Arch Microbiol
DOI 10.1007/s00203-016-1250-4

ORIGINAL PAPER

Genome mining of lipolytic exoenzymes from *Bacillus safensis* S9 and *Pseudomonas alcaliphila* ED1 isolated from a dairy wastewater lagoon

Florencia A. Ficarra[1,3] · Ignacio Santecchia[1,3] · Sebastián H. Lagorio[3] · Sergio Alarcón[2,3] · Christian Magni[1,3] · Martín Espariz[1,3]

Received: 12 February 2016 / Revised: 9 May 2016 / Accepted: 27 May 2016

Abstract Dairy production plants produce highly polluted wastewaters rich in organic molecules such as lactose, proteins and fats. Fats generally lead to low overall performance of the treatment system. In this study, a wastewater dairy lagoon was used as microbial source and different screening strategies were conducted to select 58 lipolytic microorganisms. Exoenzymes and RAPD analyses revealed genetic and phenotypic diversity among isolates. *Bacillus safensis, Pseudomonas alcaliphila* and the potential pathogens, *B. cereus, Aeromonas* and *Acinetobacter* were identified by 16S-rRNA, *gyrA, oprI* and/or *oprL* sequence analyses. Five out of 10 selected isolates produced lipolytic enzymes and grew in dairy wastewater. Based on these abilities and their safety, *B. safensis* S9 and *P. alcaliphila* ED1 were selected and their genome sequences determined. The genome of strain S9 and ED1 consisted of 3,794,315 and 5,239,535 bp and encoded for 3990 and 4844 genes, respectively. Putative extracellular enzymes with lipolytic (12 and 16), proteolytic (20) or hydrolytic (10 and 15) activity were identified for S9 and ED1 strains, respectively. These bacteria also encoded other technological relevant proteins such as amylases, proteases, glucanases, xylanases and pectate lyases.

Keywords Dairy wastewater · *Bacillus safensis* · *Pseudomonas alcaliphila* · Lipases · Esterases

Introduction

The dairy industry is one of the most polluting of food industries, and it is considered non-environment friendly since up to 10 L of highly polluted wastewater are produced per liter of processed milk (Heaven et al. 2012). This wastewater is characterized by being rich in biodegradable organic molecules such as lactose, and it usually contains high levels of fats and proteins as well (Cammarota and Freire 2006; Heaven et al. 2012; Vidal et al. 2000). Numerous aerobic or anaerobic treatment processes are used to handle this sort of wastewater including anaerobic reactors, activated sludge, trickling filters, aerated or anaerobic lagoons, or a combination of them (Cammarota and Freire 2006; Loperena et al. 2009). However, the low biodegradability and bioavailability coefficient of fats usually result in a low overall performance of the treatment system (Cammarota and Freire 2006; Loperena et al. 2006).

Bioaugmentation is a technique that improves the capacity of a contaminated matrix to remove pollution by the addition of specific competent strains or consortia of microorganisms (El Fantroussi and Agathos 2005). Bioaugmentation has been used to improve the reactor start-up or its performance, to protect the existing microbial community against adverse effects, to accelerate the onset of degradation or to compensate for organic or hydraulic overloading

Communicated by Pierre Béguin.

Electronic supplementary material The online version of this article (doi:10.1007/s00203-016-1250-4) contains supplementary material, which is available to authorized users.

✉ Martín Espariz
espariz@ibr-conicet.gov.ar

[1] Instituto de Biología Molecular de Rosario (IBR-CONICET), Suipacha 590, S2002LRK Rosario, Argentina

[2] Instituto de Química de Rosario (IQUIR-CONICET), Suipacha 531, S2002LRK Rosario, Argentina

[3] Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, S2002LRK Rosario, Argentina

Published online: 07 June 2016

(Venkata Mohan et al. 2009). Microorganisms used for bioaugmentation should be (1) nonpathogenic, (2) capable of degrading target pollutants under in situ conditions, (3) competitive and persistent after inoculation, (4) compatible with the indigenous microbial community and (5) complementary and/or synergic for the degradation of an effluent type or specific compounds (Loperena et al. 2006, 2007).

Several commercial products based on bacteria from different sources are currently available to improve the overall performance of biological wastewater treatment systems (Loperena et al. 2006). To overcome ecological barriers, microorganisms isolated from dairy wastewater treatment systems were recently proposed as a better alternative to treat such effluents. Thus, (Loperena et al. 2006, 2009) designed consortia composed of strains isolated from dairy wastewaters that decrease chemical oxygen demand (COD), proteins and fats in simulated dairy wastewater. Likewise, Rajeshkumar and Jayachandran proposed the use of *Alcaligenes* sp. MMRR7, strain isolated from a high organic load dairy wastewater, for COD reduction (Rajeshkumar and Jayachandran 2004). However, in these studies the pathogenic nature of the isolates was not considered, though they belong to pathogenic or opportunistic genera like *Alcaligenes*, *Staphylococcus* and *Pseudomonas*. This may disfavor the approval or commercialization of the inocula for their use in bioaugmentation processes.

Despite the fact that the contribution of bacteria to the biodegradation of dairy wastewater content is well known, a genome-wide approach to identify key enzymes responsible for this process has not yet been conducted. On the other hand, there are several studies investigating specific hydrolytic enzymes with the attempt of optimizing the performance of wastewater treatment (Cammarota and Freire 2006).

In this paper, the aim was to study bacterial strains with potential use in dairy wastewater treatment at genome level. Several screenings of lipolytic enzyme-producing bacteria were performed, and as a result, 58 strains were isolated from dairy wastewater and preliminary characterized. Among them, the saprophytic bacteria *B. safensis* S9 and *P. alcaliphila* ED1 showed to be capable of growing and producing lipolytic enzymes in dairy wastewater. Finally, their genome sequences were determined in an attempt to identify novel fat-removing exoenzymes and other relevant proteins.

Materials and methods

Source, screening and isolation of microorganisms

Bacteria were isolated from a facultative wastewater lagoon (33.117359S 60.597047W) from an industrial dairy plant. Samples were obtained from suspended solids at the bottom (sludge), paper-filtered liquid at the mid-depth and solids at the surface of the lagoon. The isolations were performed following three different strategies:

1. Enrichment and selection: 300 ml of sterile enrichment medium was inoculated with 30 ml of pooled lagoon samples in a 500-ml flask where sterile air was pumped. The enrichment medium was the minimum R medium (67 mM KH_2PO_4, 10.5 mM $(NH_4)_2SO_4$, 1.6 mM $MgSO_4{\cdot}7H_2O$, 3.7 μM $FeSO_4{\cdot}7H_2O$) supplemented with 1 % (w/v) powdered milk and 0.15 % (v/v) olive oil. The culture was grown for three days at 25 °C, and then, a sample of 30 ml was used to inoculate a new 500-ml reactor flask. This operation was repeated three times. Then, bacteria were isolated by sample dilution and plated onto egg yolk agar [LB containing 5 % (v/v) egg yolk].
2. Selection by direct plating: Lagoon samples were diluted in sterile PBS buffer and plated directly in egg yolk agar or oil/Rhodamine B agar [LB containing 2.5 % (v/v) olive oil and 0.001 % (w/v) Rhodamine B].
3. Spore-forming bacterium selection: Lagoon samples were diluted in sterile PBS buffer and incubated for 20 min at 80 °C or in the presence of one volume of chloroform before plating on egg yolk agar or oil/Rhodamine B agar.

