University of São Paulo
"Luiz de Queiroz" College of Agriculture

Changes in the bacterial community and overall metabolic profile in a first-generation ethanol distillery under distinct fermentative yield

Vanessa Costa Iurif

Thesis presented to obtain the degree of Doctor in
Science. Area: Bioenergy

Piracicaba
2020

Vanessa Costa Iurif
Biologist

Changes in the bacterial community and overall metabolic profile in a first-generation ethanol
distillery under distinct fermentative yield.
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **CARLOS ALBERTO LABATE**

Thesis presented to obtain the degree of Doctor in
Science. Area: Bioenergy

Piracicaba
2020

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Iurif, Vanessa Costa

Changes in the bacterial community and overall metabolic profile in a first-generation ethanol distillery under distinct fermentative yield / Vanessa Costa Iurif. versão revisada de acordo com a resolução CoPGr 6018 de 2011, Piracicaba, 2020.

99 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz". Universidade Estadual de Campinas. Universidade Estadual Paulista "Julio de Mesquita Filho"

1. Fermentação 2. Contaminantes 3. Metabarcoding 4. Metabolômica I.
Título

ACKNOWLEDGMENT

Firstly I would like to thank my Ph.D. advisor, Professor Carlos Alberto Labate, and my exchange supervisor, Professor Marcus Thomas Pius Gilbert. I am deeply grateful for their trust in my work, for all the academic support and scientific discussions.

My special regards to Conselho Nacional de Desenvolvimento Científico and Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for my Ph.D. scholarship and the financial support of my Doctoral Exchange Program in Denmark, respectively. I also want to thank the financial support received from my parents after the governmental blocking of scholarships (and also for having me back in their house), which directly affected me and almost impaired the proper conclusion of this Ph.D. study.

A special thanks to all current and former members of Max Feffer Laboratory/ESALQ-USP for all the support in the past four years, for creating such a nice and funny working environment and for “spontaneously” helping me during the experimental samplings. A huge thank to Thais and Mônica for sharing their knowledge and lab experiences so patiently; to Fabricio and the 'fermentation team' - Juliana, Maria Letícia, Mariane and Thalita - for all the scientific help with experiments and data analysis; to Alline, Gustavo and Hana for the friendship.

My sincere thanks to all the staff and members of EvoGenomics Section and GeoGenetics Center from Copenhagen University (Denmark), especially to Adam, Anna, Ashot, Camila, Christina, Dora, Emily, Fatima, George, Jonas, Marcela, Marie, Marta, Physilia, Sarah, Sama, Sarai and Sofi for all the scientific discussions and fun in nordic lands. A big thanks to Christian, Christina, Christopher, Jacob, Kimmo, Luisa, Sarah, and Shyam for also sharing their knowledge and helped me with metabarcoding technics and data processing.

A deep thanks to the current and former staff from the Ph.D. Program in Bioenergy, members of the Department of Genetics/ESALQ-USP, and from ESALQ-USP Graduate Studies Office team. A special thanks to Prof. Maria Carolina Quecine for all the support and encouragement during my teaching training. I also want to thank the assistance and patience of all collaborators from the ethanol distillery where samples were collected.

A huge thanks to my graduate colleagues for all the support, care, shared experiences (good and bad), scientific help, and fun. Specially Albânia, Carlos, Diego, Flaviane, Luis, Mariane, and Renata from ESALQ and from the Doctoral Exchange Program, Carolina and Leda.

I also thank my friends “outside the academical world” and family for their continuous support and encouragement. A big thanks to my grandmother, godparents, Camila, Cynthia, Fernanda, Marília, Thais, Vivian, and many others.

A very special thanks to my boyfriend, Pedro, who was patiently by my side during the hardest parts of this Ph.D., reminding me of what really matters.

At last but not at least, my biggest gratitude to my parents - Isabel and Denis - and my sister - Deborah - for always supporting my dreams and believe in my potential.

