FEMS Microbiology Letters, 364, 2017, fnx159

doi: 10.1093/femsle/fnx159
Advance Access Publication Date: 25 July 2017
Research Letter

RESEARCH LETTER – Biotechnology & Synthetic Biology

Characterization of the contaminant bacterial communities in sugarcane first-generation industrial ethanol production

Maria L. Bonatelli, Maria C. Quecine, Mariana S. Silva and Carlos A. Labate*

Laboratório Max Feffer de Genética de Plantas, Departamento de Genética, Escola Superior de Agricultura 'Luiz de Queiroz,' Universidade de São Paulo, Av. Pádua Dias 11, CP 83, Piracicaba, SP 13400-970, Brazil

*Corresponding author: Departamento de Genética, Universidade de São Paulo, Av. Pádua Dias 11, CP 83, Piracicaba, SP 13400-970, Brazil. Tel: 551934354755; E-mail: calabate@usp.br

One sentence summary: Metagenomics applied to ethanol distilleries.
Editor: Michael Sauer

ABSTRACT

The industrial ethanolic fermentation process is operated in distilleries, either in fed-batch or continuous mode. A consequence of the large industrial ethanol production is bacterial contamination in the fermentation tanks, which is responsible for significant economic losses. To investigate this community, we accessed the profile of bacterial contaminant from two distilleries in Brazil, each operating a different fermentation mode, throughout sugarcane harvest of 2013–2014. Bacterial communities were accessed through Illumina culture-independent 16S rDNA gene sequencing, and qPCR was used to quantify total bacteria abundance. Both ethanol production modes showed similar bacterial abundance, around 10^5 gene copies/mL. 16S rDNA sequencing showed that 92%–99% of the sequences affiliated to *Lactobacillus* genus. Operational taxonomic units differently represented belonged mainly to *Lactobacillus*, but also to *Weissella*, *Pediococcus*, *Acetobacter* and *Anaeosporobacter*, although in lower abundance. Alpha-diversity only showed a correlation through the fermentation tanks in continuous mode, where it was always higher in the second and third tanks. Beta-diversity clearly separated the two distilleries and metagenome prediction reinforces clusterization within distilleries. Despite certain variations between bacterial community in the distilleries throughout harvest season, *Lactobacillus* were the main genera reported in both distilleries and bacterial community seemed to persist along time, suggesting bacterial reinfestation.

Keywords: fermentation; bacterial contamination; 16S rDNA; diversity; qPCR

INTRODUCTION

Ethanol is the most used biofuel for transportation, and Brazil is the second largest producer in the world, United States being the main producer. In 2016, Brazil produced 7295 million of gallons (RFA 2017) that were mainly consumed in the internal market. Sugarcane is the feedstock used in the fermentation process in distilleries in Brazil, and it contains considerable amounts of readily fermentable sugars, facilitating the operational processes and decreasing the costs (Amorim, Basso and Lopes 2009; Crago *et al.* 2010). In first-generation ethanol production, these sugars—sucrose, fructose and glucose—are extracted from sugarcane and used as substrate to fermentation (de Souza Dias *et al.* 2015).

The current fermentation process to ethanol production was developed in the 1930s by Firmino Boinot, which is known as

Received: 27 February 2017; **Accepted:** 24 July 2017

Melle-Boinot process and consists in fed-batch process with high yeast cell density (10%–15% w/v) that are responsible for a very short fermentation time of 6–10 h. After the fermentation process, yeast cells are treated with dilute sulfuric acid and then recycled into another batch. In Brazil, this process uses sugarcane juice and/or water-diluted molasses as substrate to produce sugarcane must and ferment (Godoy *et al.* 2008; Basso *et al.* 2011). With the improvements made in the last 30 years, this process achieves fermentation yields of 92%–93% (Amorim *et al.* 2011).

In 1970s appeared the first continuous versions of the process developed by Boinot. In Brazil, around 85% of distilleries adopt fed-batch fermentation process, while 15% adopt continuous fermentation system (Godoy *et al.* 2008). Initially, several operational problems were detected in continuous system, as elevated level of contamination, low productivity, low yield and solid flow. Nowadays, continuous process has been optimized to achieve high productivity, high process flexibility and stability, which can make it less expensive than batch processes (Zanin *et al.* 2000; Brethauer and Wyman 2010).

Due to high organic and inorganic compounds, the sugarcane fermentation tanks are susceptible to the growth of contaminant microorganisms. The impact of contaminant microorganisms in the fermentation process is related to sugar consumption and organic acids production, which affect the efficiency of yeast fermentation (Solomon 2009).

Lactic acid bacteria are contaminants commonly found in fermentation tanks, but other genera are also present, although in lower abundance. Not all microorganisms are able to grow in this environment, the fermentation process can be stressful to them due to factors such as low pH, high ethanol concentration, elevated temperature, high osmotic pressure and others (Basso *et al.* 2011). Several works have focused on accessing the bacterial diversity in this particular microbiota, but based on cultured techniques (Gallo 1990; Skinner and Leathers 2004; Lucena *et al.* 2010; Rich *et al.* 2015).

Nonetheless, accessing the bacterial community only through cultivated microorganism usually misrepresents the community present in an environment (Rappé and Giovannoni 2003). Illumina MiSeq 16S rDNA sequencing is a robust, low cost and easy to process methodology that can be used in different works to describe diversity from different bacterial community (Bartram *et al.* 2011). Also, the sequences generated can be analyzed using the Phylogenetic Investigation Communities by Reconstruction of Unobserved States (PICRUSt) that can predict gene functions comparing these sequences with a reference genomes database (Langille *et al.* 2013).

This work focused on the cultured-independent assessment and characterization of contaminating bacterial community from sugarcane ethanol fermentation process. To this purpose, we sampled from the two fermentation processes adopted in Brazil, fed-batch and continuous fermentation process, in the state of São Paulo during the harvest season of 2012–2013. To our knowledge, this is the first time that high throughput sequencing technique was used to study bacterial contaminant community from sugarcane fermentation tank.

