Complete genome sequence and description of *Lactococcus garvieae* M14 isolated from Algerian fermented milk

M. Moumene[1,2], F. Drissi[1], O. Croce[1], B. Djebbari[1], C. Robert[1], E. Angelakis[1], D. E. Benouareth[2], D. Raoult[1,3] and V. Merhej[1]

1) *Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095, Institut Hospitalo-Universitaire, Marseille, France,* 2) *University 08 May 1945, Guelma, Algeria and* 3) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

We describe using a polyphasic approach that combines proteomic by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) analysis, genomic data and phenotypic characterization the features of *Lactococcus garvieae* strain M14 newly isolated from the fermented milk (known as raib) of an Algerian cow. The 2 188 835 bp containing genome sequence displays a metabolic capacity to form acid fermentation that is very useful for industrial applications and encodes for two bacteriocins responsible for its eventual bioprotective properties.

Keywords: Fermented milk, genome, lactic acid bacteria, *Lactococcus garvieae*, sequencing
Original Submission: 20 November 2015; **Revised Submission:** 28 December 2015; **Accepted:** 14 January 2016
Article published online: 22 January 2016

Corresponding author: V. Merhej, Aix-Marseille Université, Faculté de Médecine, 27 Bd Jean Moulin, 13385 Marseille Cedex 5, France
E-mail: vicky_merhej@univ-amu.fr
M. Moumene and F. Drissi contributed equally to this article, and both should be considered first author.

Introduction

Lactococcus garvieae is a lactic acid bacteria (LAB) that has commonly been used in the manufacture of many varieties of cheese and other fermented milk products [1,2] as well as meat products [3]. The ability of some LAB to produce proteins with bactericidal properties called bacteriocins led to their potential utilization as biopreservatives in the food industry against a range of pathogenic bacteria, including *Listeria* sp. and *Clostridium* sp. [4,5]. Also, because of bacteriocins, some LAB are thought to act as bioprotective organisms that play a major role in the composition of the microbiota [6]. First isolated from cases of bovine mastitis [7], *L. garvieae* has gained recognition as a potential pathogen of various fish species, including rainbow trout [8]. Moreover, *L. garvieae* has been involved in many clinical cases including infective endocarditis associated with septicaemia, spondylodiscitis and urinary and skin infections [9–15]. Genomic interspecies microarray hybridization and pan-genome comparative analysis of the pathogenic strain Lg2 and the nonvirulent strain YT-3 identified genes encoding for host colonization and the development of pathogenesis including a capsule gene cluster and genes encoding for a myosin cross-reactive antigen and haemolysin [16,17]. The phenotypic diversity of *L. garvieae* seems to be related to the specific animal host they colonize [18,19]. Altogether, the analysis of *L. garvieae* genomes isolated from a dairy product and its comparison with pathogenic isolates seems to be necessary to clarify the global genetic variations that may justify, at least in part, the observed phenotypic differences.

In this work, we have isolated and identified a new strain of *Lactococcus garvieae* from the fermented milk product of an Algerian cow, known as raib, as a part of the study of LAB and revealed their antibacterial activity. A total of 47 different bacterial species including *Lactococcus* and *Streptococcus* spp. as identified by API 50CHL and mass spectrometry were isolated from milk product specimens (unpublished data). *Lactococcus* spp. was the only species with antibacterial effect as shown by the agar well-diffusion assay. We accomplished deep studies

http://dx.doi.org/10.1016/j.nmni.2016.01.009

including phenotypic, polyphasic taxonomy, genotypic and phylogenetic analyses.

Here we provide a set of features for the identified *Lactococcus garvieae* strain M14, together with the description of the complete genome sequence and annotation. We present a comparative genomic analysis of all available sequences of closely related species to *L. garvieae*. This integrative approach dealing with a large data set has the potential to explore the relationship between the presence of *L. garvieae* in dairy and food safety in order to recognize and prevent a potential hazard for consumers.

Materials and Methods

Sample collection

Raw cow's milk samples were collected from a farm in Guelma, in the east of Algeria, in sterile glass bottles and transported in an isotherm container to the laboratory. A total of 200 mL of milk samples was allowed to clot at room temperature to promote the development of endogenous lactic flora according to Zadi-Karam and Karam [20].

Strain isolation

M17 agar media (Sigma-Aldrich, St. Louis, MO, USA) prepared in accordance with manufacturer's instructions was used for the growth of the LAB strain screened in this investigation. Thus, 0.1 mL of fermented milk sample was plated onto M17 to promote the bacterial flora cultivation in aerobic conditions by incubation at 37°C for 24 hours. *Lactococcus garvieae* strain M14 was then isolated and stored at −20°C until further use.

Phenotypic, genotypic and phylogenetic analyses

We used 16S ribosomal RNA gene sequencing (16S rRNA) to provide genus and species identification for the isolate [21] and taxonomic classification of strain M14. A comparison of nucleotides query sequences against the nucleotide sequence database was also performed using BLAST (Basic Local Alignment Search Tool). The 16S rRNA sequences of all *Lactococcus* strains with draft genome were searched within the scaffolds using the RNAmmer server [22]. The phylogenetic tree, based on almost complete 16S rRNA gene sequences with a minimum length of 1517 nucleotides, was reconstructed using distance matrix (neighbour joining) within the MEGA 5 software [23]. Sequences were aligned using Clustal X 2.0 [24].

Different temperatures (room temperature, 28, 37 and 45°C) were tested to determine the growth temperature range and the optimal growth temperature for the strain. Growth of the strain was tested on 5% sheep's blood agar under anaerobic and microaerophilic conditions using the GENbag anaer and GENbag microaerosystems respectively (bioMérieux, Marcy l'Etoile, France) and in aerobic conditions, with or without 5% CO_2.

L. garvieae strain M14 morphology was characterized by transmission electron microscopy (TEM) using a Morgani 268D (Philips, Amsterdam, The Netherlands) spectrometer with operating voltage of 60 kV. Polyphasic taxonomic identification by manufactured kits is widely used; the API 50CH carbohydrate fermentation strips and API ZYM enzyme test system (bioMérieux) were used to determine the biochemical profile of strain M14 according to the manufacturer's instructions.

