Genome Mining and Transcriptional Analysis of Bacteriocin Genes in Enterococcus faecium CRL1879

Suárez N, Bonacina J, Hebert EM and Saavedra L*
Laboratorio de Genética y Biología Molecular, CERELA-CONICET, Centro de Referencia para Lactobacilos, Chacabuco 145, San Miguel de Tucumán (T4000ILC), Tucumán, Argentina

Abstract
Among 151 bacterial isolates from nine artisanal cheeses, Enterococcus faecium CRL 1879 showed antibacterial activity against the food-borne pathogen Listeria monocytogenes. The isolate produced a proteinase K-sensitive compound in the cell free supernatant. Genome analysis demonstrated the presence of enterocin A, enterocin B, enterocin P, enterocin SE-K4-like and enterocin X biosynthetic gene clusters. Nucleotide sequences encoding for a putative two-component bacteriocin were detected using bioinformatics tools, here named enterocin CRL1879αβ. A transcriptional analysis of all bacteriocin genes by quantitative real time PCR analysis (qRT-PCR) revealed the transcription of each enterocin gene at different levels. Finally, analysis of bacteriocin genes distribution in 251 E. faecium bioprojects was performed and compared to those identify in E. faecium CRL1879. The discriminative analysis demonstrated that bacteriocin genes are widely distributed among Enterococcus, independently of the origin of the strain.

The results presented in this paper represent a unique finding since this is the first demonstration of an E. faecium strain isolated from an artisanal cheese with the complete genetic machinery to produce six classes II and one class III bacteriocins.

Keywords: Bacteriocins; Enterococcus faecium; Genomes; Transcriptional analysis

Introduction
Enterococci are Gram-positive bacteria that commonly inhabit the gastrointestinal-tract of healthy humans and other animals. They are also associated with fermented foods, especially with dairy products such as cheese, where they are involved in flavor and aroma development. Several studies described the genus Enterococcus as the major component of the microbial community of certain Mediterranean and Hispanic cheeses [1-3]

Enterococci are known to produce a wide array of structurally diverse antimicrobial peptides, the so-called "enterocins", which may contribute to improve the hygienic and safety properties of food products [4]. Such attractive traits have increased the interest for their biotechnological use in food preservation, especially since these compounds are particularly active against food-borne pathogens, such as Listeria monocytogenes. Artisanal products and traditional fermented milk products were reported to be excellent sources for the isolation of bacteriocin-producing lactic acid bacteria (LAB) [5-7].

Bacteriocins are antimicrobial peptides produced by bacteria, which are active against closely related species. Various classifications have been reported so far based on features such as mode of action, primary structure, presence of post-translational modifications or specific anti-bacterial activity [8-9]. The structural gene of a typical bacteriocin might be part of one to three operon-type structures being usually small. In addition, these operons also contain genes encoding for proteins involved in transport, regulation and/or immunity [10].

Classical methods used to characterize LAB bacteriocins are based on the purification, determination of the molecular mass and sequencing of the potential antimicrobial peptide present in the cell-free supernatant of the bacteriocin producer strain. However, the nature of bacteriocins entails a difficult and expensive purification process. Only a few reports have focused on the in silico genomic screening of open reading frames (ORFs) encoding for potential small peptides as a new strategy to search for novel antimicrobial peptides [11,12]. In some cases, novel bacteriocin-related genes have been identified in LAB genomes [13,14]. Moreover, bacteriocin related genes in the genomes of non bacteriocin-producing strains have also been found [13].

In the present paper we described a promising Enterococcus faecium strain (CRL1879) isolated from an artisanal cheese, which displayed a remarkable antibacterial activity against the food-borne pathogen Listeria monocytogenes. The isolate produced a proteinase K-sensitive compound in the cell free supernatant. Genomic approaches were used to in silico characterize six class II bacteriocin clusters in this strain. Transcriptional analysis of all structural bacteriocin genes detected was performed in order to validate genomic finding. In addition, a careful analysis of the distribution, the presence and type of bacteriocin genes was evaluated in 251 E. faecium bioprojects and compared with those identify in E. faecium CRL1879. These data was used to subsequently examine if a set genes would be effective in predicting a category membership meaning if specific bacteriocin genes were present in niche-specific strains.