A commercial inoculum (Glensol 3901—Glensol SA, Buenos Aires, Argentina) was used for comparative analysis. In this case, bacteria were isolated by plating on LB medium.

Chemical oxygen demand (COD), biochemical oxygen demand (BOD), fat, oil and grease (FOG, as ether soluble substances), and the pH of the wastewater were determined by the Centro de Ingeniería Sanitaria (Riobamba 245 bis, Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Rosario, Argentina).

Media and growth conditions

Cells were routinely grown in LB at 30 °C. Bacterial growth in liquid media was monitored by determining the optical density of the culture at wavelength 660 nm (OD_{660}; Beckman DU640 spectrophotometer). Bacterial growth profiles in dairy wastewater were performed in autoclaved dairy wastewater from the lagoon mid-depth diluted with one volume of a solution composed of 9.1 mM $(NH_4)_2SO_4$ and 67 mM KH_2PO_4 (pH 7) and supplemented with 1 % (v/v) olive oil or 0.5 % (w/v) glucose. Media were inoculated with 0.025 vol of each culture grown in LB for 6 h at 30 °C.

Exoenzyme activity assays

All activity assays were conducted by growing strains for 16 h at 30 °C. In order to detect lipolytic activity, isolates were grown on egg yolk agar. Clearing zones around the growing colonies evidenced lecithinase activity. Lipolytic activities were also evaluated using oil/Rhodamine B agar plates. Lipase-producing bacteria were detected by the orange fluorescence under UV light at 350 nm (Kouker and Jaeger 1987). Lipolytic activities in supernatants were determined as previously described with minor changes (Lonon and Hooke 1991). Briefly, the culture supernatant was mixed with equal volume of PBS buffer supplemented with 5 % (v/v) egg yolk and incubated for an hour at 37 °C. Then, the OD_{410} of an appropriate dilution of each reaction was determined.

In order to detect proteolytic activity, bacterial strains were grown on skimmed milk agar [LB containing 2 % (w/v) skimmed milk powder]. Protease activities were evident as a clearing zone around the growing colony. In order to detect α-amylase activities, bacterial strains were grown on starch agar [LB containing 1 % (w/v) starch]. The α-amylase producing bacteria were detected by flooding the plates with Lugol's iodine solution [1 % (w/v) iodine in 2 % potassium iodide]. The isolate showing a clear halo zone was considered as α-amylase-positive strain (Kouker and Jaeger 1987; Moreno Mde et al. 2009).

Random amplified polymorphic DNA (RAPD) analysis

DNA preparation and RAPD reactions were performed as previously described (Suárez et al. 2012). Briefly, reactions were carried out in a final volume of 25 µl, using 1 µl of DNA sample as template, 2.5 µmol l^{-1} of the RAPD19 primer (5′ GGTCGACYTTNGYNGGRTC 3′), 1U of Taq DNA polymerase, 200 µmol l^{-1} of each dNTP, 1.5 mmol l^{-1} $MgCl_2$ in a buffer solution containing 10 mmol l^{-1} Tris-HCl pH 8, and 50 mmol l^{-1} KCl. The mixture was subjected to 40 cycles at 93 °C for 60 s, 36 °C for 90 s and 72 °C for 120 s in a Techne thermocycler (Techgene). PCR products were run on a 2 % (w/v) agarose gel and stained with GelGreen nucleic acid stain (Biotium Inc., CA). The molecular weight of each band was estimated using the Gel-Pro analysis software and 100-bp DNA ladder (Invitrogen).

Genotypic identification of the isolates

The isolates were identified by 16S rRNA or GyrA coding gene sequencing. For this to be possible, *gyrA* and 16S rRNA genes were amplified by PCR using primers *gyrA*-Fwd (5′ GCDGCRGCNATGCGTTAYAC 3′) and *gyrA*-Rev (5′ AVRATYTCCATRCCKACRAC 3′), and 16S-341-Fwd (5′ CCTACGGGAGGCAGCAG 3′) and 16S-1389-Rev (5′ ACGGGCGGTGTGTACAAG 3′), respectively. The reaction mixtures were subjected to 30 cycles at 94 °C for 45 s, 50 °C for 45 s and 72 °C for 120 s. The reaction products were resolved by electrophoresis in a 0.8 % (w/v) agarose gel and purified with GFX PCR DNA and gel band purification kit (GE Healthcare). The sequences of the purified PCR products were obtained at the University of Maine DNA sequencing facility.

Average nucleotide identity (ANI) calculation and phylogenomic tree construction

ANI value was calculated as described by Repizo et al. (2014). In this case, the ANI calculator Web interface available at EzGenome was used (http://www.ezbiocloud.net/ezgenome/ani). Evolutionary analyses were conducted using *gyrA* or *oprI* and *oprL* gene sequences. *gyrA* sequences were obtained using TBLASTN tool (Zhang et al. 2000), all *Bacillus*-type strains available at EzGenome and WP_013350726.1 sequence as query. *oprI* and *oprL* sequences described in Matthijs et al. (2013) were obtained from Genbank, individually aligned and concatenated using the Perl script catfasta2phyml.pl (http://www.abc.se/~nylander/catfasta2phyml/). The best substitution pattern was computed based on Bayesian information criterion scores and the evolutionary histories inferred in MEGA6 (Tamura et al. 2013).

Genotypic clustering

Genotypic clustering of the isolates was implemented by the R package pvcluster (Suzuki and Shimodaira 2013). For the hierarchical cluster analysis, the presence or absence of each polymorphic band in the RAPD patterns was used as binary variable. The distance measurement was taken by Manhattan distance function, and the average option was used as agglomerative method.

Genome annotation and protein secretion prediction

B. safensis S9 and *P. alcaliphila* ED1 genomic sequences were determined with Illumina HiSeq System technology at Molecular Research LP (USA) and INDEAR (Argentina), respectively. *B. safensis* S9 sequence assembly was performed with SeqMan NGen 11.1.0, and an average coverage of 134. *P. alcaliphila* ED1 sequence assembly was performed with A5 pipeline software v2014-11-20 and an average coverage of 296. Sequences have been deposited at DDBJ/EMBL/GenBank under the accession numbers LIHF00000000 and LLXP00000000, respectively. The versions described in this paper are versions LIHF01000000 and LLXP01000000, respectively. Coding sequences (CDS) were identified and

Table 1 Exoenzyme patterns of the isolates

Exoenzyme activities	Number of strains	Strains
Amylase, lecithinase protease and lipase	14	EFA3, EFA4, EFA8, EFY3, FBA5, FBA12, FBA15, GSA3, GSA5, GSA13, GSA32, SAA10, SAY3 and SAY4
Amylase lecithinase protease	20	EF7, EFY1, EFY14, FB10, GS12, GSY1, GSY2, GSY7, S1, S2, S3, S4, S5, S6, S8, SA12, SA7, SA9, SAY1 and SAY2
Amylase protease	2	G13 and G24
Protease lipase	2	S7 and S9
Amylase lecithinase	3	FB8, GS13 and SA1
Lecithinase lipase	8	EFA5, EFA11, EFA13, EFA19, GSY9, ID9, ID10 and ID15
Lecithinase	7	E1, E2, E3, E4, E5, FBY1 and ID16
Lipase	4	ED1, GD1, Q1 and Q2
None	2	G20 and G21