Protein mass spectroscopy analysis was carried out by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) as previously described [25] using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). Twelve distinct deposits were made for strain M14 from 12 isolated colonies.

The 12 collected spectra for M14 were imported into the MALDI BioTyper 2.0 software (Bruker) and analysed by standard pattern matching (with default parameter settings) against 7.289 bacterial spectra including 26 spectra from three *L. garvieae* species, used as reference data, in the BioTyper database. Interpretation of scores as established by Bruker was as follows: a score >1.9 to a validly published species enabled the identification at the species level, a score >1.7 but <1.9 enabled the identification at the genus level and a score <1.7 did not enable any identification.

Growth conditions and genomic DNA preparation

L. garvieae was grown on 5% sheep's blood–enriched Columbia agar (bioMérieux) at 37°C in aerobic atmosphere. Bacteria grown on three petri dishes were collected and resuspended in 4 × 100 µL of Tris-EDTA (TE) buffer. Then 200 µL of this suspension was diluted in 1 mL TE buffer for lysis treatment. After a lysozyme incubation of 30 minutes at 37°C, lysis was performed with lauryl sarcosyl by 1% final and RNAse A treatment at 50 µG/µL final concentration during 1 hour at 37°C, followed by an overnight proteinase K incubation at 37°C. Extracted DNA was then purified using three successive phenol–chloroform extractions and ethanol precipitation at −20°C overnight. After centrifugation, the DNA was resuspended in 70 µL TE buffer. The yield and concentration were measured by the Quant-it Picogreen kit (Invitrogen; Life Technologies, Carlsbad, CA, USA) on the Genios-Tecan fluorometer at 113 ng/µL.

Genome sequencing and assembly

Genomic DNA (gDNA) of *L. garvieae* was sequenced using MiSeq Technology sequencer (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded in order

to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The gDNA was quantified by a Qubit assay with the high sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 40.8/μL. The mate pair library was prepared with 1 μg of genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The profile of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 lab chip. The DNA fragments exhibited a mean of 4.5 kb (4486 pb). No size selection was performed, and only 308.9 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal size of 652 bp on the Covaris device S2 in microtubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies). The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 10 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 42-hour run in a 2 × 251 bp read length. Total information of 8.6 Gb was obtained from a 950K/mm^2 cluster density with a cluster passing quality control filters of 93.12% (18 182 000 clusters). Within this run, the index representation for *L. garvieae* was determined to 9.73%. The whole genome shotgun strategy using Illumina sequencing technology gave 3 294 808 reads. Illumina reads were trimmed using Trimmomatic [26], then assembled with Spades software [27,28]. Contigs obtained were combined together by SSpace [29] and Opera software 1.2 [30] helped by GapFiller 1.10 [31] to reduce the set. Some manual refinements using CLC Genomics 7 software (CLC Bio, Aarhus, Denmark) and homemade tools in Python improved the genome.

Genome annotation

Noncoding genes and miscellaneous features were predicted using RNAmmer [22], ARAGORN [32], Rfam [33], PFAM [34] and Infernal [35]. Coding DNA sequences were predicted using Prodigal [36], and functional annotation was achieved using BLAST+ [37] and HMMER3 [38] against the UniProtKB database [39]. The KEGG orthology [40] annotation was done by the KAAS online server [41] using the SBH method. The pathways in which each gene might be involved were derived from the best KO hit. Gene functions were assigned by the Clusters of Orthologous Groups (COGs) database [42,43]. The bacteriocin database of the Unité des Maladies Infectieuses et Tropicales Emergentes (URMITE), known as the BUR database, was used to annotate genes encoding for bacteriocins [44].

Results and Discussion

Organism classification and features

The sequenced 16S rRNA gene of strain M14 was deposited in the GenBank database under the accession number LK985397. A BLAST search against the nucleotide database showed that strain M14 was most closely related to *Lactococcus* species, with a gene sequence identity value of 99.7% with *Lactococcus garvieae* YT-3. On the basis of the comparative sequence analysis of 16S rRNA gene sequence, strain M14 belongs to the already described species *L. garvieae* [45,46].

A neighbour-joining phylogenetic tree, based on almost-complete 16S rRNA gene sequences of *Lactococcus garvieae* M14 strain and closely related species, is shown in Fig. 1. Sequences of the closest species including *Lactococcus garvieae* strains and strains of *Lactococcus lactis* subsp. *lactis* (NR_103918), *Lactococcus lactis* subsp. *cremoris* (NR_074949), *Lactobacillus rhamnosus* GG (NR_102778), *Lactobacillus sakei* subsp. *sakei* 23K (NR_075042), *Lactobacillus plantarum* WCFS1 (NR_075041), *Lactobacillus fermentum* IFO 3956 (NR_075033), *Lactobacillus salivarius* UCC118 (NR_074589) and *Lactobacillus ruminis* ATCC 27782 (NR_102839) were aligned with the 16S rRNA gene sequence of the strain M14. The strain M14 formed together with *Lactococcus* garvieae strains, and a common lineage with the *L. lactis* species was supported by a high bootstrap value of 99% (Fig. 1).

Growth occurred for all of the tested temperatures, but optimal growth was observed at 37°C. The colonies were 1 to 6 mm in diameter and moderately opaque in facultative anaerobic conditions on enriched-blood Columbia agar (bioMérieux) and appeared whitish in colour at 28°C. The motility test was negative. Gram staining showed Gram-positive non-sporulating cocci. Cells grown on agar range in length from 0.79 to 0.93 μm (mean, 0.86 μm) and diameter from 0.59 to 0.63 μm (mean, 0.61 μm) as determined by negative staining TEM micrograph.