This is the first description of a food-related E. faecium (CRL1879) strain harboring six class II and one class III bacteriocin genes. The data presented here provides critical starting points for future functional studies on enterococcal bacteriocins.
Materials and Methods

Source and maintenance of the cultures

_E. faecium_ CRL1879 was isolated from an artisanal cheese of Northwestern Argentina and it is deposited under CEREA Culture Collection. _Listeria innocua_ 7 (Unité de Recherches Laitières et Génétique Appliquée, INRA France); _L. monocytogenes_ FBUNT (Bacteriology Department, Faculty of Biochemistry, National University of Tucumán); _L. monocytogenes_ 24, _L. monocytogenes_ 139, _L. monocytogenes_ 149, _L. monocytogenes_ 152 (Department of Dairy and Fat Technology, Institute of Chemical Technology, Prague); _L. monocytogenes_ EGD-e [15], _E. coli_ DH10 B (Invitrogen); _Staphylococcus (S) aureus._ ATCC29740. _E. faecium_ ATCC19434; _E. mundtii_ CRL 35 [6], _E. faecium_ CRL1, CRL6, CRL11, CRL23, CRL147, CRL316, CRL318, CRL486, CRL589, CRL1571 (CEREA Culture Collection) were used in this study. All enterococci strains were routinely cultivated in LAFTg broth (peptone, 15 g/l; tryptone, 10 g/l; yeast extract, 10 g/l; glucose, 10 g/l; Tween 80, 0.1 % v/v; agar 1.5%) at 37°C. _L. Staphylococcus_ strains were grown in complex media Brain Heart Infusion broth (BHI, Britania, Buenos Aires, Argentina) at 37°C and _E. coli_ DH10 B in LB medium [16] at 37°C under aerobic conditions. All cultures were grown overnight. Bacterial stock cultures were maintained in 10% (w/v) skim milk with 0.5 % (w/v) yeast extract and 10 % (v/v) glycerol at ~80°C. Agar plates were prepared using each culture medium plus 1.5% of agar (Britania, Buenos Aires, Argentina).

Antimicrobial activity assays of _E. faecium_ CRL1879

Cell-free supernatant (CFS) of _E. faecium_ CRL1879 was screened for antimicrobial activity by the agar diffusion assay [17] against the indicator strains. Briefly, CFS was obtained by growing _E. faecium_ CRL1879 in LAFTg broth for 16 h at 37°C. The bacterial culture was centrifuged at 15,600 x g for 10 min to obtain CFS. Then 5 μl of CFS was spotted in plates containing 10 ml of BHI 1.5% agar plus 10 ml of BHI soft agar (0.7%) inoculated with 10^7 colony forming unit ml^-1 of an overnight culture of the indicator strains. Plates were incubated for 24 h at the corresponding incubation temperature and examined for clear zones of growth inhibition. In order to characterize the antimicrobial compound, aliquots of 200 μl of CFS were a) adjusted to pH 7.0 with 1 N NaOH and b) added 1000 U ml^-1 of catalase, c) treated with proteinase K (20 mg/ml) -Invitrogen, Buenos Aires, Argentina) at 42°C for 2 h, d) heated at 100°C for 5 min (most bacteriocins are heat stable). In all the cases, a positive control (fresh unmodified CFS) was tested in parallel.

In silico identification of bacteriocin genes in _E. faecium_ CRL1879 genome

_E. faecium_ CRL1879 genome sequence is deposited at Genbank (AN AOUK00000000.1). Genes encoding bacteriocins and post translationally modified peptides (RIPPs) or those related with their production systems were searched in _E. faecium_ CRL1879 genome. Rapid Annotations using Subsystem Technology server (RAST) and BAGE3 software Linux version were used to perform the analysis [18,19]. Putative promoter regions were analyzed with Promoter Prediction by Neural Network (NNPP) and BPROM (Softberry Inc., Mount Kisco, NY, USA) [20]. Bacteriocin sequence analyses were done with the bacteriocin database [21,22]. The analysis of the nucleotide sequences, translation and alignments were performed with DNAstar software version 11.2.1.25 (DNASTAR Inc., USA). The identity of sequences was analyzed using Blastx service from the NCBI database.

Transcriptional analysis of bacteriocin genes in _E. faecium_ CRL1879

RNA isolation: Total RNA was extracted from _E. faecium_ CRL 1879 according to Raya et al. from three independent cultures and separated into at least two technical repeats [23]. Briefly, cells from 50 ml of a culture of 12 h were harvested by centrifugation, washed with TE buffer (10 mM Tris, 1 mM EDTA; pH 8.0) and suspended in 500 μl of cold TE buffer. The cell suspension was added to a 2-ml screw-cap microcentrifuge tube containing 0.6 g of glass beads (0.1-mm diameter), 170 μl of 2% macaloid slurry, 500 μl of Tris-HCl buffered phenol-chloroform (1:1), and 50 μl of 20% sodium dodecyl sulfate. Cells were disrupted by shaking in a Mini-Beadbeater Cell Disrupter (Model 607EUR, Biospec Products, Battleville, Okla, USA) during 6 cycles of 1min of disruption following by 1 min on ice. After centrifugation at 15,600 x g, 4°C for 15 min, the aqueous supernatant containing the RNA was two times extracted with 1volume of phenol-chloroform and precipitated with sodium acetate 3 M, pH 5.2 and absolute ethanol. Finally, RNA was suspended in RNase free water. The isolated RNA was treated with "TURBO DNA-free" kit (Ambion™, Invitrogen, Buenos Aires, Argentina) according to the manufacturer’s instructions, before the first-strand synthesis of cDNA.