MATERIALS AND METHODS

Samples collection

Sampling were made in two distilleries at Piracicaba region, São Paulo state, Brazil (22°43′31″ S, 47°38′57″ W) during the sugarcane harvest season of 2012–2013. The continuous process distillery (CPD) operates with four fermentation tanks connected in parallel, using sugarcane molasses, while the fed-batch distillery (FBD) uses sugarcane juice in four independent tanks. Both distilleries use sulfuric aqueous solution and antibiotics to reduce bacterial contamination. Three samplings were made between the months of August and December in each fermentation tank, at each tank designed sampling point, in each distillery. In the CPD, four tanks in parallel were collected and they were sequentially enumerated, whereas in the FBD, four independent tanks were collected (Table 1 and Fig. 1). All samples were stored at –80°C for further analysis.

Table 1. Temporal samples from FBD and CPD.

	FBD	CPD
First sampling	FBD1.1	CPD1.1
	FBD1.2	CPD1.2
	FBD1.3	CPD1.3
	FBD1.4	CPD1.4
Second sampling	FBD2.1	CPD2.1
	FBD2.2	CPD2.2
	FBD2.3	CPD2.3
	FBD2.4	CPD2.4
Third sampling	FBD3.1	CPD3.1
	FBD3.2	CPD3.2
	FBD3.3	CPD3.3
	FBD3.4	CPD3.4

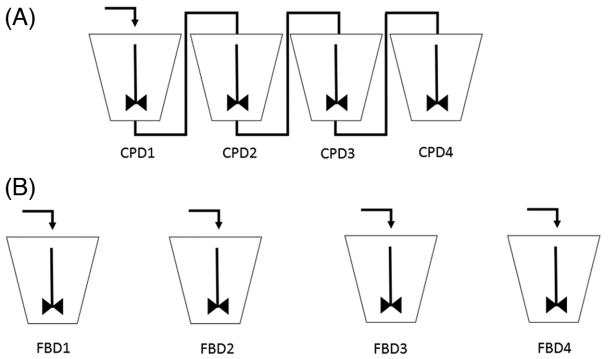

Figure 1. Diagram of sampling in fermentation tanks in the continuous process distillery (CPD) (**A**) and fed-batch distillery (FBD) (**B**).

DNA extraction

We used 1 mL of the fermentation samples to extract total DNA by Power Soil DNA Isolation kit (MO BIO, EUA) according to manufacturer's instructions.

Bacterial abundance

In the qPCR assay, the primer set P1 (5′-CCTACGGGAGGCAGCAG-3′) and P2 (5′-ATTACCGCGGCTGCTGG-3′) (Muyzer, De Waal and Uitterlinden 1993) was used to quantify bacterial abundance in CPD and FBD. The qPCR was performed with the iCycler iQ Real-Time PCR Detection System (Bio-Rad Laboratories Inc., USA) using the SYBR green I fluorescent dye detection. Amplifications were carried out in 25 μL reaction volumes containing 5 pmol of each primer (P1/P2), 12.5 μL of Platinum Syber Green qPCR SuperMix-UDG (Invitrogen), 0.02 mg.mL^{-1} of BSA and 5 ng of template DNA (1 ng.μL^{-1}). Thermal cycling parameters were set

initially in preheat for 5 min at 95°C, followed by 35 cycles at 95°C for 30 s, 62.5°C for 30 s and 72°C for 30 s for data collection and real time analysis, and 72°C –95°C for 30s increasing set point temperature after cycle 2 by 0.5°C for melt curve data collection and analysis.

Standard curves were constructed using the PCR product of the 16S rDNA gene. Amplification of 16S rDNA was carried out using the primer set R1387 (5′-CGGTGTGTACAAGGCCCGGGAACG-3′) e PO27F (5′-GAGAGTTTGATCCTGGCTCAG-3′) according to Heuer *et al.* (1997). The 16S rDNA amplified fragments were purified with polyethylene glycol methodology (Lis 1980) and the concentration measured at 260 nm with NanoCell (Thermo Scientific). Serial dilutions were performed and 10^2, 10^3, 10^4, 10^5, 10^6, 10^7 and 10^8 gene copies were used for calibration. Standard curve and all samples were run in triplicates. The average and the standard error of the threshold cycle obtained in the technical replicates were calculated and used as threshold cycle values to quantify total bacteria in each serial dilution and samples.

Bacterial 16S rDNA sequencing

The V4 region 16S rDNA was amplified using four primers forward: 16SV4FPCR 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGAY TGGGYDTAAAGNG-3′; 5′-TCGTCGGCAGCGTCAGATGT GTATAAGAGACAG NAYTGGGYDTAAAGNG-3′; 5′-TCGTCGGC AGCGTCAGATGTGTATAAGAGA CAGNNAYTGGGYDTAAAGNG-3′; 5′-TCGTCGGCAGCGTCAGATGTGTATAA GAGACAGNNNAYTG GGYDTAAAGNG-3′ and four primers reverse: 16SV4RPCR 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCGTCA ATTC MTTTRAT-3′; 5′- GTCTCGTGGGCTCGGAGATGTGTATAAGAGAC AGT CMTTTRAGT-3′; 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGNNN CCGTCAATTCMTTTRAGT-3′ based on Ribosomal Database Project (Cole *et al.* 2014). PCR mix contained 0.1 μM of each primer, 3.0 mM of $MgCl_2$, 0.3 mM of each dNTP, 10.0 $\mu g.\mu L^{-1}$ of BSA 1.5 U of *Taq* DNA polymerase, 1.2x *Taq* buffer with 25 μL of final volume. 16S rDNA amplification was initiated with a denaturation time of 3 min at 95°C, followed by 35 cycles of 45 s denaturing at 95°C, 1 min primer annealing at 57°C, 45 s extension at 72°C and final extension of 10 min at 72°C. Amplifications were conducted in Thermal cycler GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA).