Strain M14 did neither have catalase nor oxidase activity. Using the API 50CH system, a positive reaction was observed for D-ribose, D-glucose, D-fructose, D-mannose, D-galactose, D-mannitol, amygdalin, arbutin, N-acetylglucosamine, esculin, salicin, D-cellobiose, D-lactose, D-saccharose and D-trehalose. Negative reactions were observed for glycerol, erythritol, D-arabinose, L-arabinose, D-xylose, L-xylose, D-adonitol, methyl-βD-xylopyranoside, L-sorbose, L-rhamnose, dulcitol, inositol, D-sorbitol, methyl-αD-mannopyranoside, methyl-αD-glucopyranoside, D-maltose, D-melibiose, inulin, D-melezitose, D-raffinose, starch, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol and gluconate. Using the API ZYM system, negative reactions were observed

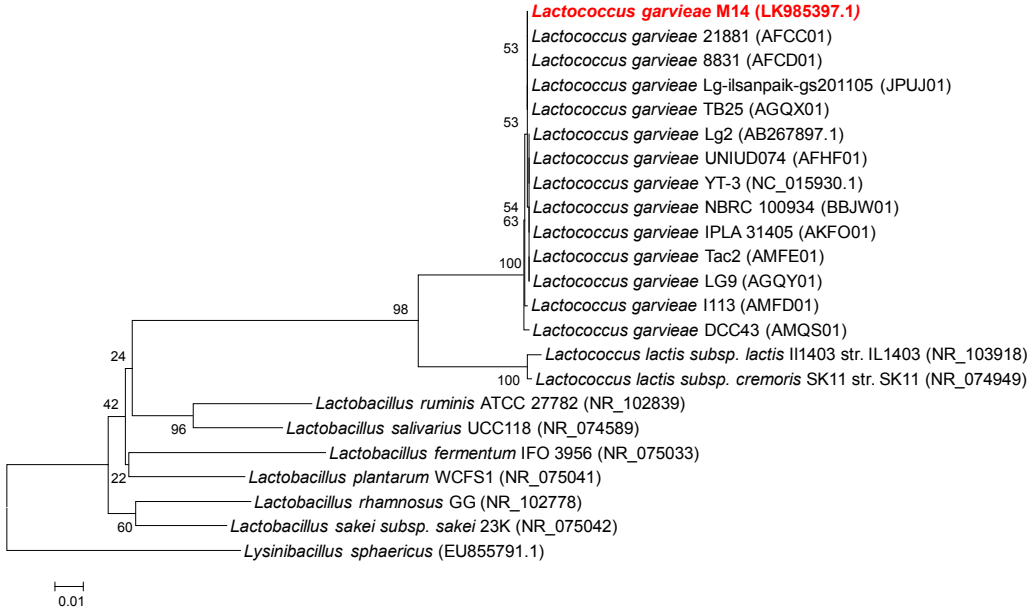

FIG. 1. Phylogenetic tree highlighting position of *Lactococcus garvieae* strain M14 (LK985397) relative to other phylogenetically close strains within genus *Lactococcus* and *Lactobacillus*, with *Lysinibacillus sphaericus* as outgroup. Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 500 times to generate majority consensus tree. Scale bar = 2% nucleotide sequence divergence.

for alkaline phosphatase, cystine arylamidase (proteases), trypsin, α-galactosidase (melibiase), β-glucosidase (cellulase), α-mannosidase and α-fucosidase, and positive reactions were observed for esterase, esterase lipase, lipase, leucine and valine arylamidase, α-chemotrypsin, acid phosphatase, β-galactosidase, β-glucuronidase and α- and β-glucosidase. The urease reaction, nitrate reduction and indole production were negative. When compared to the phylogenetically close species from *Lactococcus* and *Lactobacillus* [17,47–53], *L. garvieae* strain M14 exhibited the phenotypic differences detailed in Table 1. *L. garvieae* was susceptible to amoxicillin, imipenem, piperacillin, ciprofloxacin, ceftriaxone, erythromycin, vancomycin, nitrofurantoin, nitrofurantoin, metronidazole and rifampicin but resistant to cefoxitin and cotrimoxazole.

TABLE 1. Differential phenotypic characteristics between *Lactococcus garvieae* strain M14 and phylogenetically close species[a]

Characteristic	*L. garvieae* M14 DSM 29394	*L. garvieae* YT-3 DSM 6783	*L. lactis* subsp. *lactis* DSM 20481	*L. lactis* subsp. *cremoris* DSM 20069	*L. rhamnosus* DSM 20021	*L. sakei* subsp. *sakei* DSM 20017	*L. plantarum* DSM 20174	*L. fermentum* DSM 20052	*L salivarius* DSM 20555
Gram stain	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Cell shape	Cocci	Cocci	Rod	Rod	Rod	Rod	Rod	Rod	Rod
L-Arabinose	−	−	−	−	−	−	+	−	−
D-Ribose	+	+	+	−	+	+	+	+	−
D-Glucose	+	+	+	+	+	+	+	+	+
D-Galactose	+	+	+	+	+	+	+	+	+
D-Mannitol	+	+	+	+	+	−	+	−	+
Amygdalin	+	+	+	−	+	−	+	−	−
Arbutin	+	+	+	−	+	−	+	−	−
Esculin	+	+	+	+	+	+	+	+	+
D-Cellobiose	+	+	+	+	+	−	+	+	−
D-Lactose	+	−	+	−	+	−	+	+	+
Inulin	−	−	+	−	−	−	+	−	−
D-Melezitose	−	−	−	−	+	−	+	−	−
Glycogen	−	−	−	−	−	−	−	−	−
Alkaline phosphatase	−	−	−	−	+	−	−	−	−
Acid phosphatase	+	+	−	+	+	−	−	−	+
α-Glucosidase	+	+	+	−	+	−	−	+	+
N-acetyl-β-glucosaminidase	−	−	+	−	+	−	−	+	−
α-Mannosidase	−	−	+	−	−	−	+	+	−

DSMZ, Deutsche Sammlung von Mikroorganismen.
[a]The type strains of related lactic acid bacteria species were obtained from the DSMZ culture collection (Braunschweig, Germany). All strains were cultured according to recommendations given in the DSMZ catalogue of strains.