Lecithinase, lipase, protease and amylase activities are detected growing strains for 16 h at 30 °C on egg yolk, oil/Rhodamine B, skimmed milk and starch agar plates, respectively. All activities are evaluated at least by triplicate. In Online Resource 1, exoenzyme patterns are indicated by strains

assigned using RAST (Aziz et al. 2008) and handled using the R package seqinR (Charif and Lobry 2007). CDS encoding features of interest were selected based on its annotation using grep and one of the following search patterns: acetolactate decarboxylase, amylase, azoreductase, catalase, cellulase, esterase, glucanase, glucose isomerase, glucose oxidase, glucosyltransferase, hydrolase, laccase, ligninase, lipase, mannanase, pectate lyase, pectinase, phytase, protease, proteinase, pullulanase or xylanase. Those CDS encoding for putative esterase, lipase, hydrolase, protease or proteinase were translated to amino acid sequences and their secretion fate predicted using SignalP 4.1, tatP 1.0 and secretomeP 2.0 tools (Bendtsen et al. 2004, 2005; Petersen et al. 2011).

Statistical analyses

The uncertainty of the hierarchical cluster analysis was assessed via AU (approximately unbiased) *p* value, which was computed by multiscale bootstrap resampling with 100,000 replications using the R package pvcluster (Suzuki and Shimodaira 2013).

Results

Isolation of lipolytic strains from dairy wastewater

In an attempt to isolate strains with the capability of degrading fats from dairy wastewater, samples from a treatment lagoon of an industrial dairy plant were collected. At the moment of sampling, the COD, BOD, FOG and pH of the wastewater were 760 mg O_2 ml^{-1}, 700 mg O_2 ml^{-1}, 1900 mg ml^{-1} and 6.7, respectively. Three different screening strategies were conducted to isolate bacteria with lipolytic activity.

First, an enrichment procedure in mineral medium supplemented with milk powder and oil was performed as described in Materials and methods. Once the enrichment operation was repeated three times, the resulting culture was diluted and plated onto egg yolk agar plates. All resulting colonies (~2500) showed lipolytic activity as well as similar morphology and color. Five strains were selected (E1-5) for further characterization.

Secondly, direct isolations of strains by plating different lagoon samples on egg yolk and oil/Rhodamine B agar plates were performed. Sixteen out of 77 strains (EFY1, EFY3, EFY14, FBY1, GSY1, GSY2, GSY7, GSY9, ID9, ID10, ID15, ID16, SAY1, SAY2, SAY3 and SAY4) had lipolytic activity in an egg yolk agar plate. On the other hand, when olive oil was used as a source of low rate biodegradable fats (olive oil has approximately twice the amount of triglycerides with long chain fatty acid as milk fat), 17 out of 166 isolated strains (ED1, EFA3, EFA4, EFA5, EFA8, EFA11, EFA13, EFA19, FBA5, FBA12, FBA15, GD1, GSA3, GSA5, GSA13, GSA32 and SAA10) showed lipolytic activity.

Spores represent an advantage in terms of handling, transportation and storage, and are suitable for bioaugmentation processes. Therefore, spore-forming bacteria were selected. In order to achieve that a pretreatment of the dairy wastewater samples was carried out by incubating them with one vol of chloroform or for 20 min at 80 °C. Nine and seven strains (EF7, FB8, FB10, GS12, GS13, SA1, SA7, SA9 and SA12; S1, S2, S3, S4, S5, S6 and S8), obtained from chloroform and heat treatments, respectively, were selected for their lipolytic activity in egg yolk agar, and four strains (Q1, Q2, S7 and S9) obtained from heat treatment were selected in oil/Rhodamine B agar.

Bacterium isolation from a commercial product (Glensol 3901) was also performed by plating onto LB agar.

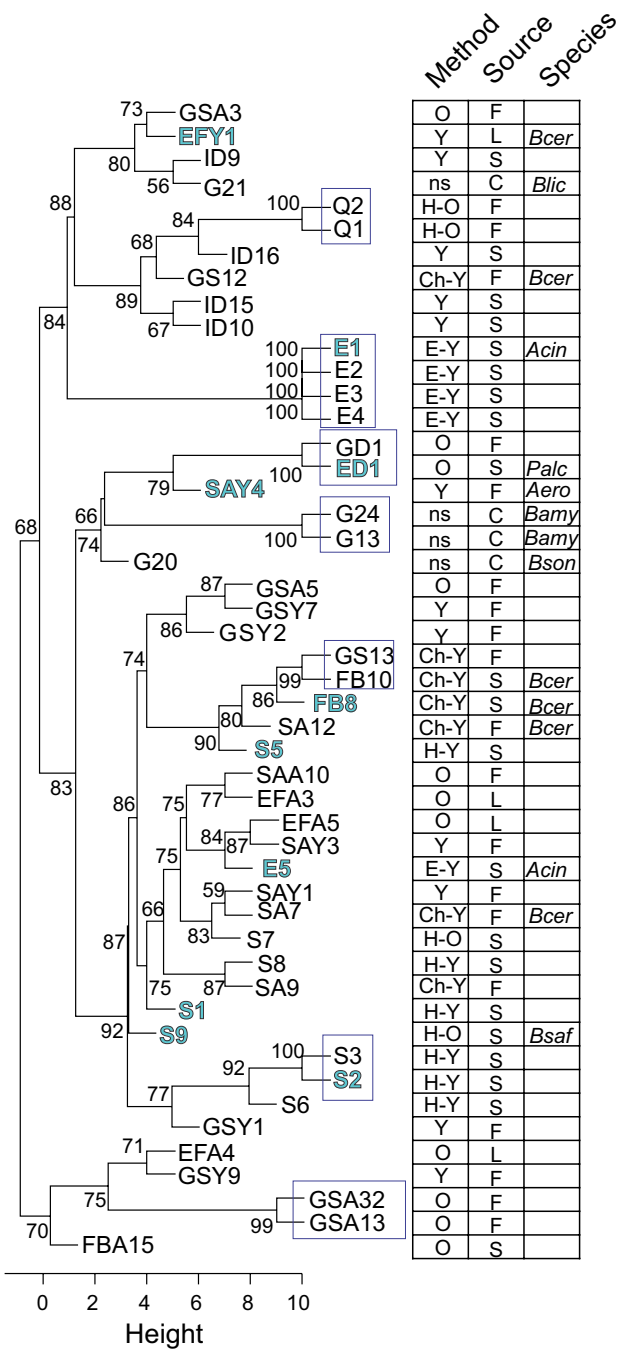


Fig. 1 Genetic clustering, source and selection method of isolates. Hierarchical cluster of the isolates are performed based on their polymorphic band pattern. AU *p* value is computed by multiscale bootstrap resampling with 100,000 replications. Those strains that cluster together with an AU *p* value >99 % are highlighted by *boxes*. Bacteria sources [paper-filtered liquid (L), sludge (S) and superficial solids (F) of the lagoon and a commercial product (C)] as well as the selection method used are indicated. Enrichment and plating on egg yolk (E-Y), direct plating on egg yolk (Y), or oil/Rhodamine B agar (O), chloroform treatment and plating on egg yolk (Ch-Y) and heat treatment and plating on egg yolk (H-Y) or oil/Rhodamine B agar (H–O)]. Strains selected for wastewater growth and lipolytic enzyme production studies are highlighted. ns, strains isolated in non-selective medium; Acin, *Acinetobacter* sp.; Aero, *Aeromonas* sp.; Bamy, *B. amyloliquefaciens*; Bcer, *B. cereus*; Blic, *B. licheniformis*; Bsaf, *B. safensis*; and Bson, *B. sonorensis*