Amplified 16S rDNA was sequenced in an Illumina Miseq according to manufacturer's instructions at the Functional Genomics Facility at the College of Agriculture Luiz de Queiroz—University of São Paulo, SP, Brazil.

Sequence analysis

Initially, paired-end Illumina Miseq sequences were joined with FLASh using default parameters: minimum overlap of 10 bp, maximum overlap of 65 bp, maximum mismatch density of 0.25 and combiner threads of 4 (Magoč and Salzberg 2011). Sequence analyses were made using QIIME (Caporaso *et al.* 2010). Fastq sequences were demultiplexed and filtered considering phred quality threshold of Q20. Operational taxonomic units (OTU) were generated with open reference OTU-picking script, considering USEARCH v 6.1 (Edgar 2010) method, using Greengenes database of May 2013, enabling reverse strand match and suppressing step four. Taxonomical analysis used Ribosomal Database Project Classifier training set No. 16 (Cole *et al.* 2014) as reference and UCLUST method (Edgar 2010) considering a minimum consensus fraction of 0.7. Alpha- and beta-diversity were estimated using subsampled OTU table that generated diversity and richness indexes—Shannon and Chao1, respectively. For beta-diversity was considered weighted-Unifrac distance to obtain three-dimensional principal coordinates analysis. OTU with statistically significant differences between sample groups was generated after filtering from OTU table OTUs with low abundance (>0.005%) (Bokulich *et al.* 2013).

Metagenome prediction

Prediction of genes families abundances was made with PICRUSt (Langille *et al.* 2013). OTUs were picked with closed reference from May 2013 Greengenes. The accuracy of metagenome prediction as measured with Nearest Sequence Taxon Index. STAMP v. 2.1.3 was used to generate principal coordinates analysis bidimensional distance matrix and gene classes comparisons (Parks *et al.* 2014).

RESULTS

Abundance of bacterial communities in sugarcane fermentation distilleries

The qPCR applied to determine the bacterial abundance in fermentation tanks from CPD and FBD revealed similar values between the distilleries. Bacterial abundance ranged from 10^5 to 10^6 gene copies/mL in CPD (Fig. 2A) and from 10^4 to 10^6 gene copies/mL in FBD (Fig. 2B), but mainly stayed around 10^5 gene copies/mL in both distilleries.

Culture-independent assessment of bacterial diversity

Paired-end sequences, after joined, resulted in median sequence read length of 367 bp. After filtration based on quality, it was obtained sequences that were clustered into 644 OTUs. Samples were subsampled to 54 844 sequences to prevent the bias that may be generated from different samples sizes.

Interestingly, in all samples from CPD and FBD more than 98% of sequences affiliated to Firmicutes phylum. However, it was also possible to find sequences belonging to Actinobacteria, Aquificae, Bacteroidetes, Cyanobacteria/Chloroplast, Deinococcus-Thermus and Proteobacteria (Fig. 3). It was reported more than 35 genera of which 92% up to 99% of sequences affiliated to the genus *Lactobacillus* in the samples (Fig. 4).

Alpha-diversity indicated that CPD presented higher value of richness (Chao1) when compared with FBD (Fig. 5A), but both presented similar values of diversity (Shannon) (Fig. 5B). Within distilleries, alpha-diversity varied among fermentation tanks in CPD and FBD along the time. Interestingly, in FBD, each fermentation tank seemed to have a particular richness and diversity, varying between 340.49 until 435.41 in Chao1 and 2.80 until 6.45 in Shannon index. In CPD, samples from the same sampling time exhibit a pattern in alpha-diversity through the fermentation tanks, where richness and diversity indexes increased from the first fermentation tank, reaching higher values in the second and third tanks, falling in the fourth fermentation tank (Table 2).

Despite of the high similarity found in taxonomic-based analysis, weight-based UniFrac beta-diversity analysis separated the samples from CPD and FBD, suggesting that there are differences in bacterial communities between distilleries (Fig. 6A). Most of the differentially represented OTUs belong to the genera *Lactobacillus* (Fig. 6B). Other genera that helped to explain this difference were *Weissella* and *Pediococcus* that were more abundant in FBD, while *Acetobacter* and *Anaeosporobacter* were more abundant in CPD.

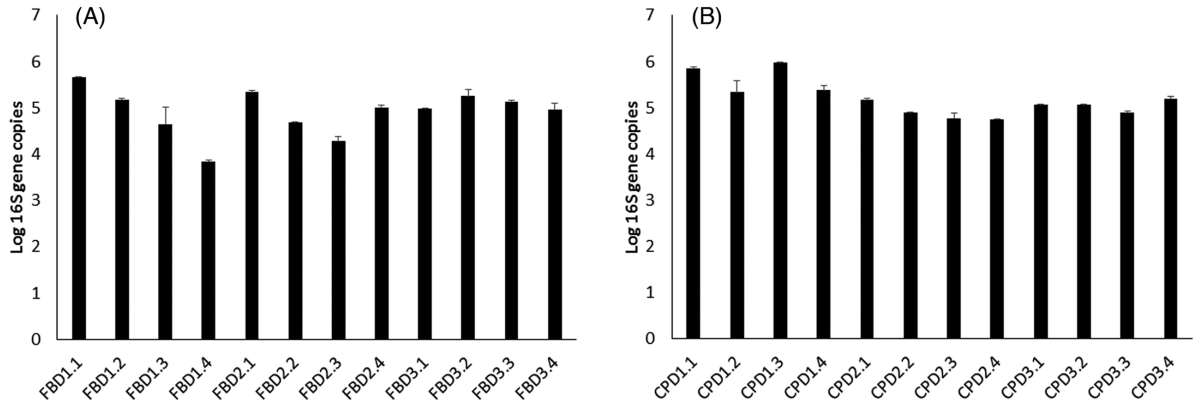

Figure 2. Bacterial abundance in fermentation tanks from fed-batch distillery (FBD) (**A**) and continuous process distillery (CPD) (**B**). The bars show the mean and error bars show the standard deviation of bacterial abundance of 3 replicates.