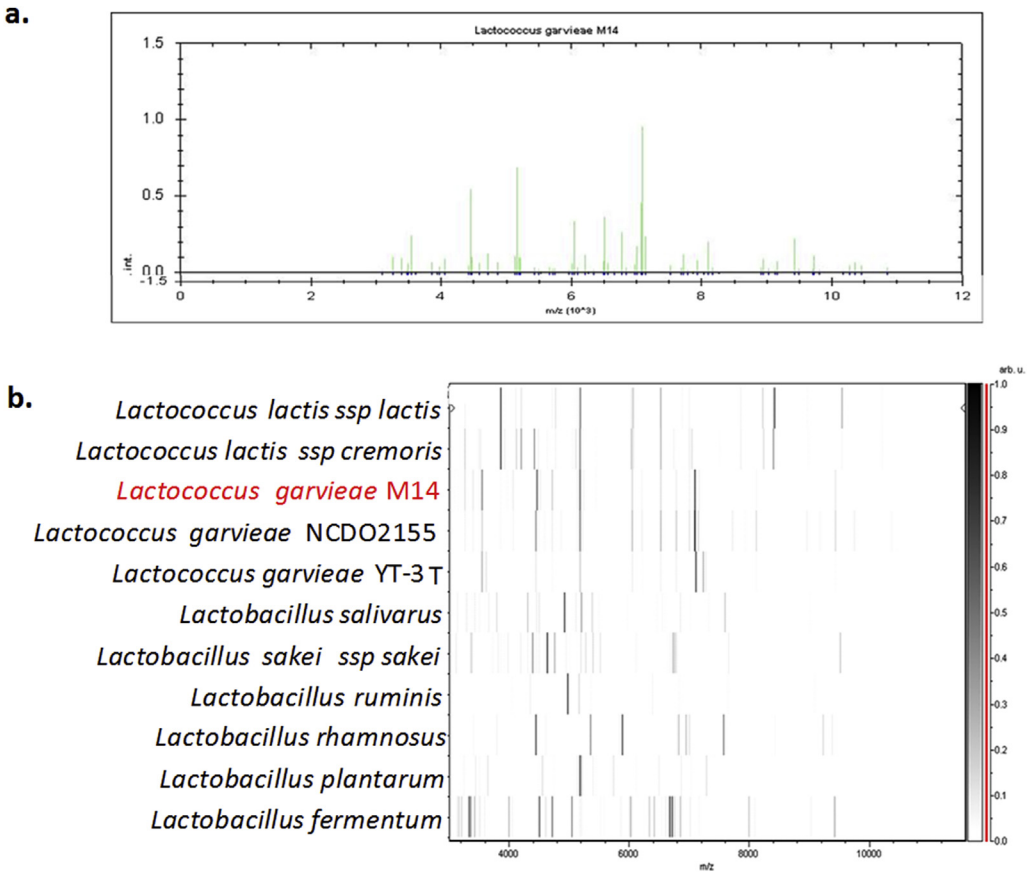

FIG. 2. (a) Reference mass spectrum from *Lactococcus garvieae* strain M14 and (b) gel view comparing *L. garvieae* M14 spectra with *Lactococcus* species (*Lactococcus lactis* subsp. *lactis*, *Lactococcus lactis* subsp. *cremoris* and two strains of *Lactococcus garvieae*) and with *Lactobacillus* species (*Lactobacillus sakei* subsp. *sakei*, *Lactobacillus salivarus*, *Lactobacillus ruminis*, *Lactobacillus rhamnosus*, *Lactobacillus plantarum* and *Lactobacillus fermentum*).

Extended features descriptions

MALDI-TOF analysis results of strain M14 showed scores ranging from 2.177 to 2.343 with *Lactococcus* spp., suggesting that our isolate was a member of *Lactococcus* species yet not a known strain. We incremented our database with the spectrum of strain M14 (Fig. 2a). Spectrum differences with others of phylogenetically close species are shown in Fig. 2b. The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel-like look. The *x*-axis records the *m*/*z* value. The left *y*-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a greyscale scheme code. The colour bar and the right *y*-axis indicate the relation between the colour a peak is displayed with and the peak intensity in arbitrary units. The compared species are indicated on the left.

Genome project history

Lactococcus garvieae is commonly used in the industry of dairy products manufactured from raw milk. Moreover, the M14 strain has shown bactericidal effects against several bacteria (unpublished data) which indicate the eventual role played by *L. garvieae* in the composition of the gut microbiota. The sequencing of this strain is part of a study of the human digestive flora aiming at describing the force balance that shapes its composition, including antibacterial potency. Indeed, *L. garvieae* strain M14 was the 46th genome from the genus *Lactococcus* and the 16th genome of *L. garvieae* sp. The European Molecular Biology Laboratory (EMBL) accession number

TABLE 2. **Project information**

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	1 mate paired
MIGS-29	Sequencing platforms	MiSeq Illumina
MIGS-31.2	Sequencing coverage	110
MIGS-30	Assemblers	Spades
MIGS-32	Gene calling method	Prodigal
	GenBank ID	CCXC01000001–CCXC01000013
	GenBank Date of Release	October 2014
MIGS-13	Source material identifier	M14
	Project relevance	Potential probiotic and biopreservative

MIGS, minimum information about a genome sequence.

