroquinolones [52]) in prophage release and gene spread within *E. faecalis* isolates. *optrA* was not identified as integrated in a prophage in this work, but a diversity of prophages was found integrated in the chromosome of several strains, with those carrying *optrA* in Tn*6674* being generally the most enriched in prophage sequences including intact phages. Strikingly, phage-mediated lateral transduction was recently described as a putative new universal mechanism of transfer promoting the efficient exchange of large portions of bacterial genomes at high frequencies [72]. *optrA* has been identified in many and diverse genetic platforms (plasmids, ICEs, phages) amongst different Gram-positive genera (enterococci, staphylococci, streptococci). Among enterococci, mostly *E. faecalis*, *optrA* has been commonly described as adjacent to *fexA* (conferring resistant to phenicols), relics of Tn*554* and Tn*558* [*optrA-erm*(A) clusters] and surrounded by ISs (e.g. IS*1216* elements) in a great variety of several and unique combinations of these and other sequences, on plasmids or chromosomal platforms [21]. In *Streptococcus suis* pig isolates, *optrA* has been recently described within a novel composite IS*1216E-araC-optrA-hp-cat*pC$_{194}$-IS*1216E* transposon integrated on a phage similar to Φm46.1 [73]. *optrA* was identified in prophages and ICEs only in *Streptococcus suis* [32, 73, 74], but the mobilization of *optrA* through phage-mediated transduction cannot be discarded also in enterococci. Wang *et al.* [75] also recently demonstrated by a metagenomic approach in swine feedlot wastewater samples that phages were more likely to harbor ABC transporter family and ribosomal protection genes including *optrA*. Adding to the fact that transfer frequency of Tn*554* via transduction is higher than most plasmids [69], we could expect a tremendously fast dissemination of *optrA* among different Gram-positive species and across different hosts and settings, as seems to be occurring more expressively in the animal production setting. It has been recently hypothesized that the virome may play a larger role in ABR transfer in non-clinical environments where phages are abundant and frequently carry ABR genes [76].

Conclusions

This work constitutes one of the most comprehensive studies in comparing *optrA*-positive *E. faecalis* genomes from disparate origins and the first to depict extended profiles in key adaptive features including antibiotic resistance, virulence, prophages and the plasmidome. Our study evidenced the amazing variability of clones and *optrA*-genetic platforms or plasmids that can be found in a single species, and on the other hand the easy possibility of finding identical strains or *optrA* genetic platforms in samples from different sources and distant countries. The analysis of publicly available *E. faecalis* strains carrying *optrA* revealed that its immediate gene vicinity was generally identical in chromosomal (Tn*6674*) and plasmid (*impB-fexA-optrA*) backbones. We herein identified a common Tn*554::fexA-optrA* chromosomal platform, the recently identified novel Tn*6674*, in different linezolid-resistant *E. faecalis* clones from livestock, meat, animal farmers and hospitalized patients across different continents, and proposed a conserved chromosomal hotspot for *optrA-fexA* integration. *optrA*-carrying *E. faecalis* strains from livestock/foodstuffs and human infections could be closely related and harbour indistinguishable key adaptive features, which reinforces the generalist lifestyle of this species and highlights the relevance of the animal reservoir and potential role of the food-chain in the transmission of *optrA*-positive strains. On the other hand, we provide pieces of evidence that *E. faecalis* can easily acquire and exchange *optrA*-carrying platforms or plasmids with other Gram-positive pathogens, besides a variety of other accessory genetic information from bacteria of the *Firmicutes* phylum. The finding of *optrA* in strains sharing identical adaptive features and genetic backgrounds across different hosts and countries suggests the occurrence of common and independent genetic events occurring in distant regions and might explain the easy *de novo* generation of *optrA*-positive strains. The usual association of *optrA* with Tn*554*-like or IS-like elements will certainly aid in a global and dramatic increase of *optrA* carriage and spread with a serious impact on the efficacy of linezolid for the treatment of Gram-positive infections.