qRT-PCR: Synthesis of cDNA were performed from 100 ng of total DNA-free RNA, using qScript™ CDNA SuperMix (Quanta-Biosciences, Genbiotech, Argentina). The primers used in qRT-PCR are listed in (Table 1). These were designed according to the entA, entB, entP, ent EntSE-K4, entXαβ and entCRL1879aβ sequences according to the genome sequence of _E. faecium_ CRL1879 with Primer Quest software (Integrated DNA Technologies, Inc.). [26]. qRT-PCR reactions were performed in duplicates using 35 ng cDNA, 0.3 μM of each primer and 10 μl of qPCR Master Mix (PerfeCta " SYBR" Green Supermix for iQ™, Quanta Biosciences™ Genbiotech, Argentina) on a final volume of 20 μl per reaction. Reactions were run on an iQ™ 5 Multicolor Real-Time PCR Detection System (BioRad, USA). PCR conditions were: initially 2 min/50°C, 3 min/95°C, 40 cycles consisting of 10 s/95°C, 30 s/50°C following by a melting curve temperature cycles. For each reaction, a non-template control (NTC) was included (without cDNA) and the invariable expression of 16S rRNA under the test conditions (data not shown) was used as house-keeping gene. Standard curves for entA, entB, entP, entSE-K4, entXαβ and entCRL1879aβ genes were performed in order to determine the copy number of each transcript. These genes were amplified by PCR from _E. faecium_ CRL 1879 genomic DNA and purified using the Wizard® SV Gel and PCR Clean-Up System kit (Promega, Madison-USA). All primers and standard curves for qRT-PCR assay displayed efficiencies between 98 and 110%. A single product-specific melting curve was observed for each primer set. The concentration of the amplified product was measured with a UV-visible spectrophotometer (Varian, 50MPR Microplate Reader -Cary) at 260 nm using a standard calibration curve, and converted to number of copies per unit volume using the following equation:[25].

(\text{DNA copy} = \frac{6.023 \times 10^{23} \text{ [copies mol}\text{-1]}}{\text{DNA length [bp]} \times 660 \text{ [g mol}\text{-1 bp}\text{-1}})\)

The concentration of the purified linear dsDNA standards was adjusted to 10^-6 copies per μl. This stock solution was tenfold serially diluted to obtain a standard series from 10^-6 to 10^0 copies per μl and it was used to construct the standard curves for enterocin structural genes. Each standard dilution was analyzed in duplicate. The Ct values were plotted against the logarithm of their initial template copy concentration to generate each standard curve by a linear regression.
**Citation:** Suárez N, Bonacina J, Hebert EM, Saavedra L (2015) Genome Mining and Transcriptional Analysis of Bacteriocin Genes in Enterococcus faecium CRL1879. J Data Mining Genomics Proteomics 6: 180. doi:10.4172/2153-0602.1000180

**Table 1:** qRT-PCR specific primers used in this study.

| Sensitive strains | Antimicrobial Activity (mm) | ATCC 19434 | CRL 1879 |
|-------------------|-----------------------------|------------|----------|
| E. coli DH10 B    | -                           | -          |          |
| E. faecium ATCC19434 | -                   | -          |          |
| E. faecium CRL1   | -                           | -          |          |
| E. faecium CRL6   | -                           | -          | ++       |
| E. faecium CRL11  | -                           | -          | ++       |
| E. faecium CRL23  | -                           | -          | +        |
| E. faecium CRL147 | -                           | -          |          |
| E. faecium CRL316 | -                           | -          |          |
| E. faecium CRL318 | -                           | -          | ++       |
| E. faecium CRL486 | -                           | -          | ++       |
| E. faecium CRL589 | -                           | -          | ++       |
| E. faecium CRL1571 | -                      | -          | ++       |
| E. mundtii CRL35  | -                           | -          | +        |
| L. innocua 7      | +                           | +          | +        |
| L. monocytogenes 139 | +                      | +          | +        |
| L. monocytogenes 149 | +                      | +          | +        |
| L. monocytogenes 152 | +                      | +          | +        |
| L. monocytogenes 24 | +                        | +          | +        |
| L. monocytogenes EGD-e | -                       | -          |          |
| L. monocytogenes FBUNT | +                       | +          | ++       |
| S. aureus ATCC29740 | -                      | -          |          |

Diameter of growth inhibition zones (mm) (++): >10 mm; (+): <10 mm; (-): no growth inhibition zone.