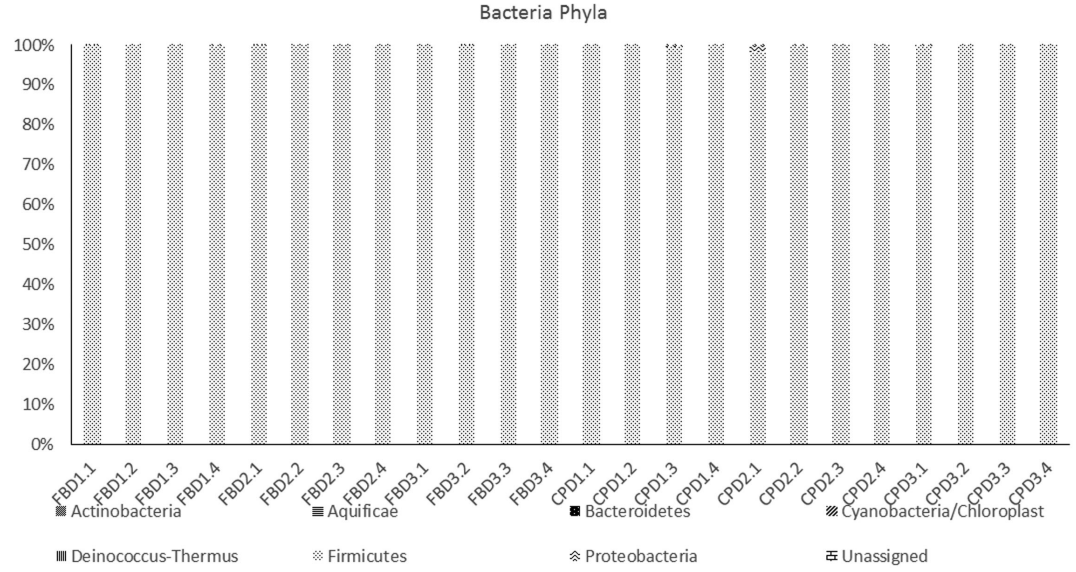

Figure 3. Taxonomic affiliation of 16S rDNA gene sequences in phyla against the Ribosomal Database Project database from fermentation tanks in the fed-batch distillery (FBD) and continuous process distillery (CPD) using QIIME.

Predicted metagenome

PICRUSt was used as a predictive exploratory tool, and samples were analyzed at level 2 of KEGG Orthology groups. Nearest Sequence Taxon Index values from FBD were 0.03 ± 0.002 and from CPD were 0.03 ± 0.006, which are considered very low Nearest Sequence Taxon Index values, indicating ideal data sets to examine prediction from PICRUSt (Table 2). Interestingly, PCA plot assembled with gene family reinforced the clustering of samples collected in each distillery (Fig. 7A). It was possible to predict 29 gene families, whereas 20 of them showed statistically significant differences (test-t $P < 0.05$). Important gene families related to microbial metabolism in sugarcane ethanolic fermentation, as carbohydrate metabolism and membrane transport, were well represented, also included in those families that show statistically significant differences (Fig. 7B).

DISCUSSION

Bacterial contamination is often regarded as a major drawback during industrial ethanol fermentation (Basso *et al.* 2011). Nonetheless, studies of this community are usually made through cultured-based assessment, which can often underestimate the community composition. In fact, cultured-independent assessment of contaminant bacterial community reported several bacterial genera that have never been reported in sugarcane fermentation tanks, and they belong to phyla Actinobacteria, Aquificae, Bacteroidetes, Cyanobacteria/Chloroplast, Deinococcus-Thermus, Firmicutes and Proteobacteria showing that the diversity in this environment can be higher than described (Rosales 1989; Gallo 1990, Lucena *et al.* 2010, Rich *et al.* 2015). Costa *et al.* (2015); accessing by uncultured-based methodology, the bacterial diversity of various stages of a sugarcane distillery also reported phyla Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria as most abundant in the process, but this work did not consider any fermentation tank.

Regarding bacterial abundance, our results showed that bacterial abundance are around 10^5 gene copies/mL in the fermentation tanks in both distilleries. Even though bacterial contamination may cause losses in ethanol production, the fermentation process can operate satisfactory with a low level of