Funding information
This work was supported by the Applied Molecular Biosciences Unit - UCIBIO which is financed by national funds from FCT (UIDB/04378/2020). It was also supported by the QREN projects NORTE-01–0145-FEDER-000024 and NORTE-01–0145-FEDER-00001. We are grateful to the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Food- and Water-borne Infection Study Group (EFWISG) Study Group. ARF acknowledges the Junior Research Position (CEECIND/02268/2017 - Individual Call to Scientific Employment Stimulus 2017) granted by FCT/MCTES through national funds. APT was financed by the Sara Borrell Research Grant CD018/0123 granted by Instituto de Salud Carlos III and co-financed by the European Development Regional Fund (A Way to Achieve Europe program).

Author contributions
A.R.F., A.P.T. and L.P. designed the study. A.R.F., A.P.T. and V.F.L. selected and provided characterized genomes and their corresponding epidemiological information, and analysed the genomic sequence data. A.R.F., A.P.T., C.N., V.F.L. and L.P. analysed the data. A.R.F., wrote the manuscript. L.P. and V.F.L., provided the analysis tools and computing resources needed for the study. All authors contributed to the editing and to a critical analysis of the manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Data Bibliography
The genome assemblies of all the genomes in this study were downloaded from the NCBI genome database, and identifiers are listed in Table 1.

References

1. Bender JK, Cattoir V, Hegstad K, Sadowy E, Coque TM *et al*. Update on prevalence and mechanisms of resistance to linezolid, tigecycline and daptomycin in enterococci in Europe: towards a common nomenclature. *Drug Resist Updat* 2018;40:25–39.
2. Sadowy E. Linezolid resistance genes and genetic elements enhancing their dissemination in enterococci and streptococci. *Plasmid* 2018;99:89–98.

3. Deshpande LM, Castanheira M, Flamm RK, Mendes RE. Evolving oxazolidinone resistance mechanisms in a worldwide collection of enterococcal clinical isolates: results from the SENTRY antimicrobial surveillance program. *J Antimicrob Chemother* 2018;73:2314–2322.
4. Brenciani A, Fioriti S, Morroni G, Cucco L, Morelli A *et al*. Detection in Italy of a porcine *Enterococcus faecium* isolate carrying the novel phenicol-oxazolidinone-tetracycline resistance gene poxtA. *J Antimicrob Chemother* 2019;74:817–818.
5. Tyson GH, Sabo JL, Hoffmann M, Hsu C-H, Mukherjee S *et al*. Novel linezolid resistance plasmids in Enterococcus from food animals in the USA. *J Antimicrob Chemother* 2018;73:3254–3258.
6. de Jong A, Simjee S, Rose M, Moyaert H, El Garch F *et al*. Antimicrobial resistance monitoring in commensal enterococci from healthy cattle, pigs and chickens across Europe during 2004-14 (EASSA study). *J Antimicrob Chemother* 2019;74:921–930.
7. Elghaieb H, Freitas AR, Abbassi MS, Novais C, Zouari M *et al*. Dispersal of linezolid-resistant enterococci carrying poxtA or optrA in retail meat and food-producing animals from Tunisia. *J Antimicrob Chemother* 2019;74:2865–2869.
8. Gawryszewska I, Żabicka D, Hryniewicz W, Sadowy E. Linezolid-Resistant enterococci in Polish hospitals: species, clonality and determinants of linezolid resistance. *Eur J Clin Microbiol Infect Dis* 2017;36:1279–1286.
9. Bender JK, Fleige C, Lange D, Klare I, Werner G. Rapid emergence of highly variable and transferable oxazolidinone and phenicol resistance gene optrA in German Enterococcus spp. clinical isolates. *Int J Antimicrob Agents* 2018;52:819–827.