CONTENTS

RESUMO	6
ABSTRACT	7
LIST OF ABBREVIATIONS AND ACRONYMS	8
THESIS STRUCTURE.....	9
1. INTRODUCTION.....	11
1.1 GLOBAL PERSPECTIVES FOR ETHANOL.....	11
1.2 BRAZILIAN FUEL ETHANOL	11
1.3 THE INDUSTRIAL ETHANOL FERMENTATION PROCESS IN BRAZIL	12
1.4 THE CONTAMINANTS OF INDUSTRIAL PROCESS.....	13
1.5 NEW TECHNOLOGIES APPLIED TO ETHANOL FERMENTATION	14
REFERENCES	15
2. AIMS	19
3. ASSESSMENT OF BACTERIAL COMMUNITIES PROFILE IN A FIRST- GENERATION ETHANOL DISTILLERY ON NORMAL AND LOW FERMENTATIVE YIELD	21
ABSTRACT	21
3.1 INTRODUCTION.....	21
3.2 METHODS.....	22
3.2.1 Sampling	22
3.2.2 DNA extraction.....	22
3.2.3 Metabarcoding PCR, library preparation and sequencing	22
3.2.4 Sequence analysis	23
3.2.5 Fermentative yield parameters	23
3.3 RESULTS.....	24
3.3.1 The metabarcoding sequence of microorganisms present in the must and in the fermentation tanks	24
3.3.2 Must communities	24
3.3.3 Tanks communities	27
3.3.4 Overall community profile	29
3.4 DISCUSSION.....	29
REFERENCES	32
SUPPLEMENTARY	35
4. UNTARGETED METABOLOMICS OF FIRST GENERATION ETHANOL: COMPARING THE METABOLIC PROFILE OF NORMAL AND LOW ETHANOL YIELDS.....	53
ABSTRACT	53
4.1 INTRODUCTION.....	53
4.2 METHODS.....	54
4.2.1 Sampling	54
4.2.2 Sample preparation for GC-MS analysis.....	55
4.2.3 GC- MS data analysis	55
4.2.4 Fermentative yield parameters	56
4.3 RESULTS.....	56
4.3.1 Investigation of differences in metabolic profile across the continuous fermentation line	56
4.3.2 Metabolic profile of Low and Normal productivity.....	57
4.4 DISCUSSION.....	63
REFERENCES	65
SUPPLEMENTARY	68
5. UNTARGETED INTRACELLULAR METABOLIC PROFILE OF FERMENTATIVE TANKS FROM A BRAZILIAN ETHANOL DISTILLERY UNDER DISTINCT ETHANOL YIELDS.....	79

ABSTRACT	79
5.1 INTRODUCTION	79
5.2 METHODS	80
5.2.1 Sampling	80
5.2.2 Sample preparation for GC-MS analysis	80
5.2.3 GC- MS data analyses	80
5.2.4 Fermentative Yield parameters.....	81
5.2RESULTS	81
5.3 DISCUSSION	86
REFERENCES.....	88
SUPPLEMENTARY	90
6 AUTHOR’S FINAL CONSIDERATIONS AND FUTURE PERSPECTIVES	97

RESUMO

Alterações na composição da comunidade bacteriana e no perfil metabólico de uma usina de primeira geração de etanol em diferentes condições de produtividade