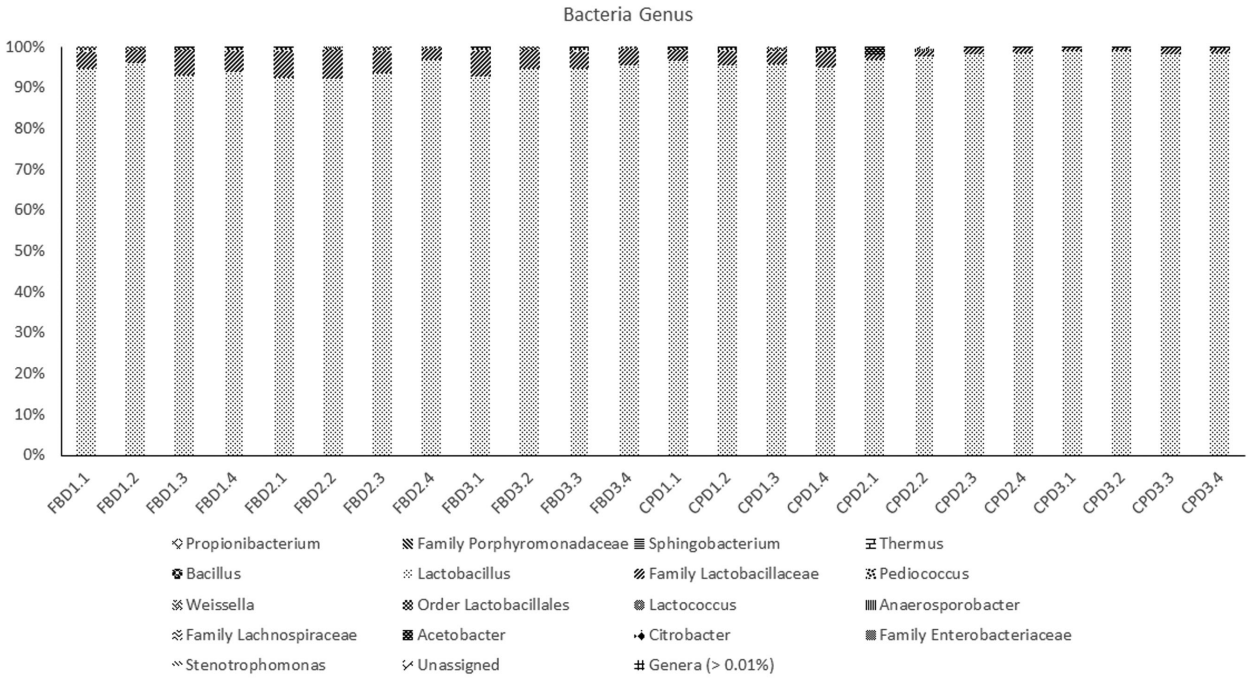

Figure 4. Taxonomic affiliation of 16S rDNA gene sequences in genera against the Ribosomal Database Project database from fermentation tanks in the fed-batch distillery (FBD) and continuous process distillery (CPD) using QIIME.

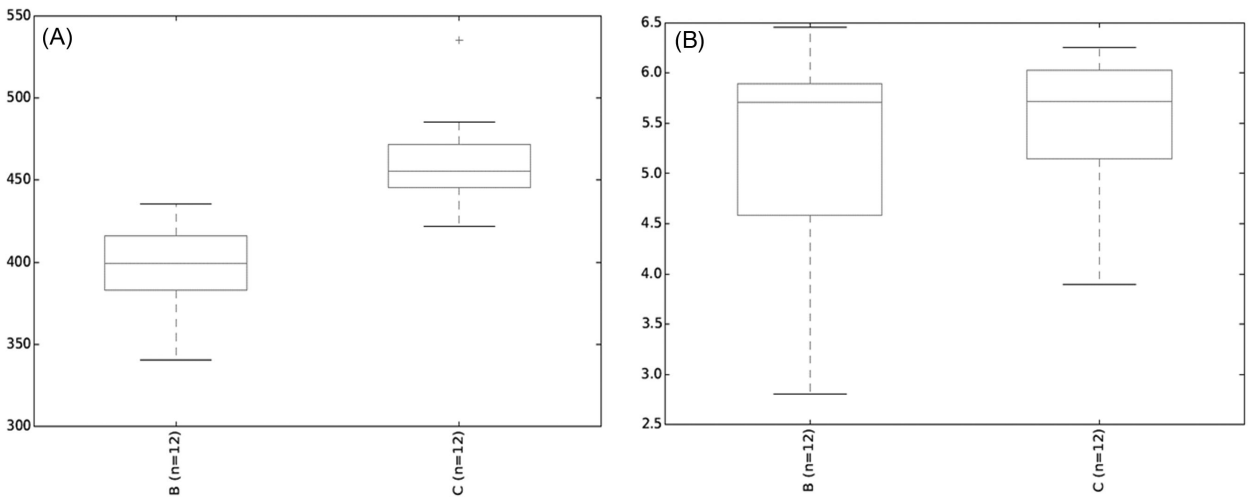

Figure 5. Alpha-diversity analysis of richness considering Chao1 index (**A**) and diversity considering Shannon index (**B**) from fermentation tanks in the fed-batch distillery (FBD) and continuous process distillery (CPD) using QIIME.

contaminant. Skinner and Leathers (2004) showed that with bacterial population around 10^6 CFU/mL, or even higher, the process may not be significantly affected. This is an interesting result since continuous system was reported in the literature as more susceptible to bacterial contamination, when compared to fed-batch process (Godoy *et al.* 2008). Our work suggests that, when well managed, both the processes can operate with low contaminating bacterial abundance.

The most abundant genera reported in all works that studied sugarcane fermentation tank is *Lactobacillus* and, particularly in our work, 92%–99% of the sequences affiliated to this genus in both distilleries. It was considerably more abundant when compared to other studies that investigated similar conditions, they reported 45%–60% of this genus (Rosales 1989; Gallo 1990). In related studies, Costa *et al.* (2015) reported 62.2% of *Lactobacillus* when accessing wine diversity of a sugarcane distillery. Corn dry grind facilities can also be affected by bacterial contaminant, in a percentage higher than reported in sugarcane distilleries. Skinner and Leathers (2004) reported 69%–87% of *Lactobacillus* sp. in a corn batch dry grind facility, while Rich *et al.* (2015) reported 92% of isolates belonging to *Lactobacillus* sp. in the same kind of facility. Our work suggests that the importance of this genus for sugarcane ethanolic fermentation may be higher than the one previously reported.

Despite the abundance similarity, differences between CPD and FBD were reported in OTU-based alfa- and beta-diversity analysis. CPD showed higher richness values when compared with FBD, and alpha-diversity showed correlation between

Table 2. Sequence analysis from Illumina 16S rDNA sequencing through QIIME[a] and PICRUSt[b], showing richness (Chao1) and diversity (Shannon) indexes, number of OTUs, coverage (%) of the sequencing and Nearest Sequence Taxon Index.