10. Sassi M, Guérin F, Zouari A, Beyrouthy R, Auzou M *et al*. Emergence of optrA-mediated linezolid resistance in enterococci from France, 2006-16. *J Antimicrob Chemother* 2019;74:1469–1472.
11. Tsilipounidaki K, Gerontopoulos A, Papagiannitsis C, Petinaki E. First detection of an *optrA*-positive, linezolid-resistant ST16 *Enterococcus faecalis* from human in Greece. *New Microbes New Infect* 2019;29:100515.
12. Càmara J, Camoez M, Tubau F, Pujol M, Ayats J, Enterococci AL, mara J C *et al*. Detection of the novel optrA gene among linezolid-resistant enterococci in Barcelona, Spain. *Microb Drug Resist* 2019;25:87–93.
13. Cui L, Wang Y, Lv Y, Wang S, Song Y *et al*. Nationwide surveillance of novel oxazolidinone resistance gene optrA in Enterococcus isolates in China from 2004 to 2014. *Antimicrob Agents Chemother* 2016;60:7490–7493.
14. Wang Y, Lv Y, Cai J, Schwarz S, Cui L *et al*. A novel gene, optrA, that confers transferable resistance to oxazolidinones and phenicols and its presence in Enterococcus faecalis and Enterococcus faecium of human and animal origin. *J Antimicrob Chemother* 2015;70:2182–2190.
15. Tamang MD, Moon DC, Kim S-R, Kang HY, Lee K *et al*. Detection of novel oxazolidinone and phenicol resistance gene optrA in enterococcal isolates from food animals and animal carcasses. *Vet Microbiol* 2017;201:252–256.
16. Sharkey LKR, O'Neill AJ. Antibiotic resistance ABC-F proteins: bringing target protection into the limelight. *ACS Infect Dis* 2018;4:239–246.
17. Na SH, Moon DC, Choi M-J, Oh S-J, Jung D-Y *et al*. Detection of oxazolidinone and phenicol resistant enterococcal isolates from duck feces and carcasses. *Int J Food Microbiol* 2019;293:53–59.
18. Cavaco LM, Korsgaard H, Kaas RS, Seyfarth AM, Leekitcharoenphon P *et al*. First detection of linezolid resistance due to the optrA gene in enterococci isolated from food products in Denmark. *J Glob Antimicrob Resist* 2017;9:128–129.
19. Cai J, Schwarz S, Chi D, Wang Z, Zhang R *et al*. Faecal carriage of optrA-positive enterococci in asymptomatic healthy humans in Hangzhou, China. *Clin Microbiol Infect* 2019;25:630.e1–63630.
20. Bender JK, Fleige C, Klare I, Werner G. Development of a multiplex-PCR to simultaneously detect acquired linezolid resistance genes cfr, optrA and poxtA in enterococci of clinical origin. *J Microbiol Methods* 2019;160:101–103.
21. He T, Shen Y, Schwarz S, Cai J, Lv Y *et al*. Genetic environment of the transferable oxazolidinone/phenicol resistance gene optrA in Enterococcus faecalis isolates of human and animal origin. *J Antimicrob Chemother* 2016;71:1466–1473.
22. Chen H, Wang X, Yin Y, Li S, Zhang Y *et al*. Molecular characteristics of oxazolidinone resistance in enterococci from a multicenter study in China. *BMC Microbiol* 2019;19:162.
23. Tan SC, Chong CW, Teh CSJ, Ooi PT, Thong KL. Occurrence of virulent multidrug-resistant *Enterococcus faecalis* and *Enterococcus faecium* in the pigs, farmers and farm environments in Malaysia. *PeerJ* 2018;6:e5353.
33. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
25. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
26. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 2012;3:217–223.
27. Arndt D, Grant JR, Marcu A, Sajed T, Pon A *et al*. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016;44:W16–W21.
28. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011;27:1009–1010.