Table 2 Growth rates and lipolytic activity of selected isolates

Strain	μ_{max} (h^{-1})		Lipolytic activity (min^{-1})	
	Glu	Oil	Glu	Oil
E1	0.012	0.068	1.87	0.95
E5	0.116	0.004	1.27	1.07
ED1	0.111	0.112	5.73	5.73
EFY1	0.110	0.069	25.21	6.81
FB8	0.126	0.071	14.97	2.49
S1	0.124	0.065	18.37	4.77
S2	0.018	0.008	1.97	0.43
S5	0.001	0.010	1.99	0.81
S9	0.093	0.087	2.69	4.81
SAY4	0.007	0.004	1.05	0.81

Strains are grown in the WW medium supplemented with the indicated carbon source and with agitation at 30 °C for seven days. Maximum bacterial growth rate (μ_{max}) and maximal lipolytic activity present in supernatant are indicated. Values are the mean of at least three determinations with no more than 5 % SD among them

Remarkably, none of the strains isolated (G13, G20, G21, G24) showed lipolytic activity on egg yolk or oil/Rhodamine B agar plates.

Finally, exoenzyme profiles were performed for the 58 lipolytic enzyme-producing strains (Table 1 and Online Resource 1). Thirty-four strains were found to have proteolytic, amylase and lipolytic activities. Twenty of them had lecithinase activity but did not hydrolyze oil.

Genetic and physiological characterization of the isolates

RAPD analyses were performed to evaluate the presence of different genotypes among 49 isolates from which adequate DNA quality and quantity could be obtained. As shown in Fig. 1, six pairs and a quartet of isolates belong to the same genotype ($p < 0.95$). Based on their morphology (not shown), RAPD and exoenzyme profiles, ten strains were selected to evaluate their capability to adapt to wastewater-specific environmental conditions. In order to do this, isolates were grown individually in a sterile wastewater medium (WW) supplemented with different carbon sources [0.5 % (w/v) Glucose (WW + Glu) or 1 % (v/v) Oil (WW + Oil)]. Lipolytic activities in supernatants and bacterial growth rate (μ) were monitored for seven days as described in Materials and methods. Isolates ED1, EFY1, FB8, S1 and S9 showed the highest μ_{max} in both media (Table 2). In turn, these strains also showed the highest lipolytic activities among the strains under study. Interestingly, while all strains showed a higher activity in WW + Glu, the highest activity for strain S9 was observed in WW + Oil.

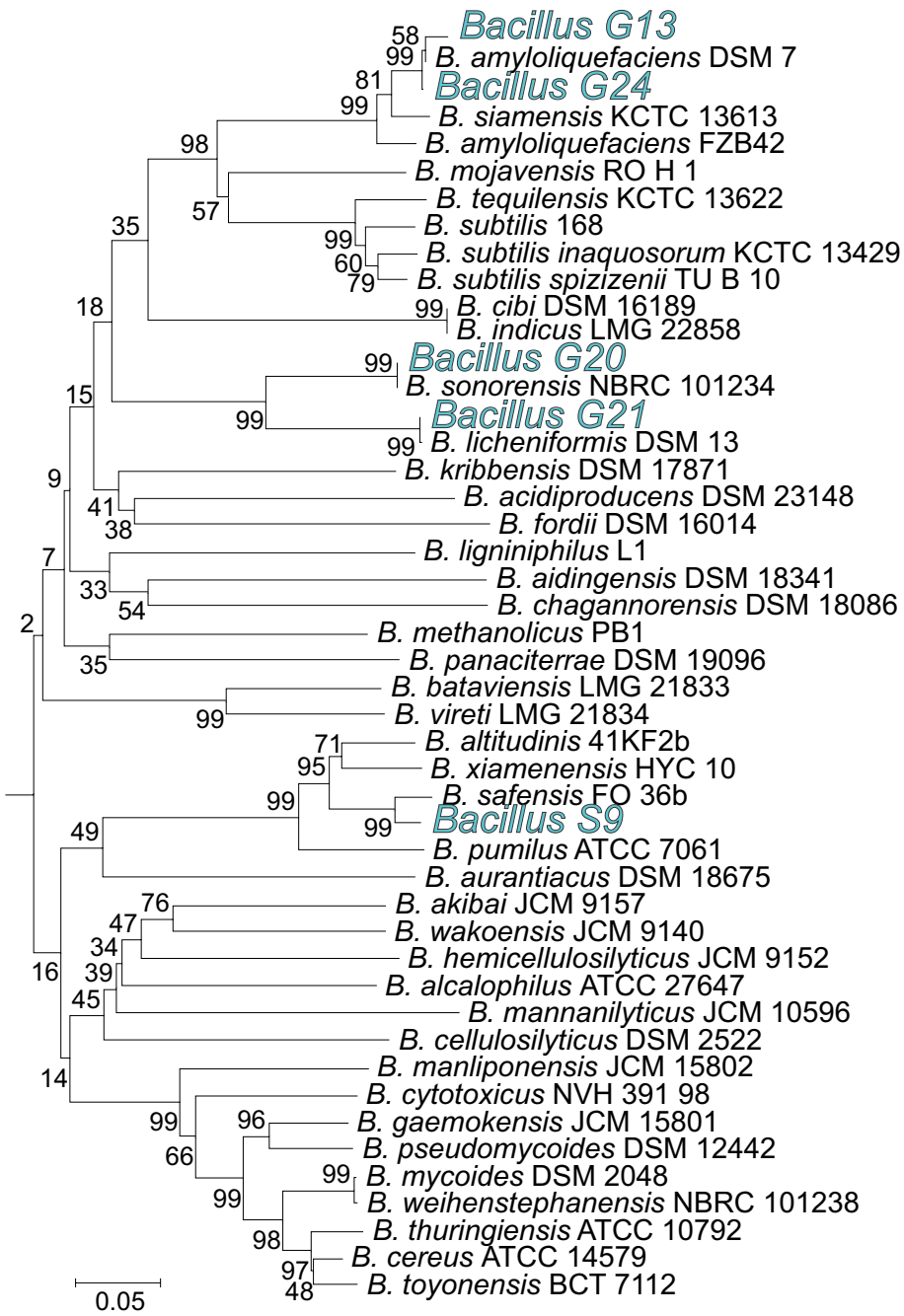


Fig. 2 Evolutionary analyses of *Bacillus* strains using *gyrA* gene sequence. The evolutionary history of the strains is inferred using the neighbor-joining method in MEGA6 (Tamura et al. 2013). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10,000 replicates) is shown next to the branches. The evolutionary distances of the subtree are computed using the Tamura-Nei method and are in the units of the number of base substitutions per site. The complete phylogenetic tree is shown in Online Resource 2