Samples	chao1[a]	Shannon[a]	OTU[a]	Coverage (%)[a]	Nearest Sequence Taxon Index values[b]
FBD1.1	394.31	5.88	348	99.93	0.0367
FBD1.2	381.04	5.50	336	99.91	0.0365
FBD1.3	413.16	4.60	367	99.89	0.0372
FBD1.4	396.69	5.74	361	99.92	0.0360
FBD2.1	437.33	5.67	364	99.91	0.0383
FBD2.2	419.40	5.95	360	99.90	0.0336
FBD2.3	411.87	4.49	320	99.91	0.0328
FBD2.4	333.45	2.80	298	99.92	0.0360
FBD3.1	396.16	5.74	371	99.93	0.0390
FBD3.2	418.00	6.28	369	99.90	0.0333
FBD3.3	429.53	6.44	398	99.91	0.0350
FBD3.4	388.98	3.18	346	99.90	0.0358
CPD1.1	471.62	5.39	406	99.91	0.0412
CPD1.2	472.47	5.97	408	99.91	0.0407
CPD1.3	505.24	6.26	459	99.87	0.0379
CPD1.4	461.00	5.52	408	99.91	0.0418
CPD2.1	447.18	6.02	413	99.91	0.0267
CPD2.2	482.03	6.08	419	99.89	0.0343
CPD2.3	448.04	6.06	403	99.91	0.0286
CPD2.4	450.62	5.91	385	99.90	0.0266
CPD3.1	424.72	3.87	380	99.90	0.0448
CPD3.2	435.03	4.08	378	99.90	0.0442
CPD3.3	474.00	5.28	429	99.90	0.0413
CPD3.4	428.12	4.74	391	99.89	0.0427

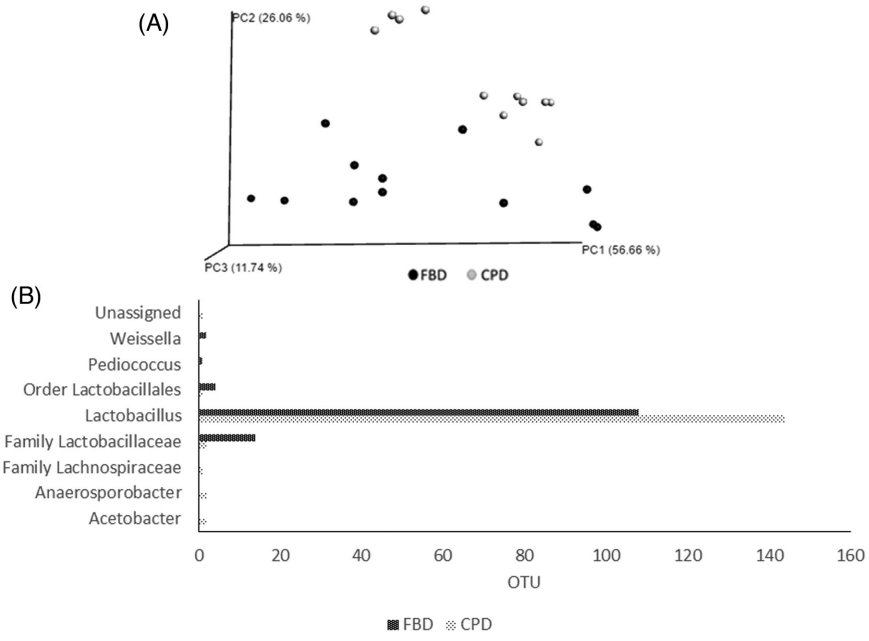

Figure 6. Beta-diversity and differentially represented OTUs analysis using QIIME from fermentation tanks in the fed-batch distillery (FBD) and continuous process distillery (CPD). Principal coordinates analysis tridimensional distance considering weighted-UniFrac (**A**) and differentially represented OTUs genera (**B**).

fermentation tanks in CPD but did not in FBD. This result may be expected because in CPD the fermentation tanks are connected with each other.

Main difference was noticeable through beta-diversity that separated bacterial community from CPD and FBD. Taxonomical analysis reported *Lactobacillus* as the most abundant genus in all distilleries, but OTU considers sequences that share at least 97% of similarity, thus being able to report differences between the distilleries. Several OTUs that belong to genera *Lactobacillus* were differentially represented in the distilleries, wherein CPD presented more of these OTUs. Lucena *et al.* (2010) sampled four distilleries in Northeast of Brazil and reported differences

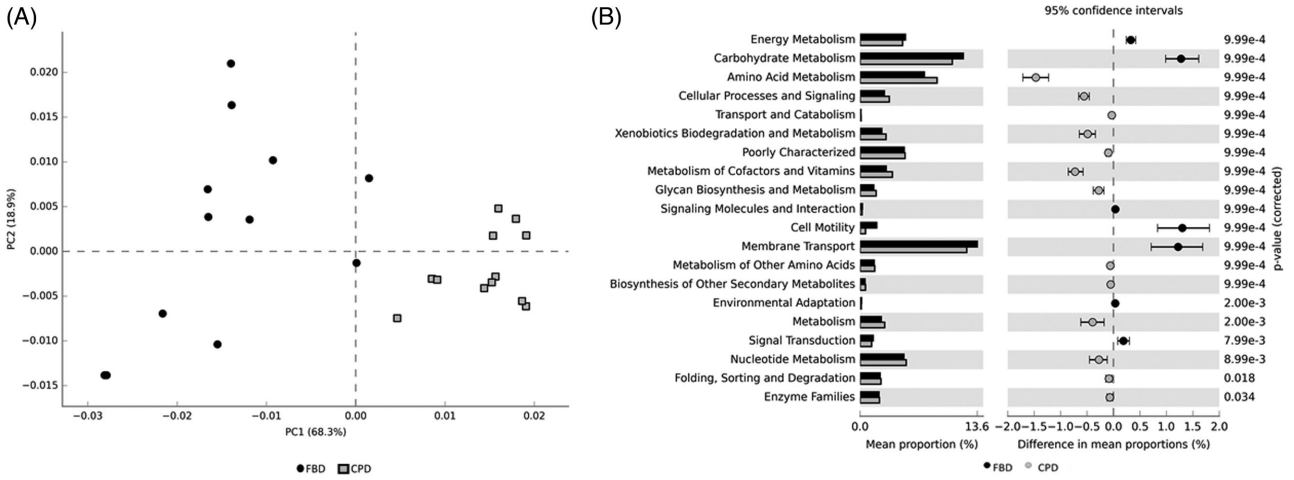

Figure 7. Metagenome prediction analysis using PICRUSt. Principal coordinates analysis bidimensional distance matrix between continuous process distillery (CPD) (gray) and fed-batch distillery (FBD) (black) (**A**) and gene classes that were statistically different between CPD (gray) and FBD (black) (**B**), considering P value < 0.05.