29. Peterson J, Garges S, Giovanni M, McInnes P, Wang L *et al*. The NIH human microbiome project. *Genome Res* 2009;19:2317–2323.
30. Chen M, Pan H, Lou Y, Wu Z, Zhang J *et al*. Epidemiological characteristics and genetic structure of linezolid-resistant *Enterococcus faecalis*. *Infect Drug Resist* 2018;11:2397–2409.
31. He Q, Hou Q, Wang Y, Li J, Li W *et al*. Comparative genomic analysis of *Enterococcus faecalis*: insights into their environmental adaptations. *BMC Genomics* 2018;19:527.
32. Huang J, Chen L, Wu Z, Wang L. Retrospective analysis of genome sequences revealed the wide dissemination of optrA in Gram-positive bacteria. *J Antimicrob Chemother* 2017;72:614–616.
33. Cavaco LM, Bernal JF, Zankari E, Léon M, Hendriksen RS *et al*. Detection of linezolid resistance due to the optrA gene in *Enterococcus faecalis* from poultry meat from the American continent (Colombia). *J Antimicrob Chemother* 2017;72:678–683.
34. Zhou W, Gao S, Xu H, Zhang Z, Chen F *et al*. Distribution of the optrA gene in Enterococcus isolates at a tertiary care hospital in China. *J Glob Antimicrob Resist* 2019;17:180–186.
35. Lanza VF, Baquero F, de la Cruz F, Coque TM, AcCNET CTM. AcCNET (accessory genome constellation network): comparative genomics software for accessory genome analysis using bipartite networks. *Bioinformatics* 2017;33:283–285.
36. Garsin DA, Frank KL, Silanpää J, Ausubel FM, Hartke A *et al*. Pathogenesis and Models of Enterococcal Infection. In: Gilmore MS, Clewell DB, Ike Y, Shankar N (editors). *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*; 2014.
37. Freitas AR, Elghaieb H, León-Sampedro R, Abbassi MS, Novais C *et al*. Detection of optrA in the African continent (Tunisia) within a mosaic Enterococcus faecalis plasmid from urban wastewaters. *J Antimicrob Chemother* 2017;72:3245–3251.
38. Werner G, Fleige C, Ewert B, Laverde-Gomez JA, Klare I *et al*. High-Level ciprofloxacin resistance among hospital-adapted Enterococcus faecium (CC17). *Int J Antimicrob Agents* 2010;35:119–125.
39. Kim YB, Seo HJ, Seo KW, Jeon HY, Kim DK *et al*. Characteristics of high-level ciprofloxacin-resistant Enterococcus faecalis and Enterococcus faecium from retail chicken meat in Korea. *J Food Prot* 2018;81:1357–1363.
40. Haller C, Berthold M, Wobser D, Kropec A, Lauriola M *et al*. Cell-Wall glycolipid mutations and their effects on virulence of E. faecalis in a rat model of infective endocarditis. *PLoS One* 2014;9:e91863.

41. Nannini EC, Teng F, Singh KV, Murray BE. Decreased virulence of a gls24 mutant of Enterococcus faecalis OG1RF in an experimental endocarditis model. *Infect Immun* 2005;73:7772–7774.
42. Singh KV, La Rosa SL, Somarajan SR, Roh JH, Murray BE. The fibronectin-binding protein EfbA contributes to pathogenesis and protects against infective endocarditis caused by *Enterococcus faecalis*. *Infect Immun* 2015;83:4487–4494.
43. Frank KL, Barnes AMT, Grindle SM, Manias DA, Schlievert PM *et al*. Use of recombinase-based *in vivo* expression technology to characterize Enterococcus faecalis gene expression during infection identifies in vivo-expressed antisense RNAs and implicates the protease Eep in pathogenesis. *Infect Immun* 2012;80:539–549.