Taxonomic identification of selected isolates

Economic, regulatory and safety restrictions for the production of bacteria at high-scale levels and their application in industrial wastewater treatment require that those strains belong to nonpathogenic species (El Fantroussi and Agathos 2005). Therefore, 16S-rRNA sequences from selected isolates were determined and compared to NCBI databases using BLAST (Zhang et al. 2000). Isolates EFY1, FB8, FB10, GS12, SA7, SA12 and S9 were found to belong to the *Bacillus* genera. While the latter isolate shared 99 % identity with 16S rRNA gene of *B. safensis* and *B. pumilus* strains, the remaining six shared ≥99 % identity with *B. cereus* group strains. Interestingly, S9 was selected by its lipolytic activity in oil/Rhodamine B, whereas strains EFY1, FB8, FB10, GS12, SA7 and SA12 were selected in egg yolk agar plates, indicating that such procedure may favor the selection of *B. cereus* group strains. It has previously been reported that the sequence of gyrase A encoding gene (*gyrA*) enables a more accurately identification among *Bacillus* species (Reva et al. 2004). Therefore, *gyrA* sequences of G13, G20, G21, G24 and S9 were determined

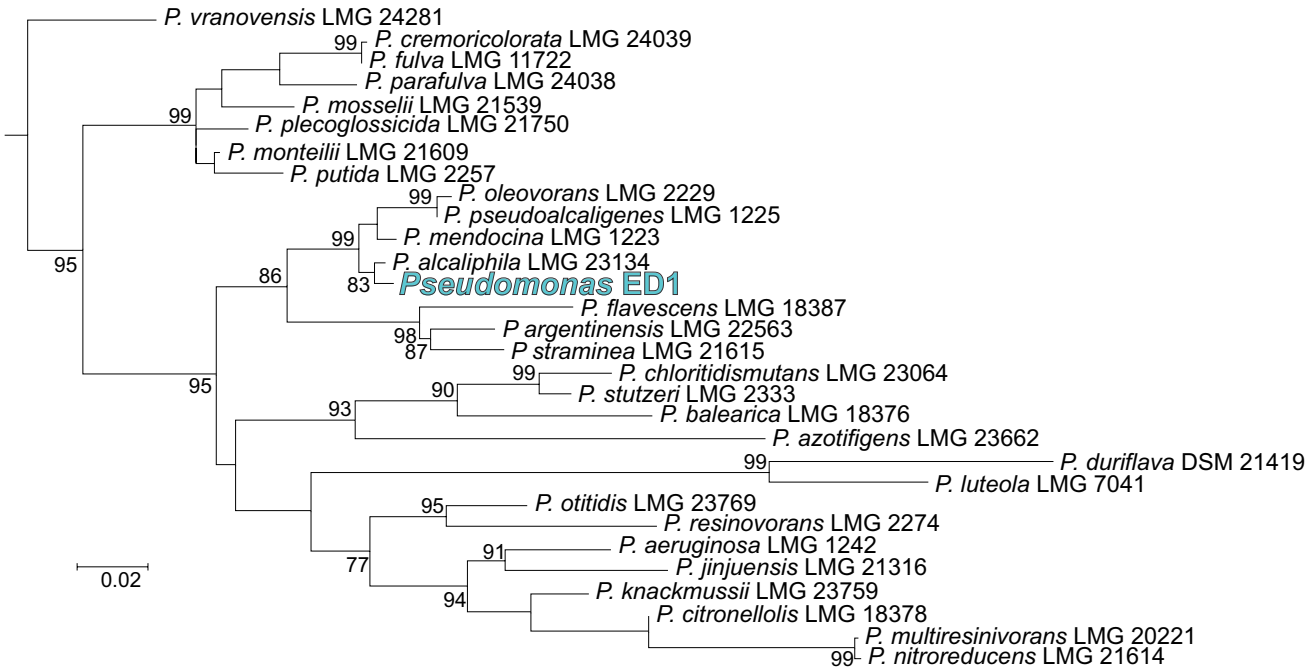


Fig. 3 Evolutionary analyses of *Pseudomonas* ED1 strain using the concatenated sequence of *oprI* and *oprL* genes. The evolutionary history of the strains is inferred by using the maximum likelihood method based on the general time reversible model in MEGA6 (Tamura et al. 2013). The tree with the highest log likelihood (−6882.9545) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The subtree is drawn to scale, with branch lengths measured in the number of substitutions per site. The complete phylogenetic tree is shown in Online Resource 3

and their phylogenomic relationship inferred. As shown in Fig. 2, G13 and G24 clustered with *B. amyloliquefaciens* strains and were more related to the *amyloliquefaciens* subspecies. G20 and G21 grouped together with members of *B. sonorensis* and *B. licheniformis* species, respectively. On the other hand, S9 clustered with strains of the *B. pumilus* group and was more related to the *B. safensis* FO-36b. *B. pumilus* group is composed of seven species that share over 99.5 % of 16S rRNA gene identity. The selection of marker genes that provide prokaryotic species boundaries at a resolution higher than 16S rRNA or *gyrA* is a challenging but necessary task to reconstruct the genealogy of this group of bacteria (Rossello-Mora 2012; Sunagawa et al. 2013). On the other hand, the average nucleotide identity (ANI) among shared genes was recently proposed for species circumscription (Richter and Rossello-Mora 2009). As expected, the ANI value between S9 and FO-36b was 96.1 %, which is slightly higher than the 95–96 % cutoff for species boundary (Richter and Rossello-Mora 2009).

The genes coding for 16S rRNA were also analyzed in isolates SAY4, E1, E5 and ED1. The former shared 99 % identity with *Aeromonas* spp. On the other hand, 16S rRNA encoding genes of E1 and E5 shared 99 % identity with *Acinetobacter* strains and ED1 shared 100 % identity with *Pseudomonas* spp. The sequence of *oprI* and *oprL*, coding for the outer membrane lipoprotein I (OprI) and the peptidoglycan-associated lipoprotein L (OprL), respectively, were proposed as specific molecular markers of the genera (Matthijs et al. 2013). Then, 89 *oprI* and *oprL* sequences from *Pseudomonas* spp. were aligned with the respective sequences of ED1 strain and their phylogenomic relationship inferred. This analysis indicated that ED1 is closely related to *Pseudomonas alcaliphila* species (Fig. 3), which is consistent with the fact that the ANI value between ED1 and *P. alcaliphila* 34 was 96.9 %.