**Table 2** Inhibitory activity of heated and neutralized cell-free supernatants of *E. faecium* CRL1879 and ATCC 19434 strains as determined by the agar spot test method.

When assaying the samples for each enterocin expression, the corresponding standard series were run under the same conditions and the copy numbers of samples was determined by reading off the standards series with the Ct values of the samples.

**Genomic mining of putative bacteriocin genes in *E. faecium* bioprojects**

251 *E. faecium* genomes bioprojects available at NCBI Genome database were analyzed by BAGEL3 software Linux version, BLASTx algorithm and bacteriocin database Bactibase [20-23] (Table S1). The resulting data from BAGEL3 was subjected to linear discriminant analysis (LDA). LDA was used to establish whether the presence of a set of *bac* genes was effective in predicting category membership, i.e, specific *bac* genes present in niche-specific strain. LDA was performed with Statistica 10.0 software. The discrete distribution of *bac* genes among strains isolated from human, animal, food and those of unknown provenance was determinate. All genomes analyzed were previously classified depending on the origin of isolation of the strains, i.e, animal (AN); human (H); food (F) sources and strains with unknown provenance (U) (Table S2).

**Results**

**Antibacterial activity of *E. faecium* CRL1879**

*E. faecium* CRL1879 was isolated from an artisanal cheese made in Northwestern region of Argentina. This strain was selected due to the interesting antimicrobial activity observed in cell-free supernatant (CFS); *E. faecium* ATCC19434 (enterocin A and B producer, unpublished data) was used as positive control (Table 2). All L. monocytogenes strains, including biofilm-forming strains 24,139,149,152, were sensitive to CFS except EGD-e strain. Regarding enterococci used as indicators, it is surprisingly that CFS from CRL1879 turned out to be active against almost all strains tested as compared to the antimicrobial activity of the ATCC19434 CFS, suggesting the potential production of different antimicrobial compounds among strains. On the other hand, CFS did not show antimicrobial activity against *E.coli* DH10B, *S. aureus* ATCC 29740. To further characterize the antimicrobial compound, CFS aliquots were i) neutralized to pH 7, ii) treated with catalase and iii) with proteinase K and tested against *L. monocytogenes* FBUNT. Untreated CFS from CRL1879 was used as positive control. Growth inhibition halos were only sensitive to proteinase k demonstrating that the antimicrobial compound present in CFS was of proteinaceous nature, likely a bacteriocin.

**Genome identification of bacteriocin production systems in *E. faecium* CRL 1879**

The genome sequence of *E. faecium* CRL 1879 was acquired using a whole-genome shotgun (WGS) strategy with an Ion Torrent personal genome machine. Quality filtered reads were *in silico* assembled...
using the DNASTAR Ngen assembler, giving 73 large contigs (AN AOUK00000001) [24]. RAST Server, Blast algorithms and BAGEL3 analysis demonstrated the existence of enterocin P, enterocin B, enterocin X αβ, enterocin A and enterocin SE-K4-like biosynthetic clusters (Figure 1).

Sequence analysis of the enterocin P gene cluster (contig_152, AN AOUK01000025) revealed the same genetic organization as previously described by Cintas et.al. [26] (Figure 1). The deduced amino acid sequence of entP gene showed 95% identity to enterocin P described by Cintas et al. with two amino acids changes in the leader peptide at positions 9 (A-T) and 13 (I-K) and one at the position 69 (M-I) of the mature peptide [26].

Enterocin B biosynthetic cluster from E. faecium CRL1879 (contig_71, AN AOUK01000001) was identical to the one previously described by Franz et al. [27]. By using the software BAGEL 3, genomic sequence analysis of the same contig revealed the presence of ORFs encoding for two -component bacteriocins, enterocin X previously described in E. faecium KU-B5 [28]. A putative immunity gene for enterocin X (entiX) was located 584 nucleotides upstream of the enterocin B precursor gene, in the same direction. The putative two components of enterocin X locus containing enterocin Xa (entiXa) and the enterocin Xβ genes (entiXβ) was located upstream entiX (Figure 1).