44. Xu Y, Singh KV, Qin X, Murray BE, Weinstock GM. Analysis of a gene cluster of *Enterococcus faecalis* involved in polysaccharide biosynthesis. *Infect Immun* 2000;68:815–823.
45. Teng F, Singh KV, Bourgogne A, Zeng J, Murray BE. Further characterization of the EPA gene cluster and EPA polysaccharides of Enterococcus faecalis. *Infect Immun* 2009;77:3759–3767.
46. Thurlow LR, Thomas VC, Hancock LE. Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *J Bacteriol* 2009;191:6203–6210.
47. Chowdhury SA, Nallapareddy SR, Arias CA, Murray BE. The majority of a collection of U.S. endocarditis *Enterococcus faecalis* isolates obtained from 1974 to 2004 lack capsular genes and belong to diverse, non-hospital-associated lineages. *J Clin Microbiol* 2014;52:549–556.
48. Raven KE, Reuter S, Gouliouris T, Reynolds R, Russell JE *et al*. Genome-Based characterization of hospital-adapted Enterococcus faecalis lineages. *Nat Microbiol* 2016;1:15033.
49. Zischka M, Künne CT, Blom J, Wobser D, Sakınç T *et al*. Comprehensive molecular, genomic and phenotypic analysis of a major clone of *Enterococcus faecalis* MLST ST40. *BMC Genomics* 2015;16:175.
50. 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.
51. Tedim AP, Ruiz-Garbajosa P, Corander J, Rodríguez CM, Cantón R *et al*. Population biology of intestinal *Enterococcus* isolates from hospitalized and nonhospitalized individuals in different age groups. *Appl Environ Microbiol* 2015;81:1820–.
52. Matos RC, Lapaque N, Rigottier-Gois L, Debarbieux L, Meylheuc T *et al*. Enterococcus faecalis prophage dynamics and contributions to pathogenic traits. *PLoS Genet* 2013;9:e1003539.
53. Duerkop BA, Palmer KL, Horsburgh MJ. *Enterococcal Bacteriophages and Genome Defense*. Boston, MA: Massachusetts Eye and Ear Infirmary; 2014.
54. Yasmin A, Kenny JG, Shankar J, Darby AC, Hall N *et al*. Comparative genomics and transduction potential of *Enterococcus faecalis* temperate bacteriophages. *J Bacteriol* 2010;192:1122–1130.
55. Cheng S, Xing S, Zhang X, Pei G, An X *et al*. Complete genome sequence of a new *Enterococcus faecalis* bacteriophage, vB_EfaS_IME197. *Genome Announc* 2016;4:e00827–16.
56. Stevens RH, Ektefaie MR, Fouts DE. The annotated complete DNA sequence of *Enterococcus faecalis* bacteriophage ϕEf11 and its comparison with all available phage and predicted prophage genomes. *FEMS Microbiol Lett* 2011;317:9–26.
57. Denes T, Vongkamjan K, Ackermann H-W, Moreno Switt AI, Wiedmann M *et al*. Comparative genomic and morphological analyses of *Listeria phages* isolated from farm environments. *Appl Environ Microbiol* 2014;80:4616–4625.
58. La Rosa SL, Snipen L-G, Murray BE, Willems RJL, Gilmore MS *et al*. A genomic virulence reference map of *Enterococcus faecalis* reveals an important contribution of phage03-like elements in nosocomial genetic lineages to pathogenicity in a Caenorhabditis elegans infection model. *Infect Immun* 2015;83:2156–2167.
59. Lanza VF, Tedim AP, Martínez JL, Baquero F, Coque TM. The plasmidome of firmicutes: impact on the emergence and the spread of resistance to antimicrobials. *Microbiol Spectr* 2015;3:PLAS-0039-2014.
60. Clewell DB, Weaver KE, Dunny GM, Coque TM, Francia MV *et al*. *Extrachromosomal and Mobile Elements in Enterococci: Transmission, Maintenance, and Epidemiology*. Boston: Massachusetts Eye and Ear Infirmary; 2014.