between bacterial communities' composition. *Lactobacillus* can affect yeast fermentation due to the accumulation of both lactic and acetic acids and through competition for nutrient (Narendranath *et al.* 1997; Narendranath 2003) provoking loss of ethanol yield (Basso *et al.* 2014), yeast flocculation (Carvalho-Neto *et al.* 2015) and decrease in yeast viability (Thomas, Hynes and Ingledew 2001).

Through cultured-independent assessment, our work showed other genera that differed significantly between distilleries as *Weissella* and *Pediococcus*, which was more abundant in FBD, and *Acetobacter* and *Anaeosporobacter* were more abundant in CPD. Genera *Weissella*, *Pediococcus* and *Acetobacter* were already reported in this environment (Rosales 1989; Gallo 1990), but *Anaerosporobacter* was not.

Although these bacteria are often reported as contaminant of industrial fermentation involving *Saccharomyces cerevisiae*, little is known about their actual influence on yeast metabolism. *Weissella* and *Pediococcus* are genera of the Lactobacillales family, which as *Lactobacillus* are also known as lactic acid bacteria; in this way, they may affect yeast fermentation in a similar way that *Lactobacillus* does. *Acetobacter* belongs to acetic acid bacteria, and the production of this acid combined with low pH, reduced yeast ethanol production (Graves *et al.* 2006). Finally, although *Anaerosporobacter* was never reported in this environment, this genus belongs to the same order of *Clostridium*, the Clostridiales. Actually, *Anaerosporobacter* is a newly described genus (Jeong *et al.* 2007), and there is little information about its physiology in the literature and for our information, this is the first report of this genus in such environment.

Furthermore, PICRUSt metagenomics prediction also reported the difference found in beta-diversity analysis, that is, gene family analysis separated CPD and FBD, and some of them were closely related to fermentation process in *Lactobacillus*. If there are differences at the metabolic level, the control that is usually made with expensive broad-spectrum antibiotic with high antimicrobial activity—that represents high costs for the fermentation process, and also could lead to bacterial resistance (Muthaiyan, Limayem and Ricke 2010)—should be changed to an approach that is more specific and may be cheaper.

Finally, this clusterization within distilleries indicates the persistence of the bacterial community through harvest season. This assumption was first made by Skinner and Leathers (2004) when studying wet mill and dry grind corn facilities through a period of 1 year. They reported that individual production facilities appeared to have distinct bacterial community and that this was due to persistent endemic infections. The persistence of bacteria through time may be related with their capacity of formatting biofilm. Rich *et al.* (2015) isolated bacteria from a corn dry grind facility and found that 7% of 768 isolates showed capacity of produce biofilm. This indicates that distilleries should pay more attention to equipment cleaning and decontamination.

Despite of the fermentation process per se, it is known that bacterial community can be influenced by other factors. Sugarcane can have different microbiota depending on the place where it was cultivated (Mendes *et al.* 2007; Magnani *et al.* 2010), climate is also another factor that was related with the amendment of microflora of plants and soil (Bossio *et al.* 1998; Chazarenc, Brisson and Merlin 2010). Besides that, equipment maintenance and cleaning can be crucial to the appearance or reinfestation of the contaminants (Amorim *et al.* 2011). This being said, others works should better explore how each of these variables may affect bacterial contamination.

When accessing contaminating bacterial community in sugarcane fermentation process through cultured-independent techniques, we could see that *Lactobacillus* genus may be even more important to this environment than ever thought, that each distillery appeared to have a distinct microbiome, considering both OTU and predictive gene families, and that these communities seemed to persist over time. These results suggest that distilleries may invest in methods that control *Lactobacillus* rather than any other bacteria and take care of tanks cleaning to prevent bacterial reinfestation.

FUNDING

Financial support was provided by FAPESP with a postdoc scholarship to Mariana Souza e Silva (Proc. N. 2011/23816-0) and a PhD scholarship to Maria Leticia Bonatelli (Proc. N. 2013/08431-0).

Conflict of interest. None declared.

REFERENCES

Amorim HV, Basso LC, Lopes ML. Sugar cane juice and molasses, beet molasses and sweet sorghum: composition and usage. In: Ingledew WM, Kelsall DR, Austin GD *et al.* (eds). *The Alcohol Text Book*. 5th edn. Nottingham: University Press, 2009, 39–46.

Amorim HV, Lopes ML, de Castro Oliveira JV *et al.* Scientific challenges of bioethanol production in Brazil. *Appl Microbiol Biot* 2011;**91**:1267–75.

Bartram AK, Lynch MD, Stearns JC *et al.* Generation of multimillion-sequence 16S rDNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Appl Environ Microb* 2011;**77**:3846–52.

Basso LC, Basso TO, Rocha SN. Ethanol production in Brazil: the industrial process and its impact on yeast fermentation. In: Bernardes MAS (ed). *Biofuel Production-Recent Developments and Prospects*. 2011 InTech. ISBN: 978-953-307-478-8.

Basso TO, Gomes FS, Lopes ML *et al.* Homo-and heterofermentative lactobacilli differently affect sugarcane-based fuel ethanol fermentation. *A Van Leeuw J Microb* 2014;**105**:169–77.

Bokulich NA, Subramanian S, Faith JJ *et al.* Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 2013;**10**:57–9.