BLAST and RAST analysis of contig_191 (AN AOUK01000051) demonstrated that a fragment of 700 bp display 94% of identity with E. faecalis pAMS1 (NG_036515.1) and 92% with pEK4S (AB092692.1) plasmids, both known for their bacteriocin-encoding genes [29,30]. The biosynthetic cluster of this enterocin SE-K4-like bacteriocin is organized in an operon-like structure consisting in two consecutive ORFs (Figure 1). The first ORF 1 encodes a 71 amino acid protein corresponding to enterocin SE-K4-like pre-peptide. The deduced amino acid sequence of the putative bacteriocin encoding gene in CRL1879 genome exhibited only 77% of identity to enterocin SE-K4 first described in E. faecalis K4 (BAC20326.1) with six amino acid deletion at the C-terminal domain (Figure 1). However, this nucleotide sequence is widely distributed and conserved among E. faecium genomes strains deposited at NCBI (EEV57603.1, EFS32511.1, EFR67967.1, EJX61666.1, EJX72841.1, among others). ORF2 gene, encoding a putative immunity protein composed of 113 amino acids, was immediately found downstream of the bacteriocin gene. A putative ribosome binding site (RBS) for the bacteriocin structural gene (AAGGTG) was located 7 bases upstream of the initiation codon. Possible -10 and -35 promoter sequences were also detected.

The enterocin A locus was located in contig_279 (AN AOUK01000059) (Figure 2). Sequence analysis showed the presence of two putative operons: (I) the bacteriocin operon with the enterocin A gene (entA), the immunity gene (entI), and three component regulatory system the peptide pheromone gene (entF), the histidine protein kinase gene (entK), and the response regulator (entR), and (II) the transporter operon consisting of the two genes, the ABC transporter (entT) and its accessory gene (entD) as previously described by O’ Keeffe et al. [31]. Analysis of 1474 bp intergenic sequence located between enterocin A-regulatory and transport operons revealed three consecutive ORFs oriented in opposite direction to the complete entA cluster. Both, ORF 1 and ORF2, encode putative double-glycine-type pre-peptide with a 19 and 21-residue leader sequence with theoretical pl/MW 9.31/514.86 and 7.78/4338.99, respectively (Figure 2). Although deduced amino acid sequence of ORF1 showed 97-99% identity with annotated as putative bacteriocin-like proteins present in several E. faecium genomes (YP_008396028.1; YP_007395047.1; EEV45206.1;
EEF20746.1), no clear homology was found to any biochemically characterized bacteriocin. ORF2 deduced amino acid sequence showed 97-99% identity with hypothetical proteins (YP_007395046.1; YP_008396027.1; EEVE45207.1) but no homology to any previously known bacteriocin. Sequence analysis of the deduced amino acid sequence of these two putative peptides also showed the GXXXG motif, suggested to be involved in helix-helix interaction between two component bacteriocins [32]. No obvious rho-independent transcriptional terminator (invert repeats) was found between ORF1 and ORF2 suggesting that may belong to the same transcriptional unit. Thus, we named here enterocin CRL1879αβ as a potential two-component bacteriocin. In addition, ORF3 located downstream of ORF2 predicted a 63-residue, hydrophobic and cationic (pI 10.47) protein with one putative trans-membrane domain resembling an immunity-like protein, but no sequence identity could be determined either.

### E. faecium CRL 1879 expresses multiple bacteriocin genes

In order to validate genomic findings we performed transcriptional analyses of all putative bacteriocin genes found in E. faecium CRL 1879. qRT-PCR experiments revealed the presence of entP, entSE-K4, entB, entA, entXαβ subunits as well as entCRL1879αβ transcripts (Table 3). The gene encoding for the enterocin P structural peptide presented the highest level of expression (2.4 x 10^7 copies μl^-1), while enterocin Xα and Xβ gene transcripts were present in an order of 1.6 x 10^4 and 2.7 x 10^4 copies μl^-1 respectively. The entA gene was transcript in an order of 2.4 x 10^4 copies μl^-1. Interestingly, entCRL1879αβ genes were transcript at 7.7 x 10^5 and 7.4 x 10^4 copies μl^-1, entSE-K4 transcript was present at 2.5 x 10^5 copies μl^-1 and finally entB transcript was in an order of 9.2 x 10^4 copies μl^-1.