Prospective exoenzymes identification in *B. safensis* S9 and *P. alcaliphila* ED1

In order to identify enzymes that contribute to the degradation of dairy wastewater components, the genome sequences of S9 and ED1 strains were determined. The draft genome sequences obtained consist of 3,794,315 and 5,239,535 bp and were assembled into 22 and 19 contigs, respectively. Genome sequences were automatically annotated using the RAST server (Aziz et al. 2008). The genome of strain S9 encoded a total of 3900 CDS and 90 structural RNAs (74 tRNAs), whereas ED1 encoded a total of 4781 CDS and 63 structural RNAs (58 tRNAs). A total of 1747 (45 %) and 2562 (54 %) genes were assigned to specific subsystem categories by RAST for S9 and ED1,

Table 3 Features encoded in S9 and ED1 genomes

Feature	Strains	
	S9	ED1
Cofactors, vitamins, prosthetic groups, pigments	204	318
Cell wall and capsule	140	193
Virulence, disease and defense	57	**138**
Potassium metabolism	7	21
Miscellaneous	30	29
Phages, prophages, transposable elements, plasmids	4	16
Membrane transport	76	**242**
Iron acquisition and metabolism	31	46
RNA metabolism	151	209
Nucleosides and nucleotides	111	112
Protein metabolism	158	293
Cell division and cell cycle	40	33
Motility and chemotaxis	93	177
Regulation and cell signaling	59	105
Secondary metabolism	4	5
DNA metabolism	97	132
Fatty acids, lipids and isoprenoids	102	172
Nitrogen metabolism	13	**43**
Dormancy and sporulation	**109**	4
Respiration	56	139
Stress response	86	197
Metabolism of aromatic compounds	6	**56**
Amino acids and derivatives	360	588
Sulfur metabolism	43	67
Phosphorus metabolism	24	50
Carbohydrates	446	368

The most noteworthy differences between ED1 and S9 cellular features are in bold

respectively. An overview of the genome features found in S9 and ED1 strains is provided in Table 3.

In accordance with the evidence of lipolytic activities found in plaque and in the presence of dairy wastewater (Tables 1, 2), both S9 and ED1 genome sequences contained several candidates for esterases or lipases (Fig. 4). SignalP (Petersen et al. 2011), tatP (Bendtsen et al. 2005) and secretomeP (Bendtsen et al. 2004) tools were used to predict which of these enzymes have secretion fate. Eventually, 12 out of 31 and 16 out of 66 lipolytic enzymes were predicted to be secreted for S9 and ED1, respectively (Fig. 4).

As S9 showed proteolytic activity in skimmed milk agar plates (Table 1), proteases and proteinases were also analyzed in its genome sequence. Thirty-five putative proteases and proteinases were identified, 20 out of which were predicted to be secreted (Fig. 4). Remarkably, 20 out of 33 putative proteases and proteinases were also identified

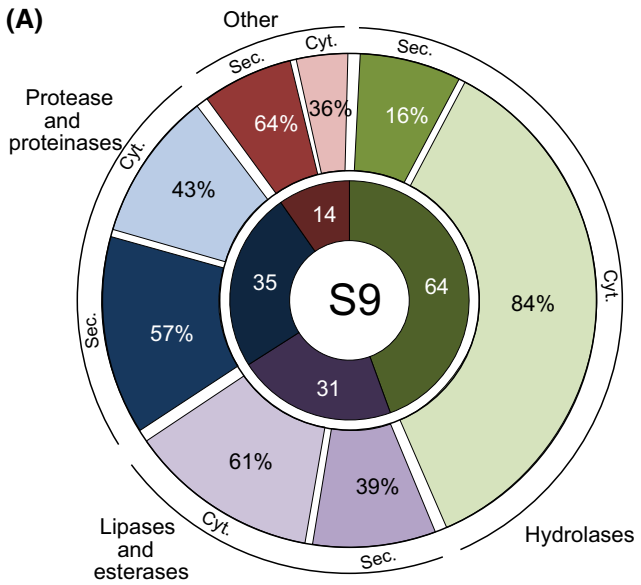
Fig. 4 Hydrolytic and biotechnological relevant enzyme repertories of S9 and ED1 strains. CDS in LIHF00000000 and LLXP00000000 sequences are identified and assigned using RAST. These assignations are used to construct circular plots for S9 (**a**) and ED1 (**b**) where the numbers of putative proteins with the depicted activities are indicated (*inner circles*). The percentage of protein predict to be secreted (Sec.) or cytoplasmic (Cyt.) is indicated in *outer circles*. For this prediction, SignalP, tatP and secretomeP are used. The RAST assignation of each putative lipolytic exoenzymes encoded in S9 or ED1 is described (with minor changes) below its corresponding circle, and the most interesting enzymes are highlighted. Other activities include amylase, glucanase, xylanase, pectate lyase, glucosyltransferase, azoreductase, catalase and acetolactate decarboxylase activities. See Online Resources 4 and 5 for further information

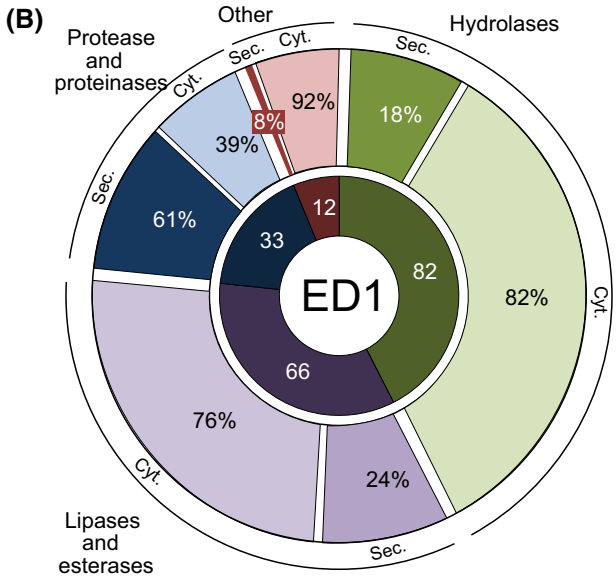
and predicted to be secreted in ED1 genome sequence, though its protease activity was not detected in our in vivo assays. To broaden our in silico study, hydrolase activity was searched in S9 and ED1 genomes, and as a result, 10 and 15 extracellular hydrolases were found, respectively (Fig. 4). The complete list of putative lipolytic, proteolytic and hydrolytic exoenzymes identified, as well as the secretion system involved, is described In Online Resource 4. Finally, biotechnological relevant enzymes were included in our analysis, and one amylase, three glucanases, two xylanases, one pectate lyase, three glucosyltransferases, seven azoreductases, eight catalases and one acetolactate decarboxylase were found (Table 4).

Discussion

In this study, three screening strategies were followed to isolate microorganisms that produce extracellular lipolytic enzymes and that grow in wastewater media. Strains were analyzed with respect to their RAPD and exoenzyme profiles finding diversity among the isolates. It is noteworthy that four groups of isolates (Q1 and Q2; E1, E2, E3 and E4; S2 and S3; and GSA13 and GSA32) that individually showed identical exoenzyme and RAPD patterns were selected by the same method and from the same sample types (Fig. 1). This may indicate that they are organisms with clonal origin. Conversely, strains ED1 and GD1, which also showed identical patterns, were isolated from different lagoon sample types (Fig. 1), suggesting that some microorganisms could be present in more than one lagoon section.