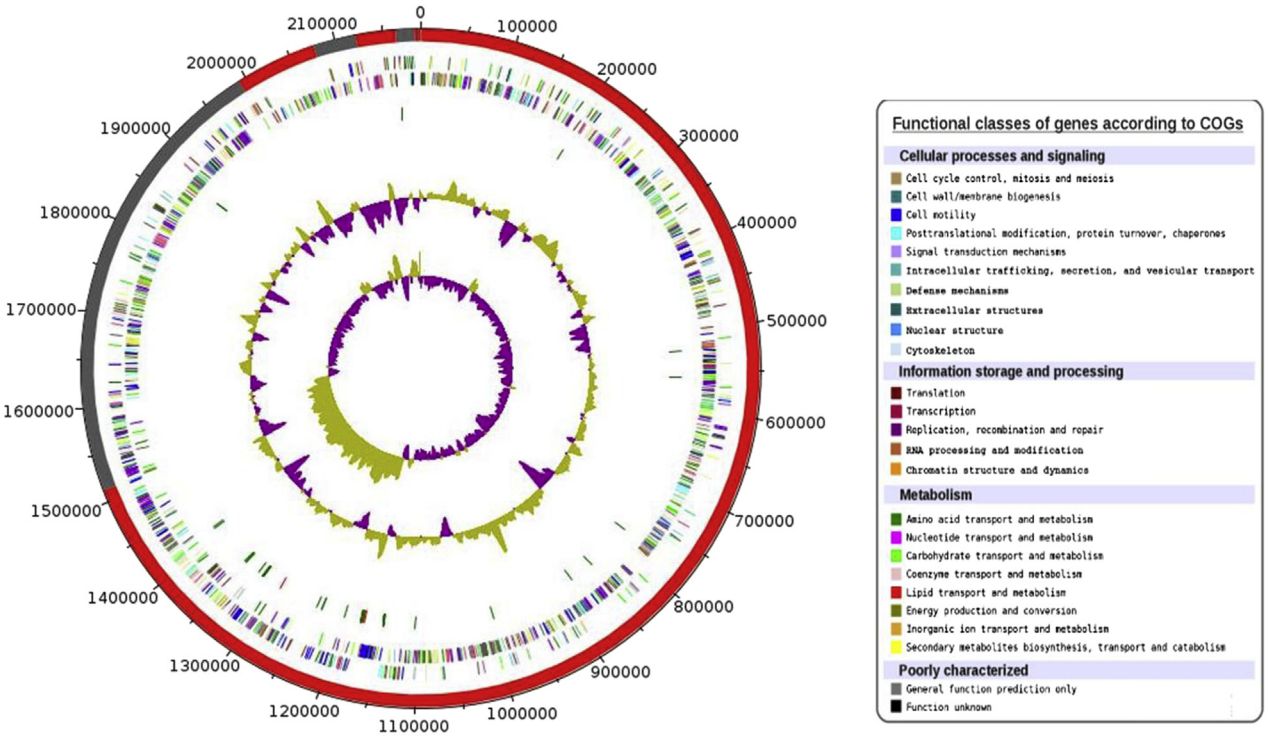

FIG. 3. Circular representation of *Lactococcus garvieae* strain M14 genome. Circles from outside to center: Contigs (red/grey), genes colored according to categories determined in COGs on forward and reverse strands (two circles), rRNAs red, tRNA green, GC content and GC skew (green/purple).

is CCXC01000001–CCXC01000013 and consists of 13 contigs without gaps, including four contigs that have been assigned to four plasmids (Table 2).

Genome properties

The draft genome of *L. garvieae* M14 consists of 13 contigs of sizes ranging between 889 and 1 512 971 bp. The genome of M14 is composed of a single linear chromosome (2 188 835 bp; 37.69% G+C content) and four plasmids ranging in size from 42 306 to 1095 bp, including one circular plasmid (Fig. 3, Tables 3 and 4). The chromosome contains 91 predicted RNA including five rRNA (one 16S, one 23S and three 5S), 45 tRNA, one tmRNA and 40 miscellaneous RNA and 2214 protein-coding genes which represent 1 934 957 (88.40% of the genome). A total of 1515 genes (68.42%) were assigned a putative function (by COGs) [42,43]. We found that 23.63% of the genes encode for information storage and processing (J, A, K, L and B categories), 21.11% were involved in cellular processes and signaling (D, V, T, M, N, U, O and X categories), 40.29% participate in metabolism (C, G, E, F, H, I, P and Q categories) and 14.97% were poorly characterized (R and S categories). The distribution of genes into COGs functional categories is presented in Fig. 4.

Genome comparison

When comparing *L. garvieae* M14 with nine *Lactobacillus* species and two *Lactococcus* strains that have similar 16S rRNA sequences, we found that the genome sequence of *L. garvieae* M14

TABLE 3. Nucleotide content and gene count levels of genome

Attribute	Value	% of total[a]
Genome size (bp)	2 188 835	100.00
DNA coding region (bp)	1 934 957	88.40
DNA G+C content (bp)	827 233	37.79
Total genes	2264	100.00
rRNA	5	0.21
tRNA	45	1.90
tmRNA	1	0.04
miscRNA	40	1.69
Protein-coding genes	2214	97.79
Genes with function prediction	1651	72.92
Genes assigned to COGs	1515	68.42

COGs, Clusters of Orthologous Groups database.
[a]Total is based on either size of genome (bp) or total number of protein-coding genes in annotated genome.

TABLE 4. Nucleotide content and gene count levels of plasmids

Attribute	Value	% of total[a]
Size (bp)	64 869 (42 306; 16 485; 4983; 1095)[b]	100
DNA coding region (bp)	55 608	85.72
DNA G+C content (bp)	22 256	34.31
Total genes	75	100
rRNA	0	0
Protein-coding genes	75	100
Genes with function prediction	20	35.09
Genes assigned to COGs	10	17.54

COGs, Clusters of Orthologous Groups database.
[a]Total is based on either size of plasmids (bp) or total number of protein-coding genes in annotated sequences.
[b]Sizes are indicated in parentheses.