As crescentes preocupações sobre os efeitos da queima de combustíveis fósseis nas mudanças climáticas intensificaram as buscas por fontes alternativas de energia. Nesse cenário, o bioetanol despontou como uma alternativa viável para a substituição de combustíveis fósseis na matriz energética. No mundo todo, a fermentação é a via principal de produção de etanol. No Brasil, que é o segundo maior produtor de etanol no mundo, a produção do álcool é obtida principalmente por meio da fermentação de derivados da cana-de-açúcar por leveduras. No entanto, a presença de contaminantes bacterianos é comum em todas as etapas da produção. É de amplo conhecimento que a presença de contaminantes pode levar a uma diminuição na produtividade do etanol e levar a indústria a grandes perdas financeiras. Para avaliar se existem diferenças na comunidade bacteriana em situações distintas de produtividade, amostras de mosto e de dornas de fermentação foram coletadas de uma destilaria industrial de produção de etanol na região de Piracicaba (SP-Brasil), operando em sistema contínuo de alimentação do mosto. As comunidades bacterianas de ambas as situações foram identificadas por meio da metodologia de metabarcoding que sequencia regiões específicas do gene 16S rDNA. Com base nessa metodologia foi possível observar diferenças na microbiota dos grupos de produtividade baixa e normal de etanol. No mosto observamos diferentes proporções de gêneros bacterianos em situações distintas de produtividade. Já nas dornas de fermentação, como esperado, observamos a predominância de bactérias pertencentes ao gênero *Lactobacillus*. Entretanto, observamos algumas diferenças nos padrões de OTUs atribuídas a *Lactobacillus* nos grupos de produtividade normal e baixa. Considerando que nossos experimentos demonstraram que há diferenças na comunidade bacteriana da destilaria em diferentes condições de produtividade, nosso segundo propósito foi examinar o perfil metabólico, com o objetivo de encontrar possíveis metabólitos-chave para o monitoramento de processos. Utilizando a cromatografia gasosa acoplada a espectrometria de massas não direcionada, o exometaboloma e o perfil metabólico intracelular da mesma destilaria foi explorado em condições de produtividade baixa e normal. Com essa metodologia foi possível selecionar metabólitos no meio extracelular com o potencial de serem explorados como futuros marcadores de monitoramento da produtividade, como os metabólitos *Oleic acid*, *Tagatose*, *3-O-methyl-Glucose*, *5 beta-Dihydrotestosterone*, *NA 1045* e *NA901*. Também foi possível identificar metabólitos provavelmente relacionados a alterações em vias metabólicas. Notamos uma maior abundância de metabólitos relacionados a vias do metabolismo energético em produtividade Normal, como: *Glucose*, *Glucose-6-phosphate*, *Lactulose*, *Maltotriose* e *Sorbose*. Ao analisarmos o perfil intracelular também detectamos metabólitos relacionados com alterações na comunidade microbiológica em condições distintas de produtividade, como o metabólito mannitol, por exemplo.

Palavras-chave: Fermentação, Contaminantes, Metabarcoding, Metabolômica

ABSTRACT

Changes in the bacterial community and overall metabolic profile in a first-generation ethanol industry under distinct fermentative yield

Growing concerns about the greenhouse gases emissions, produced by the fossil fuel combustion, on climate change have intensified the search for alternative sources of energy. In this scenario, bioethanol was recognized as a viable alternative for the replacement of certain fossil fuels. The main route for ethanol production is the ethanolic fermentation of sugary feedstocks. Through the fermentation of sugarcane derivatives, Brazil is the second-largest ethanol producer in the world. The presence of bacterial contaminants is common in all stages of Brazilian ethanol production. It is well known that the presence of contaminants can lead to a decrease in ethanol productivity and cause financial losses for the ethanol industry. In order to evaluate if there are differences in the bacterial community of distinct ethanol productivity yield, samples from must and fermentation tanks were collected from an industrial ethanol production distillery in the Piracicaba region (SP-Brazil) running in a continuous system of fermentation. Bacterial communities were identified through metabarcoding sequencing of regions of the 16S rDNA gene. We were able to observe differences in the microbiota from samples representing low and normal ethanol productivity. In the must samples we observed different proportions of bacterial genera. In the fermentation tanks, as expected, we observed a predominance of *Lactobacillus*. However, despite microbes being from this genus, we observe some differences in OTU patterns in samples from Low and Normal productivity. Since our data showed that there are differences in the distillery community on different productivity conditions, our second goal was to examine the metabolic profile aiming to find potential key metabolites for process monitoring. Using untargeted GC-MS the extracellular and intracellular metabolic profile of the same distillery were explored in low and normal productivity conditions. This approach proved to be effective to select metabolites from the extracellular medium to be further explored for the distillery as potential monitoring markers of productivity, such as the metabolites Oleic acid, Tagatose, 3-O-methyl-Glucose, 5 beta-Dihydrotestosterone, NA 1045 and NA901. Through intracellular analysis it was possible to identify metabolites related to changes in metabolic pathways. We noticed a greater abundance of energy-related metabolism pathways in Normal productivity, such as: Glucose, Glucose-6-phosphate, Lactulose, Maltotriose and Sorbose. It was also possible to detect metabolites related to changes in the microbiological community under different conditions of productivity, such as the metabolite mannitol, for example to select metabolites to be further explored for the distillery as potential monitoring markers of productivity. It was also possible to identify, through GC-MS approaches, some metabolites probably related to changes of the microbial community under distinct fermentative yield situations.