Regarding their taxonomy, strains from *B. safensis, P. alcaliphila, B. cereus, Aeromonas* and *Acinetobacter* were identified. *Pseudomonas* and *Bacillus* species with high-fat degradation effectiveness had been isolated from wastewater sources (Loperena et al. 2009). While the former genera were mainly associated with xenobiotic cleanup of industrial wastes (Golovleva et al. 1992), both were proposed to be used to improve wastewater treatment process





CDS number	Description of putative lipolytic exoenzymes encoded in *B. safensis* S9
229	2',3'-Cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16) / 5'-nucleotidase (EC 3.1.3.5)
248	Carbohydrate esterase family 4 protein
415	Polysaccharide deacetylase, Carbohydrate esterase family 4
1236	4-Hydroxybenzoyl-CoA thioesterase family active site containing protein
1339	Rhamnogalacturonan acetylesterase
1714	**Lysophospholipase-like family protein**
1869	**para-Nitrobenzyl esterase**
2156	**Lipase precursor**
2226	**Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)**
2726	**Hypothetical lysophospholipase**
3235	**Phospholipase/carboxylesterase**
3640	**Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)**

CDS number	Description of putative lipolytic exoenzymes encoded in *P. alcaliphila* ED1
178	**Lipase precursor (EC 3.1.1.3)**
242	**Patatin-like phospholipase**
379	**Lipase**
414	**Lipase**
506	**Lipase (EC 3.1.1.3)**
850	Diguanylate cyclase/phosphodiesterase with PAS/PAC sensors
985	**Phospholipase A1 precursor (EC 3.1.1.32, EC 3.1.1.4); Outer membrane phospholipase A**
1030	Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
1906	**Lipase class 3 family protein**
3326	**Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)**
3501	Diguanylate cyclase/phosphodiesterase with PAS/PAC sensors
3690	**Esterase**
3928	**Lysophospholipase (EC 3.1.1.5)**
3994	Diguanylate cyclase/phosphodiesterase with PAS/PAC sensors
4267	**Lipase precursor (EC 3.1.1.3)**
4272	**Arylesterase precursor (EC 3.1.1.2)**

Table 4 Putative biotechnological relevant enzymes encoded in *B. safensis* S9 and *P. alcaliphila* ED1

Strain	CDS number	Description
S9	3056	Alpha-acetolactate decarboxylase (EC 4.1.1.5)
	2470	FMN-dependent NADH-azoreductase
	2732	FMN-dependent NADH-azoreductase
	1880	Catalase (EC 1.11.1.6)
	3511	Catalase (EC 1.11.1.6)
	394	Catalase (EC 1.11.1.6)
	816	Manganese catalase (EC 1.11.1.6)
	1073	Endoglucanase 4 precursor (EC 3.2.1.4)
	1074	Exoglucanase II precursor (EC 3.2.1.91)
	2386	Endo-beta-1,3-1,4 glucanase (Licheninase) (EC 3.2.1.73)
	3013	Poly(glycerol-phosphate) alpha-glucosyltransferase (EC 2.4.1.52)
	3303	Pectate lyase(EC:4.2.2.2)
	1275	Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8)
	1354	Xylanase
ED1	23	Alpha-amylase (EC 3.2.1.1)
	1176	FMN-dependent NADH-Azoreductase
	1434	FMN-dependent NADH-Azoreductase
	1439	Azoreductase
	1586	FMN-dependent NADH-Azoreductase
	2931	FMN-dependent NADH-Azoreductase
	3	Catalase (EC 1.11.1.6)
	4065	Catalase (EC 1.11.1.6)/Peroxidase (EC 1.11.1.7)
	4662	Catalase (EC 1.11.1.6)
	681	Catalase (EC 1.11.1.6)
	1261	Glucans biosynthesis glucosyltransferase H (EC 2.4.1.-)
	2205	UDP-glucose:(heptosyl) LPS alpha1,3-glucosyltransferase (EC 2.4.1.-)

Data correspond to sequences LIHF01000000 and LLXP01000000, respectively

CDS of *B. safensis* S9 and *P. alcaliphila* ED1 are identified and assigned using RAST

(Dhall et al. 2012; Loperena et al. 2006, 2009). Consistently, strains G13, G20, G21 and G24, isolated from a commercial product (Glensol 3901), belong to *Bacillus* genera. Vrints et al. (2007) also studied bioaugmentation products showing that *Bacillus* were the dominant cultivable genus present in all 18 analyzed inoculants. However, none of the strains isolated from the commercial inoculum used as reference in this work showed lipolytic and/or proteolytic activities, highlighting the importance of using bacterial strains selected by its specific traits (i.e., lipolytic enzyme-producing strains) for wastewater treatment of certain industries.

While occasionally proposed to be used in wastewater treatment process (Ahmad et al. 2010; Lee et al. 2008; Wang et al. 2013; Yao et al. 2013), some strains of *B. cereus, Aeromonas* and *Acinetobacter* were shown to be human pathogens (Janda and Abbott 2010; Juni 1978; Panel 2005). Taking this into account, strains E1, E2, EFY1, FB8, FB10, GS12, SA7, SA12 and SAY4 should be carefully evaluated for their use. In general, caution should be taken when non-identified bacteria are intended to be used as inoculum in industrial processes. However, this seems not to be a common practice since the presence of *B. cereus* strains was detected in 16 out of the 18 products analyzed by Vrints et al. (2007).

A common characteristic of *Pseudomonas* is their considerable metabolic versatility (Rojo 2010). Concordantly, ED1 encodes for 166, 81, 30 and 50 more genes than S9 featured in membrane transport, defense (mainly metal resistance and multirug efflux pumps), nitrogen metabolism and aromatic compound metabolism, respectively (Table 4). Moreover, ED1 encodes for approximately twice as many genes for each cellular feature as S9. The exception was genes involved in dormancy and sporulation functions, as expected, but also those involved in carbohydrate metabolism (Table 4).

Finally, the whole-genome sequencing strategy performed in this study allows us to identify a repertoire of 68 technologically relevant putative proteins, including 28 lipolytic exoenzymes (Fig. 4; Table 4). Proteolytic

exoenzymes were identified in silico for ED1, whereas no protease activity was detected in vivo. This may indicate that these putative enzymes were not expressed or active at assayed conditions or that they were involved in physiological activities not related to extracellular protein hydrolysis. Further studies will help to understand the biochemistry and regulation of S9 and ED1 exoenzymes and, therefore, comprehend their contribution to the removal of fats and proteins from dairy wastewaters.

Acknowledgments This work was supported by grants from the Secretaría de Estado de Ciencia, Tecnología e Innovación, Provincial de Santa Fe [SL grant: Programa 1, Instrumento 1.1 (2009); SL and CM grant: Programa 1, Instrumento 1.3 (2010)] and SL and ME personal funding. We sincerely thank Drincovich M.F. for her assistance and encouragement at the initial stage of this study. This research paper would not have been possible without her support. We also thank Dr. Repizo G. for critical reading of the manuscript and the staff from the English Department (Facultad de Ciencias Bioquímicas y Farmacéuticas, UNR) for the language correction of the manuscript.