is smaller than those of *Lactobacillus plantarum* WCFS1, *Lactobacillus rhamnosus* GG, *Lactococcus lactis* subsp. *cremoris* SK11 and *Lactococcus lactis* subsp. *lactis* II1403 (3.35, 3.01, 2.60 and 2.37 MB respectively), but larger than those of *Lactobacillus salivarius* UCC118, *Lactobacillus fermentum* IFO 3956, *Lactobacillus ruminis* ATCC 27782, *L. garvieae* Lg2, *L. garvieae* YT-3, and *Lactobacillus sakei* subsp. *sakei* 23K (2.13, 2.10, 2.07, 1.96, 1.95 and 1.88 Mb respectively) (Table 5). The G+C content of *L. garvieae* M14 is smaller than those of *L. fermentum* IFO 3956, *L. rhamnosus* GG, *L. plantarum* WCFS1, *L. ruminis* ATCC 27782, *L. sakei* subsp. *sakei* 23K, *L. garvieae* YT-3 and *L. garvieae* Lg2 (51.47, 46.69, 44.42, 43.47, 41.26, 38.83 and 38.76% respectively), but larger than those of *L. lactis* subsp. *cremoris* SK11, *L. lactis* subsp. *lactis* II1403 and *L. salivarius* UCC118 (35.78, 35.33 and 33.04% respectively) (Table 5). The gene content of *L. garvieae* M14 is smaller than those of *L. plantarum* WCFS1, *L. rhamnosus* GG and *L. lactis* subsp. *cremoris* SK11 and *L. lactis* subsp. *lactis* II1403 (3063, 2944, 2504 and 2277 respectively), but larger than those of *L. salivarius* UCC118, *L. sakei* subsp. *sakei* 23K, *L. ruminis* ATCC 27782 and *L. fermentum* IFO 3956 (2014, 1885, 1862 and 1843 respectively) (Table 5). The proportion of gene count (in percentage) related to each COGs category was similar among the studied strains of *L. garvieae*. However, the distribution of genes into COGs category was not entirely similar in all the compared genomes (Fig. 4). *L. garvieae* M14 have an important number of genes participating in carbohydrate transport and metabolism (175 genes), yet less important than those of *L. plantarum* WCFS1 and *L. rhamnosus* GG (267 and 263 genes respectively). *L. garvieae* M14 and *L. rhamnosus* GG have the highest number of genes involved in defence mechanisms (51 and 63 genes respectively), compared

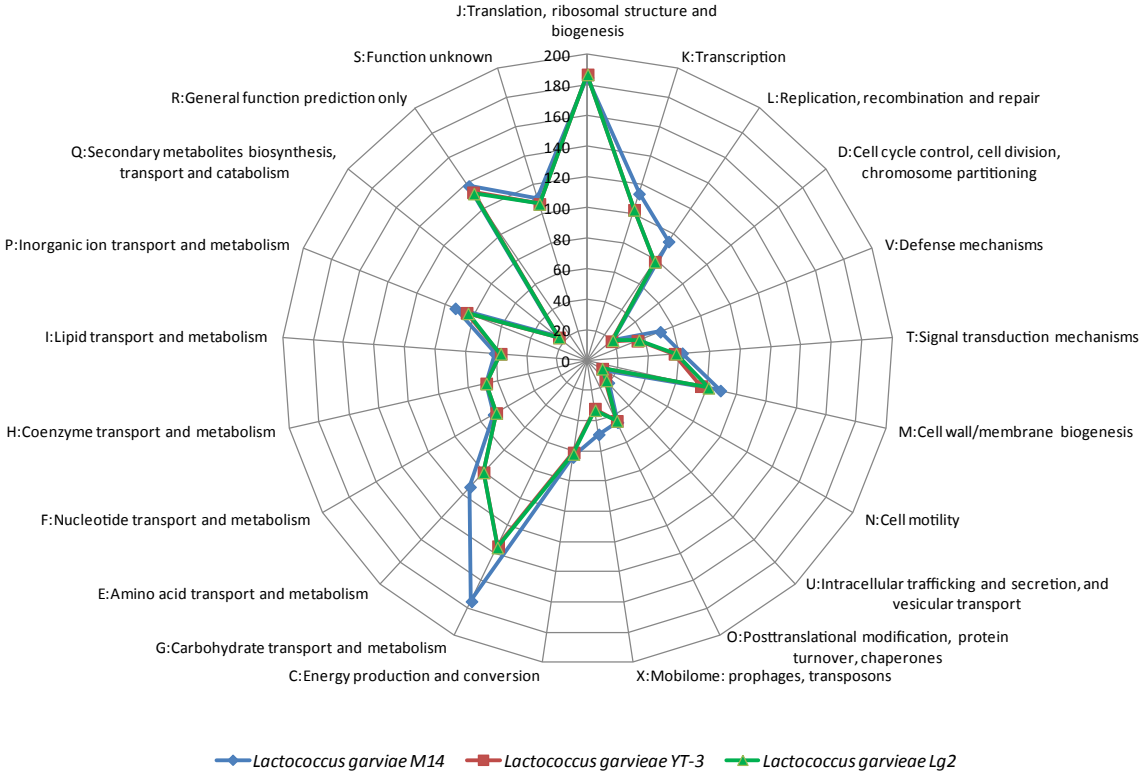

FIG. 4. Functional classification of genes encoded by *Lactococcus garvieae* M14 and its comparison with *Lactococcus garvieae* YT-3 and *Lactococcus garvieae* Lg2. Protein-coding sequences are classified according to COGs categories. COGs, Clusters of Orthologous Groups (COGs) database.

TABLE 5. General genome features

Strain	Size (Mb)	G+C%	Gene content
Lactobacillus plantarum WCFS1	3.35	44.42	3063
Lactobacillus rhamnosus GG	3.01	46.69	2944
Lactococcus lactis subsp. *cremoris* SK11	2.60	35.78	2504
Lactococcus lactis subsp. *lactis* II1403	2.37	35.33	2277
Lactococcus garvieae M14	2.19	37.79	2264
Lactobacillus salivarius UCC118	2.13	33.04	2014
Lactobacillus fermentum IFO	2.10	51.47	1843
Lactobacillus ruminis ATCC 27782	2.07	43.47	1862
Lactococcus garvieae Lg2	1.96	38.76	1968
Lactococcus garvieae YT-3	1.95	38.83	1947
Lactobacillus sakei subsp. *sakei* 23K	1.88	41.26	1885

to the other analysed genomes (37 genes in average). Unlike *L. garvieae* YT-3, *L. garvieae* strain M14 possesses plasmids the sequences of which were closely related to *L. garvieae* strain 21881 plasmid pGL5, *L. garvieae* strain IPLA31405 plasmid pIG42, *L. lactis* plasmid pSRQ900 and *L. lactis* strain MJC15 plasmid pCD4 sequences. The genes of the plasmids encode for a type IV secretory pathway and for proteins with hypothetical functions.