### Comparative genomic analysis of putative bacteriocin genes among E. faecium strains

A detailed in silico analysis of the distribution, the presence and type of bac genes was evaluated in E. faecium bioprojects and compared with those identify in E. faecium CRL1879. 251 partial or complete E. faecium genomes, deposited at the NCBI Genome database, from different origin were mine searching bacteriocin genes using BAGEL3 software and BLASTx algorithm (Table S1). Exhaustive analysis demonstrated that i) almost all strains contain one to six putative bacteriocin genes except to PRJNA82503 and PRJNA179542 which exhibited no putative bacteriocin genes ii) 24 bacteriocin class II encoding genes and one for enterolysin A, a class III bacteriocin, were detected; iii) known enterocins (enterocin A, B, P SE-K4, X, NKR-5-3A, L50A and B) showed a highly conserved sequence among the genus; iv) not all strains carrying enterocin A genes have enterocin B gene v) enterolysin A, a bacteriocyn, is widely distributed among E. faecium genomes; vi) as a general rule, enterocin Xαβ is disseminated among E. faecium strains that concomitantly harbors entB gene. One strain showed enterocin Xαβ in the absence of entB gene. Conversely, six genomes showed only entB gene; vii) interestingly, we found genes encoding for bacteriocins described in the Lactobacillus genus (sakacin Q and sakacin T) and viii) E. faecium CRL1879 is the only food related strain harboring six class II and one class III bacteriocin genes. The resulting data from BAGEL3 was subjected to LDA statistical analysis. Each individual strain was a priori classified according to its origin of isolation (Table S1). Once each bacteriocin gene was detected by BAGEL3, LDA was used to establish whether the presence of a set of bac genes was effective in predicting category membership i.e. the origin of a strain (Figure3) (Table S2). This analysis measure the percentage of similarities according to the presence or absence of variables used to identify each individual. Results demonstrated that 53% of strains analyzed were classified in the same groups as the a priori classification (Table 4). LDA showed that 63.33% of total strains a priori classify as Human could be assigned in the same group. However, the discriminative analysis demonstrated bacteriocin genes are widely distributed among Enterococcus, independently of the origin of the strain and there is no strong association between the origin and the number of bacteriocin genes present in a specific strain.

### Discussion

Northwestern Argentinan cheeses have been prolific substrates where numerous LAB strains with technological properties were isolated and characterized [17]. Several studies described the genus

---

**Table 3:** Enterocin transcripts copy number obtained by qRT-PCR.

| E. faecium CRL1879 | Ct and estimated enterocin copy number by RT-PCR absolute quantification* |
|-------------------|--------------------------------------------------------------------------------|
| Bacteriocin mRNA copies (copies μl^-1) | 2.4 x 10^6 | 9.2 x 10^4 | 2.4 x 10^7 | 2.5 x 10^6 | 1.6 x 10^5 | 2.7 x 10^5 | 7.7 x 10^4 | 7.2 x 10^5 |

*Estimated copy number of enterocin genes transcripts in E. faecium CRL 1879 cells at laboratory conditions, by RT-PCR absolute quantification.

| **Classification Matrix** |
|---------------------------|
| **A priori classification** | Human (H) | Unknown (U) | Animal (AN) | Food (F) |
|----------------------------|-----------|-------------|-------------|---------|
| % of strains defined in the corresponding category classification* | 63.6 | 77 | 11 | 28 | 5 | 7 |
| **Total** | 53.4 | 129 | 33 | 70 | 19 |

* A priori classification according to the strain isolation source information obtained from the Bioproject at the NCBI Genome database.

**: Percentage of strains that effectively correspond to their a priori established classification after the LDA analysis.
Enterococcus as the major component of the microbial community of certain Mediterranean and Hispanic cheeses [2,3,33]. The predominance of enterococci into the microbial population of various fermented foods such as meat or cheese can be explained by the frequent production of bacteriocins, as well as its tolerance to heat, dry and saline environments [34,35]. In the present paper, we demonstrated that CFS of CRL1879 strain is active against six different Listeria strains including biofilm-forming bacteria (strains 139, 149, 152, 24) except to the widely used L. monocytogenes EGD-e. In contrast, CFS was not active against E. coli DNH108 neither to S. aureus ATCC 29740 used as indicator strains. In general, bacteriocins produced by Gram-positive bacteria have no effect on Gram-negative bacteria without the addition of any membrane-active compound, but there is some exceptions to the rule [36].

*E. faecium* CRL1879 strain have known structural genes encoding for class II bacteriocins: enterocin A, enterocin B, enterocin P, enterocin X and enterocin SE-K4-like. Enterocin A and B first discovered and identified in *E. faecium* T136 are frequently present in *E. faecium* strains from various sources [37,38]. Enterocin P is a sec-dependent bacteriocin widely distributed among *E. faecium* strains. We have previously shown that the presence of more than one bacteriocin gene in *Enterococcus* genus is a common finding [39]. Interestingly, the gene encoding for enterocin SE-K4 was only described in *E. faecalis* strains [29]. The co-existence of genes encoding for the production of enterocin SE-K4-like bacteriocin with other *bac* genes in *E. faecium* strains have not been described in the literature yet. To our knowledge, this is the first report of an *E. faecium* strain that possess the structural gene encoding for enterocin SE-K4-like bacteriocin.