Bossio DA, Scow KM, Gunapala N *et al.* Determinants of soil microbial communities: effects of agricultural management, season, and soil type on phospholipid fatty acid profiles. *Microbial Ecol* 1998;**36**:1–12.

Brethauer S, Wyman CE. Review: continuous hydrolysis and fermentation for cellulosic ethanol production. *Bioresource Technol* 2010;**101**:4862–74.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

Carvalho-Netto OV, Carazzolle MF, Mofatto LS *et al. Saccharomyces cerevisiae* transcriptional reprograming due to bacterial contamination during industrial scale bioethanol production. *Microb Cell Fact* 2015,**14**:1.

Chazarenc F, Brisson J, Merlin G. Seasonal and spatial changes of microorganism communities in constructed wetlands: a community level physiological profiling analysis. *Int J Chem Eng* 2010, DOI:10.1155/2010/490240.

Cole JR, Wang Q, Fish JA *et al.* Ribosomal database project: data and tools for high throughput rDNA analysis. *Nucleic Acids Res* 2014, DOI: 10.1093/nar/gkt1244.

Costa OYA, Souto BM, Tupinambá DD *et al.* Microbial diversity in sugarcane ethanol production in a Brazilian distillery using a culture-independent method. *J Ind Microbiol Biot* 2015;**42**:73–84.

Crago CL, Khanna M, Barton J *et al.* Competitiveness of Brazilian sugarcane ethanol compared to US corn ethanol. *Energy Policy* 2010;**38**:7404–15.

de Souza Dias MO, Maciel Filho R, Mantelatto PE *et al.* Sugarcane processing for ethanol and sugar in Brazil. *Environ Dev* 2015;**15**:35–51.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1.

Gallo CR. Determinação da microbiota bacteriana de mosto e de dornas de fermentação alcoólica. Ph.D. Thesis. University of Campinas 1990.

Godoy A, Amorim HV, Lopes ML *et al.* Continuous and batch fermentation processes: advantages and disadvantages of these processes in the Brazilian ethanol production. *Int Sugar J* 2008;**110**:175–81.

Graves T, Narendranath NV, Dawson K *et al.* Effect of pH and lactic or acetic acid on ethanol productivity by *Saccharomyces cerevisiae* in corn mash. *J Ind Microbiol Biot* 2006;**33**:469–74.

Heuer H, Krsek M, Baker P *et al.* Analysis of actinomycete communities by specific amplification of genes encoding 16S rDNA and gel-electrophoretic separation in denaturing gradients. *Appl Environ Microb* 1997;**63**:3233–41.

Jeong H, Lim YW, Yi H *et al. Anaerosporobacter mobilis* gen. nov., sp. nov., isolated from forest soil. *Int J Syst Evol Micr* 2007;**57**:1784–7.

Langille MG, Zaneveld J, Caporaso JG *et al.* Predictive functional profiling of microbial communities using 16S rDNA marker gene sequences. *Nat Biotechnol* 2013;**31**:814–21.

Lis JT. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Method Enzymol* 1980**65**:347–53.

Lucena BT, dos Santos BM, Moreira JL *et al.* Diversity of lactic acid bacteria of the bioethanol process. *BMC Microbiol* 2010;**10**:298.

Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;**27**:2957–63.

Magnani GS, Didonet CM, Cruz LM *et al.* Diversity of endophytic bacteria in Brazilian sugarcane. *Genet Mol Res* 2010;**9**:250–8.

Mendes R, Pizzirani-Kleiner AA, Araujo WL *et al.* Diversity of cultivated endophytic bacteria from sugarcane: genetic and biochemical characterization of *Burkholderia cepacia* complex isolates. *Appl Environ Microb* 2007;**73**:7259–67.

Muthaiyan A, Limayem A, Ricke SC. Antimicrobial strategies for limiting bacterial contaminants in fuel bioethanol fermentations. *Prog Energ Combust* 2011;**37**:351–70.

Muyzer G, De Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rDNA. *Appl Environ Microbiol* 1993;**59**:695–700.

Narendranath NV. Bacterial contamination and control in ethanol production. In: Jacques KA, Lyons TP, Kelsall DR (eds). *The Alcohol Textbook*. 4th edn. Nottingham: Nottingham University Press, 2003, 287–98.

Narendranath NV, Hynes SH, Thomas KC *et al.* Effects of lactobacilli on yeast-catalyzed ethanol fermentations. *Appl Environ Microbiol* 1997;**63**:4158–63.

Parks DH, Tyson GW, Hugenholtz P *et al.* STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 2014;**30**:3123–4.

Rappé MS, Giovannoni SJ. The uncultured microbial majority. *Annu Rev Microbiol* 2003;**57**:369–94.

Renewable Fuels Association (RFA). *World Fuel Ethanol Production* 2017. http://www.ethanolrfa.org/resources/industry/statistics/#1454099103927-61e598f7-7643 (4 May 2017, date last accessed).

Rich JO, Leathers TD, Bischoff KM *et al.* Biofilm formation and ethanol inhibition by bacterial contaminants of biofuel fermentation. *Bioresource Technol* 2015;**196**:347–54.

Rosales SYR. Contaminantes bacterianos da fermentação etanólica: isolamento em meios diferenciais, identificação e avaliação de desinfetantes. *Ph.D. Thesis*. Universidade Estadual Paulista "Julio de Mesquista" 1989.

Skinner KA, Leathers TD. Bacterial contaminants of fuel ethanol production. *J Ind Microbiol Biot* 2004;**31**:401–8.

Solomon S. Post-harvest deterioration of sugarcane. *Sugar Tech* 2009;**11**:109–23.

Thomas KC, Hynes SH, Ingledew WM. Effect of lactobacilli on yeast growth, viability and batch and semi-continuous alcoholic fermentation of corn mash. *J Appl Microbiol* 2001;**90**:819–28.

Zanin GM, Santana CC, Bon EPS *et al.* Brazilian bioethanol program. *Appl Biochem Biotech* 2000;**84**:1147–61.