Energy metabolism and transporters

The coding DNA sequences annotated by the COGs database revealed that as much as 10.5% of the genomes corresponded to genes involved in carbohydrate transport and metabolism. Like all obligately homofermentative strains, *L. garvieae* M14 was found to possess the fructose bisphosphate aldolase (EC 4.1.2.13) in its genome, which is a key enzyme of the glycolysis pathway, whereas it lacks the phosphoketolase enzyme (EC 4.1.2.9) of the pentose phosphate pathway, present only in heterofermentative bacteria genomes. All the genes required for the degradation of the glucose to pyruvate are present in the genome, as well as the lactate dehydrogenase gene which allows the conversion of pyruvate into lactic acid. Several enzymes acting on pyruvate, including α-acetolactate synthase, pyruvate-formate lyase, lactate dehydrogenase and pyruvate oxidase, have also been identified in the strain M14 genome. Further, genome examination indicates that some enzymes needed for the full citrate cycle and for the gluconeogenesis are missing. The phosphotransferase system (PTS) for fructose, galactose, mannose, maltose, lactose, sucrose, trehalose, mannitol and cellobiose were present in the genome, while PTS for xylose, gluconate and ribose were absent. On the basis of its metabolic profile, *L. garvieae* M14 produces primarily lactic acid from hexoses using glycolysis. This ability of homolactic fermentation is useful for industrial applications.

Defense mechanism

We identified in the genome of *L. garvieae* M14 several phage-related genes and 51 proteins involved in defence features including a glycopeptide antibiotics resistance protein and two bacteriocins that are localized in the chromosome. The first bacteriocin has a length of 64 amino acids, and the use of the BUR database allowed us to identify a very similar bacteriocin sequence in the genomes of *L. garvieae* strains YT-3, Lg2 and TRF1. These sequences have been previously annotated as encoding for hypothetical proteins in the genomes of *L. garvieae* strains YT-3 and Lg2. The second bacteriocin has a length of 184 amino acids and corresponds to a colicin V, also found in the genome of the strain Lg2. Garviecin L1-5 was the first bacteriocin detected in a *Lactococcus garvieae* strain [54]. It has been shown to inhibit the growth of species relatively closely related to the producer but also of the human pathogen *Listeria monocytogenes*. Altogether, the production of bacteriocins gives *L. garvieae* strains a competitive advantage within their environment, allowing them to directly inhibit other bacteria and proliferate.

Conclusions

We have presented the phenotypic, phylogenetic and genomic analyses that allowed the description of *Lactococcus garvieae* strain M14. This new bacterial strain, isolated from a fermented milk sample from an Algerian cow, is essential in the manufacture of dairy products and seems to play a major role as a biopreservative in the food industry. If the infectious risk is definitively ruled out, it may be a potential probiotic.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process. VM was supported by a Chairs of Excellence program from the Centre National de la Recherche Scientifique (CNRS). This study was funded by the Fondation Méditerranée Infection.

Conflict of Interest

None declared.

References

[1] El-Baradei G, Delacroix-Buchet A, Ogier JC. Biodiversity of bacterial ecosystems in traditional Egyptian Domiati cheese. Appl Environ Microbiol 2007;73:1248–55.