61. Freitas AR, Tedim AP, Francia MV, Jensen LB, Novais C *et al*. Multilevel population genetic analysis of vanA and vanB *Enterococcus faecium* causing nosocomial outbreaks in 27 countries (1986-2012). *J Antimicrob Chemother* 2016;71:3351–3366.
62. Manson JM, Hancock LE, Gilmore MS. Mechanism of chromosomal transfer of *Enterococcus faecalis* pathogenicity island, capsule, antimicrobial resistance, and other traits. *Proc Natl Acad Sci U S A* 2010;107:12269–12274.
63. Morroni G, Brenciani A, Antonelli A, D'Andrea MM, Di Pilato DV*et al*. Characterization of a multiresistance plasmid carrying the *optrA* and *cfr* resistance genes from an *Enterococcus faecium* clinical isolate. *Front Microbiol* 2018;9:2189.
64. Elghaieb H, Tedim AP, Abbassi MS, Novais C, Duarte B *et al*. From farm to fork: identical clones and Tn*6674*-like elements in linezolid-resistant *Enterococcus faecalis* from food-producing animals and retail meat. *J Antimicrob Chemother* 2020;75:30–35.
65. Runyen-Janecky LJ, Hong M, Payne SM. The virulence plasmid-encoded impCAB operon enhances survival and induced mutagenesis in Shigella flexneri after exposure to UV radiation. *Infect Immun* 1999;67:1415–1423.
66. Sletvold H, Johnsen PJ, Hamre I, Simonsen GS, Sundsfjord A *et al*. Complete sequence of Enterococcus faecium pVEF3 and the detection of an omega-epsilon-zeta toxin-antitoxin module and an ABC transporter. *Plasmid* 2008;60:75–85.
67. Phillips S, Novick RP. Tn554--a site-specific repressor-controlled transposon in Staphylococcus aureus. *Nature* 1979;278:476–478.
67. Li D, Li X-Y, Schwarz S, Yang M, Zhang S-M *et al*. Tn*6674* a novel enterococcal *optrA*-carrying multiresistance Transposon of the Tn*554* faFamily. *Antimicrob Agents Chemother* 2019;63:e00809–00819.
69. Murphy E, Phillips S, Edelman I, Novick RP. Tn554: isolation and characterization of plasmid insertions. *Plasmid* 1981;5:292–305.
70. Murphy E, Reinheimer E, Huwyler L. Mutational analysis of att554, the target of the site-specific transposon Tn554. *Plasmid* 1991;26:20–29.
71. Oliveira PH, Touchon M, Cury J, Rocha EPC. The chromosomal organization of horizontal gene transfer in bacteria. *Nat Commun* 2017;8:841.
72. Chen J, Quiles-Puchalt N, Chiang YN, Bacigalupe R, Fillol-Salom A *et al*. Genome hypermobility by lateral transduction. *Science* 2018;362:207–212.
73. Du F, Lv X, Duan D, Wang L, Huang J. Characterization of a Linezolid- and Vancomycin-Resistant *Streptococcus suis* isolate that harbors *optrA* and *vanG* Operons. *Front Microbiol* 2019;10:10.
74. Shang Y, Li D, Hao W, Schwarz S, Shan X *et al*. A prophage and two ICESa2603-family integrative and conjugative elements (ICEs) carrying optrA in Streptococcus suis. *J Antimicrob Chemother* 2019;74:2876–2879.
75. Wang M, Xiong W, Liu P, Xie X, Zeng J *et al*. Metagenomic insights into the contribution of phages to antibiotic resistance in water samples related to swine feedlot wastewater treatment. *Front Microbiol* 2018;9:2474.
76. Lekunberri I, Subirats J, Borrego CM, Balcázar JL. Exploring the contribution of bacteriophages to antibiotic resistance. *Environ Pollut* 2017;220:981–984.