Since the first enterococcal genome sequenced - that of *Enterococcus faecalis* V583, now a veritable avalanche of complete or draft genome sequences of various enterococcal strains and species are available [40]. Recently as part of this paper, *E. faecium* CRL1879 genome has been sequenced by a whole-genome shotgun [24]. Based on this data, in silico screening strategies were employed to mine for bacteriocins or ORFs encoding for potential small peptides in DNA sequences using specially designed software such as BAGEL [20]. Genome sequence analysis of *E. faecium* CRL1879 confirmed the presence of gene clusters involved in enterocin P, B, A and SE-K4-like bacteriocin production, transport and immunity. They are chromosomally encoded and they show high degree of similarity to those previously described in the literature except for enterocin SE-K4 as described in the results section. In addition, a gene cluster encoding for enterocin X production was also found. Enterocin X cluster was located in the vicinity of enterocin B in similar fashion as previously described in the literature [28]. BAGEL3 analyses also revealed a potential two-component bacteriocin at the intergenic sequence located between enterocin A-regulatory and transport operons. Although, three consecutive ORFs (ent CRL1879αβ and CRL1879αγ) oriented in opposite direction to the complete entA cluster are present in several *E. faecium* genomes, we demonstrated that *ent CRL1879αβαγ* are transcribed under our laboratory conditions (Table S2). Transcriptional analyses showed also that *entP* transcript showed the highest copy number which might be explained since enterocin P is exported via the general secretory (Sec-dependent) pathway. Such a mechanism of export is potentially beneficial to the host due to its apparent metabolic and genetic economy. Only a few studies have analyzed the transcriptional profile of the bacteriocin production in LAB. A reverse transcription-PCR analysis showed the expression of *entA, entB, and munK3* bacteriocin genes in *E. faecium* CWB1-B1340 and *entP* and *munK5* genes in *E. mundtii* CWB1-B1431 [41]. Moreover, enterococcal strains characterized in the literature so far bear up to four bacteriocin genes [42, 9]. qRT-PCR results obtained here validated the in silico bacteriocin genome findings demonstrating that the *E. faecium* CRL1879 strain isolated from an Argentinean northwestern artisanal cheese possess the complete genetic machinery necessary to produce six antimicrobial peptides.

**Figure 3:** Raw discrete data plot performed with Statistica 10.0 software. Each data point in this plot is represented as a rectangular region with two different colors, black or grey, that correspond to the 1 or 0 values which indicate the presence or absence of bac genes for each of the *E. faecium* genomes classified as animal (AN); human (H); food (F) sources and strains with unknown provenience (U) depending on the origin of isolation of the strains. Values within each series are presented along the x-axis, with each series plotted along the y-axis. Bacteriocin genes identified: enterocin A (EA), enterocin B (EB), enterocin P (EP), enterocin SE-K4 (ESEK4), enterocin Xa (Exa), enterocin Xb (Exb), enterocin E397 (E397), enterocin LSGA_B (ELSOAB), enterocin 96 (E96), enterocin Q (EQ), acidocin LF221B-gassocininK7B (Acid), lactacin_F subunit (LacF), lactococcin 972 (Lac972), hiracin JM79 (HJM79), bac 32, uviolin B (UviB), sakacin Q (SakQ), sakacin T αβ (Sak αβ), broocinin C (BroC), enterocin NKR-5-3D, enterocin NKR-5-3A, glycin, sactipeptides (Sactipeptide), enterolysin A (Elysin A), zoocin A.

In summary, bacteriocin genes are extensively distributed among Enterococcus genus, independently of the origin of the strain and there is no strong association between the origin and the number of bacteriocin genes present in a specific strain. This is first demonstration of a bacterial strain, *E. faecium* CRL 1879, able to produce six class II bacteriocins. *E. faecium* CRL 1879 has a promising potential in food preservation. Ongoing research is evaluating the bacteriocin producing...
machinery in a model food system.

Acknowledgments

This study was carried out with the financial support from CONICET (PIP/0406/12) and MinCyT (PICT2011 N°0175) from Argentina.

Authors are grateful to Lic. Elena Bru for her collaboration with statistical analysis.