[2] El-Baradei G, Delacroix-Buchet A, Ogier JC. Bacterial biodiversity of traditional Zabady fermented milk. Int J Food Microbiol 2008;121:295–301.
[3] Aquilanti L, Santarelli S, Silvestri G, Osimani A, Petruzzelli A, Clementi F. The microbial ecology of a typical Italian salami during its natural fermentation. Int J Food Microbiol 2007;120:136–45.
[4] Schoster A, Kokotovic B, Permin A, Pedersen PD, Bello FD, Guardabassi L. *In vitro* inhibition of *Clostridium difficile* and *Clostridium perfringens* by commercial probiotic strains. Anaerobe 2013;20:36–41.
[5] Monteagudo-Mera A, Rodríguez-Aparicio L, Rúa J, Martínez-Blanco H, Navasa N, García-Armesto MR, et al. *In vitro* evaluation of physiological probiotic properties of different lactic acid bacteria strains of dairy and human origin. J Funct Foods 2012;4:531–41.
[6] Bernbom N, Licht TR, Brogren C-H, Jelle B, Johansen AH, Badiola I, et al. Effects of *Lactococcus lactis* on composition of intestinal microbiota: role of nisin. Appl Environ Microbiol 2006;72:239–44.
[7] Collins MD, Farrow JAE, Phillips BA, Kandler O. *Streptococcus garvieae* sp. nov. and *Streptococcus plantarum* sp. nov. J Gen Microbiol 1983;129:3427–31.
[8] Vendrell D, Balcázar JL, Ruiz-Zarzuela I, de Blas I, Gironés O, Múzquiz JL. *Lactococcus garvieae* in fish: a review. Comp Immunol Microbiol Infect Diss 2006;29:177–98.
[9] Fihman V, Raskine L, Barrou Z, Kiffel C, Riahi J, Bercot B, et al. *Lactococcus garvieae* endocarditis: identification by 16S rRNA and sodA sequence analysis. J Infect 2006;52:e3–6.
[10] Chan JF, Woo PC, Teng JL, Lau SK, Leung SS, Tam FC, et al. Primary infective spondylodiscitis caused by *Lactococcus garvieae* and a review of human *L. garvieae* infections. Infection 2011;39:259–64.
[11] Wang CY, Shie HS, Chen SC, Huang JP, Hsieh IC, Wen MS, et al. *Lactococcus garvieae* infections in humans: possible association with aquaculture outbreaks. Int J Clin Pract 2007;61:68–73.
[12] Vinh DC, Nichol KA, Rand F, Embil JM. Native-valve bacterial endocarditis caused by *Lactococcus garvieae*. Diagn Microbiol Infect Dis 2006;56:91–4.
[13] Wilbring M, Alexiou K, Reichenspurner H, Matschke K, Tugtekin SM. *Lactococcus garvieae* causing zoonotic prosthetic valve endocarditis. Clin Res Cardiol 2011;100:545–6.
[14] Li WK, Chen YS, Wann SR, Liu YC, Tsai HC. *Lactococcus garvieae* endocarditis with initial presentation of acute cerebral infarction in a healthy immunocompetent man. Intern Med 2008;47:1143–6.
[15] Yiu KH, Siu CW, To KK, Jim MH, Lee KL, Lau CP, et al. A rare cause of infective endocarditis; *Lactococcus garvieae*. Int J Cardiol 2007;114:286–7.
[16] Aguado-Urda M, Lopez-Campos GH, Fernandez-Garayzabal JF, Martin-Sanchez F, Gibello A, Dominguez L, et al. Analysis of the genome content of *Lactococcus garvieae* by genomic interspecies microarray hybridization. BMC Microbiol 2010;10:79.
[17] Morita H, Toh H, Oshima K, Yoshizaki M, Kawanishi M, Nakaya K, et al. Complete genome sequence and comparative analysis of the fish pathogen *Lactococcus garvieae*. PLoS One 2011;6:e23184.
[18] Miyauchi E, Toh H, Nakano A, Tanabe S, Morita H. Comparative genomic analysis of *Lactococcus garvieae* strains isolated from different sources reveals candidate virulence genes. Int J Microbiol 2012;2012:728276.
[19] Kawanishi M, Yoshida T, Yagashiro S, Kijima M, Yagyu K, Nakai T, et al. Differences between *Lactococcus garvieae* isolated from the genus *Seriola* in Japan and those isolated from other animals (trout, terrestrial animals from Europe) with regard to pathogenicity, phage susceptibility and genetic characterization. J Appl Microbiol 2006;101:496–504.
[20] Zadi-Karam H and Karam N-E. Lactic acid bacteria from camel milk. Poster session presented at: 12th Annual International French Congress of *Rencontres recherches ruminants*; 2005 Dec 7–8; Paris, France.
[21] Fox GE, Magrum LJ, Balch WE, Wolfe RS, Woese CR. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. Proc Natl Acad Sci U S A 1977;74:4537–41.
[22] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[23] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9.
[24] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and clustal x version 2.0. Bioinformatics 2007;23:2947–8.
[25] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543–51.
[26] Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, et al. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res 2012;40(Web Server issue):W622–7.
[27] Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 2013;20:714–37.
[28] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455–77.
[29] Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 2011;27:578–9.
[30] Gao S, Sung WK, Nagarajan N. Opera: reconstructing optimal genomic scaffolds with high-throughput paired-end sequences. J Comput Biol 2011;18:1681–91.
[31] Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. Genome Biol 2012;13:R56.
[32] Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 2004;32:11–6.
[33] Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. Nucleic Acids Res 2003;31:439–41.
[34] Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic Acids Res 2012;40(Database issue):D290–301.
[35] Nawrocki EP, Kolbe DL, Eddy SR. Infernal 1.0: inference of RNA alignments. Bioinformatics 2009;25:1335–7.
[36] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 2010;11:119.
[37] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinform 2009;10:421.
[38] Eddy SR. Accelerated profile HMM searches. Nucleic Acids Res 2011;7:e1002195.
[39] UniProt Consortium. Ongoing and future developments at the Universal protein Resource. Nucleic Acids Res 2011;39(Database issue):D214–9.
[40] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30.
[41] Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 2007;35(Web Server issue):W182–5.
[42] Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 2000;28:33–6.

[43] Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. Science 1997;278(5338):631–7.
[44] Drissi F, Buffet S, Raoult D, Merhej V. Common occurrence of antibacterial agents in human intestinal microbiota. Nucleic Acids Res 2015;6:441.
[45] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[46] Tindall BJ, Rossello-Mora R, Busse HJ, Ludwig W, Kampfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2010;60(pt 1):249–66.
[47] Schleifer KH, Kraus J, Dvorak C, Kilpper-Bälz R, Collins MD, Fischer W. Transfer of *Streptococcus lactis* and related streptococci to the genus *Lactococcus* gen. nov. Syst Appl Microbiol 1985;6:183–95.
[48] Morita H, Toh H, Oshima K, Murakami M, Taylor TD, Igimi S, et al. Complete genome sequence of the probiotic *Lactobacillus rhamnosus* ATCC 53103. J Bacteriol 2009;191:7630–1.
[49] Chaillou S, Champomier-Verges MC, Cornet M, Crutz-Le Coq AM, Dudez AM, Martin V, et al. The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K. Nat Biotechnol 2005;23:1527–33.
[50] Kleerebezem M, Boekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, Leer R, et al. Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc Natl Acad Sci U S A 2003;100:1990–5.
[51] Morita H, Toh H, Fukuda S, Horikawa H, Oshima K, Suzuki T, et al. Comparative genome analysis of *Lactobacillus reuteri* and *Lactobacillus fermentum* reveal a genomic island for reuterin and cobalamin production. DNA Res 2008;15:151–61.
[52] Claesson MJ, Li Y, Leahy S, Canchaya C, van Pijkeren JP, Cerdeno-Tarraga AM, et al. Multireplicon genome architecture of *Lactobacillus salivarius*. Proc Natl Acad Sci U S A 2006;103:6718–23.
[53] Forde BM, Neville BA, O'Donnell MM, Riboulet-Bisson E, Claesson MJ, Coghlan A, et al. Genome sequences and comparative genomics of two *Lactobacillus ruminis* strains from the bovine and human intestinal tracts. Microb Cell Fact 2011;10(Suppl. 1):S13.
[54] Villani F, Aponte M, Blaiotta G, Mauriello G, Pepe O, Moschetti G. Detection and characterization of a bacteriocin, garviecin L1-5, produced by *Lactococcus garvieae* isolated from raw cow's milk. J Appl Microbiol 2001;90:430–9.