References

Ahmad WA, Zakaria ZA, Khasim AR, Alias MA, Ismail SM (2010) Pilot-scale removal of chromium from industrial wastewater using the ChromeBac system. Bioresour Technol 101:4371–4378. doi:10.1016/j.biortech.2010.01.106

Aziz RK et al (2008) The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. doi:10.1186/1471-2164-9-75

Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S (2004) Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel PEDS 17:349–356. doi:10.1093/protein/gzh037

Bendtsen JD, Nielsen H, Widdick D, Palmer T, Brunak S (2005) Prediction of twin-arginine signal peptides. BMC Bioinform 6:167. doi:10.1186/1471-2105-6-167

Cammarota MC, Freire DM (2006) A review on hydrolytic enzymes in the treatment of wastewater with high oil and grease content. Bioresour Technol 97:2195–2210. doi:10.1016/j.biortech.2006.02.030

Charif D, Lobry JR (2007) A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Bastolla U, Porto M, Roman HE, Vendruscolo M (eds) Structural approaches to sequence evolution: molecules, networks, populations. Springer, Berlin, pp 207–232. doi:10.1007/978-3-540-35306-5_10

Dhall P, Siddiqi TO, Ahmad A, Kumar R, Kumar A (2012) Restructuring BOD: COD ratio of dairy milk industrial wastewaters in BOD analysis by formulating a specific microbial seed. Sci World J 2012:105712. doi:10.1100/2012/105712

EFSA panel (2005) Opinion of the Scientific Panel on biological hazards (BIOHAZ) on *Bacillus cereus* and other *Bacillus* spp in foodstuffs. EFSA J. doi:10.2903/j.efsa.2005.175

El Fantroussi S, Agathos SN (2005) Is bioaugmentation a feasible strategy for pollutant removal and site remediation? Curr Opin Microbiol 8:268–275. doi:10.1016/j.mib.2005.04.011

Golovleva LA, Aliyeva RM, Naumova RP, Gvozdyak PI (1992) Microbial bioconversion of pollutants. Rev Environ Contam Toxicol 124:41–78

Heaven MW et al (2012) Physicochemical properties and trace organic compounds in a dairy processor's aerobic bioreactor. Bioresour Technol 124:119–128. doi:10.1016/j.biortech.2012.08.005

Janda JM, Abbott SL (2010) The genus *Aeromonas*: taxonomy, pathogenicity, and infection. Clin Microbiol Rev 23:35–73. doi:10.1128/CMR.00039-09

Juni E (1978) Genetics and physiology of *Acinetobacter*. Annu Rev Microbiol 32:349–371. doi:10.1146/annurev.mi.32.100178.002025

Kouker G, Jaeger KE (1987) Specific and sensitive plate assay for bacterial lipases. Appl Environ Microbiol 53:211–213

Lee C, Kim J, Shin SG, Hwang S (2008) Monitoring bacterial and archaeal community shifts in a mesophilic anaerobic batch reactor treating a high-strength organic wastewater. FEMS Microbiol Ecol 65:544–554. doi:10.1111/j.1574-6941.2008.00530.x

Lonon MK, Hooke AM (1991) A microassay for the quantitative measurement of egg yolk-reactive enzyme activity produced by *Pseudomonas cepacia*. Curr Microbiol 23:81–84

Loperena L, Saravia V, Murro D, Ferrari MD, Lareo C (2006) Kinetic properties of a commercial and a native inoculum for aerobic milk fat degradation. Bioresour Technol 97:2160–2165. doi:10.1016/j.biortech.2005.09.032

Loperena L et al (2007) Performance of a commercial inoculum for the aerobic biodegradation of a high fat content dairy wastewater. Bioresour Technol 98:1045–1051. doi:10.1016/j.biortech.2006.04.030

Loperena L et al (2009) Isolation and selection of native microorganisms for the aerobic treatment of simulated dairy wastewaters. Bioresour Technol 100:1762–1766. doi:10.1016/j.biortech.2008.09.056

Matthijs S et al (2013) Evaluation of *oprI* and *oprL* genes as molecular markers for the genus *Pseudomonas* and their use in studying the biodiversity of a small Belgian River. Res Microbiol 164:254–261. doi:10.1016/j.resmic.2012.12.001

Moreno Mde L, Garcia MT, Ventosa A, Mellado E (2009) Characterization of Salicola sp. IC10, a lipase- and protease-producing extreme halophile. FEMS Microbiol Ecol 68:59–71. doi:10.1111/j.1574-6941.2009.00651.x

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786. doi:10.1038/nmeth.1701

Rajeshkumar K, Jayachandran K (2004) Treatment of dairy wastewater using a selected bacterial isolate, *Alcaligenes* sp. MMRR7. Appl Biochem Biotechnol 118:65–72

Repizo GD, Espariz M, Blancato VS, Suarez CA, Esteban L, Magni C (2014) Genomic comparative analysis of the environmental *Enterococcus mundtii* against enterococcal representative species. BMC Genomics 15:489. doi:10.1186/1471-2164-15-489

Reva ON, Dixelius C, Meijer J, Priest FG (2004) Taxonomic characterization and plant colonizing abilities of some bacteria related to *Bacillus amyloliquefaciens* and *Bacillus subtilis*. FEMS Microbiol Ecol 48:249–259. doi:10.1016/j.femsec.2004.02.003

Richter M, Rossello-Mora R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 106:19126–19131. doi:10.1073/pnas.0906412106

Rojo F (2010) Carbon catabolite repression in Pseudomonas: optimizing metabolic versatility and interactions with the environment. FEMS Microbiol Rev 34:658–684. doi:10.1111/j.1574-6976.2010.00218.x

Rossello-Mora R (2012) Towards a taxonomy of Bacteria and Archaea based on interactive and cumulative data repositories. Environ Microbiol 14:318–334. doi:10.1111/j.1462-2920.2011.02599.x

Suárez C, Repizo G, Espariz M, Blancato BS, Magni C, Alarcón S (2012) Detection and identification of *tyrDC*$^+$ enterococcal strains from pasteurized commercial cheeses. Food Sci Biotechnol 21:603–606

Sunagawa S et al (2013) Metagenomic species profiling using universal phylogenetic marker genes. Nat Methods 10:1196–1199. doi:10.1038/nmeth.2693

Suzuki R, Shimodaira H (2013) pvclust: Hierarchical Clustering with P-values via Multiscale Bootstrap Resampling, R package version 1.2.2 edn. http://www.sigmath.es.osaka-u.ac.jp/shimo-lab/prog/pvclust/

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. doi:10.1093/molbev/mst197

Venkata Mohan S et al (2009) Bioaugmentation of microbial communities in laboratory and pilot scale sequencing batch biofilm reactors using the TOL plasmid. Bioresour Technol 100:1746–1753. doi:10.1016/j.biortech.2008.09.048

Vidal G, Carvalho A, Méndez R, Lema JM (2000) Influence of the content in fats and proteins on the anaerobic biodegradability of dairy wastewaters. Bioresour Technol 74:231–239

Vrints M, Bertrand S, Collard JM (2007) A bacterial population study of commercialized wastewater inoculants. J Appl Microbiol 103:2006–2015. doi:10.1111/j.1365-2672.2007.03444.x

Wang JH et al (2013) Bioaugmentation of activated sludge with *Acinetobacter* sp. TW enhances nicotine degradation in a synthetic tobacco wastewater treatment system. Bioresour Technol 142:445–453. doi:10.1016/j.biortech.2013.05.067

Yao S, Ni J, Ma T, Li C (2013) Heterotrophic nitrification and aerobic denitrification at low temperature by a newly isolated bacterium, *Acinetobacter* sp. HA2. Bioresour Technol 139:80–86. doi:10.1016/j.biortech.2013.03.189

Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol J Comput Mol Cell Biol 7:203–214. doi:10.1089/10665270050081478