References

1. Abriouel H, Martin-Platero A, Maqueda M, Valdivia E, Martinez-Bueno M (2008) Biodiversity of the microbial community in a Spanish farmhouse cheese analyzed by PCR-dependent and culture-independent methods. Int J Food Microbiol 127: 200-208.
2. Casalta E, Sorba JM, Aigle M, Ogier JC (2009) Diversity and dynamics of the microbial community during the manufacture of Calenzana, an Artisanal Corsican cheese. Int J Food Microbiol 133: 243-251.
3. Reny JA Jr., Somkuti GA, Paul M, Van Hekken DL (2009) Characterization of antilisterial bacteriocins produced by Enterococcus faecium and Enterococcus durans isolates from Hispanic-style cheeses. J Ind Microbiol Biotechnol 36: 261-268.
4. Hauge HH, Nissen-Meyer J, Nes IF, Eijsink VG (1998) Amphiphilic alpha-helices are important structural motifs in the alpha and beta peptides that constitute the bacteriocin lactococcin G enhancement of helix formation upon alpha-beta interaction. Eur J Biochem 251: 565-572.
5. Andrighetto C, Kniff E, Lombardi A, Torriani S, Vancanneyt M, et al. (2001) Phenotypic and genetic diversity of enterococci isolated from Italian cheeses. J Dairy Res 68: 303-316.
6. Saavedra L, Minakh C, de Ruiz Holgado AP, Sesma F (2004) Enhancement of the enterocin CRL35 activity by a synthetic peptide derived from the NH2-terminal sequence. Antimicrob Agents Chemother 48: 2778-2781.
7. Terzic-Vidojevic A, Vukasinovic M, Veljovic K, Ostojic M, Topisirovic L (2007) Characterization of microflora in homemade semi-hard white Chiaro cheese. Int J Food Microbiol 114: 36-42.
8. Cotter PD, Hill C, Ross RP (2005) Bacteriocins: developing innate immunity for food. Nat Rev Microbiol 3: 777-788.
9. Perez RH, Himeno K, Ishibashi N, Masuda Y, Zendo T, et al. (2012) Monitoring of the multiple bacteriocin production by Enterococcus faecium NKR-S-3 through a developed liquid chromatography and mass spectrometry-based quantification system. J Biosci Bioeng 114: 490-496.
10. Snyder AB, Worobo RW (2014) Chemical and genetic characterization of bacteriocins: antimicrobial peptides for food safety. J Sci Food Agric 94: 24-44.
11. Begley M, Cotter PD, Hill C, Ross RP (2009) Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM of the multiple bacteriocin production by Enterococcus faecium. J Bacteriol 191: 3183-3188.
12. O’Keeffe T, Hill C, Ross RP (1999) Characterization and heterologous expression of the genes encoding enterocin a production, immunity, and regulation in Enterococcus faecium DPC1146. Appl Environ Microbiol 65: 1506-1515.
13. Rogne P, Finnland G, Nissen-Meyer J, Kristiansen PE (2008) Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G. Biochim Biophys Acta 1784: 543-554.
14. Javed I, Ahmed S, Ali M, Ahmad B, Ghumro PB, et al. (2010) Bacteriocinogenic potential of newly isolated strains of Enterococcus faecium and Enterococcus faecalis from dairy products of Pakistan. J Microbiol Biotechnol 20: 1506-1515.
15. Aymerich T, Garriga M, Ylla J, Vallier J, Monfort JM, et al. (2000) Application of a developed liquid chromatography and mass spectrometry-based quantification system. J Immunol Methods 278: 261-269.
16. Fontana C, Cocconcelli PS, Vignolo G, Saavedra L (2015) Occurrence of antilisterial structural bacteriocins genes in meat borne lactic acid bacteria. Food Control 47: 53-59.
40. McShan W, Shankar N (2002) The Genome of *Enterococcus faecalis* V583: a Tool for Discovery: The Enterococci. (ed ASM Press), Washington, DC.

41. Aguilar-Galvez A, Dubois-Dauphin R, Campos D, Thonart P (2011) Genetic determination and localization of multiple bacteriocins produced by *Enterococcus faecium* CWBI-B1430 and *Enterococcus mundtii* CWBI-B1431. Food Sci Biotechnol 20: 289-296.

42. Khan H, Flint S, Yu PL (2010) Enterocins in food preservation. Int J Food Microbiol 141: 1-10.

43. Strompfova V, Laukova A, Simonova M, Marcinakova M (2008) Occurrence of the structural enterocin A, P, B, L50B genes in enterococci of different origin. Vet Microbiol 132: 293-301.

44. Nilsen T, Nes IF, Holo H (2003) Enterolysin A, a Cell-Wall-Degrading Bacteriocin from *Enterococcus faecalis* LMG 2333. Appl Environ Microbiol 69: 2975–2984