Keywords: Fermentation, Contaminants, Metabarcoding, Metabolomics

LIST OF ABBREVIATIONS AND ACRONYMS

°C	Degree Celsius
bp	Base pair
DNA	Deoxyribonucleic acid
DNTP	Deoxyribonucleotide triphosphate
FC	Fold change
FDR	False discovery rate
GC	Gas chromatography
mg	Milligram
ml	Milliliter
OTU	Operational Taxonomic Unit
PCA	Principal component analysis
PCR	Polymerase chain reaction
PCoA	Principal coordinates analysis
PERMANOVA	Permutational multivariate analysis of variance
qPCR	Quantitative PCR
rDNA	Ribosomal DNA
MS	Mass spectrometry
μl	Microliter
μM	Micromolar

THESIS STRUCTURE

This thesis presents a brief introduction followed by three chapters and the author's final considerations. Chapters are presented here in the format of drafted manuscripts, since they are in preparation for submission to peer-review journals.

In the chapter entitled “ASSESSMENT OF BACTERIAL COMMUNITIES PROFILE IN A FIRST-GENERATION ETHANOL DISTILLERY ON NORMAL AND LOW FERMENTATIVE YIELD”, by sequencing the 16S region of ribosomal DNA, samples from must, tank 1 and 3 of a distillery operating in continuous system were evaluated. Samples were collected throughout the Brazilian sugarcane harvest 2017/2018. To our knowledge, this is the first study relating bacterial community and fermentative yield in Brazilian distilleries. We expect that this work contributes to a better understanding of bacterial dynamics, helps in decision-making, and assists future researches and industrial practices aiming to improve ethanol production.

The chapter “UNTARGETED METABOLOMICS OF FIRST GENERATION ETHANOL: COMPARING THE METABOLIC PROFILE OF NORMAL AND LOW ETHANOL YIELDS” displays an exploratory analysis of the metabolic profile from the ethanol distillery. Samples from the must, acid treatment, tanks 1, 2, and 3 were collected under different conditions of productivity, between August and November 2017. We hope that this exploratory analysis be a starting point for future studies aiming to develop technologies to monitor the fermentation process

In the chapter “UNTARGETED INTRACELLULAR METABOLIC PROFILE OF FERMENTATIVE TANKS FROM A BRAZILIAN ETHANOL DISTILLERY UNDER DISTINCT ETHANOL YIELDS” the same sample set of the previous chapter were used but for assessing the intracellular metabolites profile.

This thesis is concluded with the author's final considerations and future perspectives.

1. INTRODUCTION

1.1 Global Perspectives for Ethanol

The increase in greenhouse gas emissions is addressed as the major cause responsible for global warming and environmental changes (IPCC, 2014). In 2018, the Intergovernmental Panel on Climate Change – IPCC – issued a special report (IPCC, 2018) on the impacts of global warming and emphasized the need for changes in agriculture, energy, transport, cities, and social structure. In this context, the adoption of renewable energy matrices intended for natural resources conservation and mitigation of carbon dioxide emission is critical.

The use of biomass as an energy source has been widely discussed and ethanol seems a good alternative to replace fossil fuels. Nowadays the US and Brazil lead global ethanol production (RFA, 2019). Brazilian ethanol is mainly produced using sugarcane as a feedstock, while in US corn is the principal raw material. Moreover, the industrial process in both countries is considerably different and the acquired knowledge in one of these processes cannot be directly applied to the other (AMORIM et al., 2011; LOPES et al., 2016).

The use of ethanol as fuel in Brazil is recognized worldwide as a successful renewable energy program that encouraged its use in many countries to reduce pollutant emissions (LOPES et al., 2016). The IPCC report (2018) highlighted the use of bioethanol in Brazil as a fuel and its benefits, such as reducing carbon dioxide and ultra-fine polluting particles emission. Environmental concerns regarding deforestation by croplands, impacts on biodiversity, water resources and food security were also expressed. Facing all these aspects can be a challenge for policymakers, businesses, and scientists, but new technologies and innovation should aid in improving ethanol production and minimizing these risks.

Brazilian ethanol production has always been related to national energy policy and economics aspects. Thus, to make projections about the future of Brazilian ethanol, it is important to take these aspects into account.

1.2 Brazilian Fuel Ethanol

The use of ethanol in the Brazilian fuel sector has begun with the decree 19.717 from 1931, which made it compulsory to add locally produced ethanol to imported gasoline. However, the ethanol industry was strengthened after the first oil crisis, with the creation of the National Alcohol Program - Proálcool. Originally the program was launched to stimulate the production of anhydrous ethanol to be mixed with gasoline but after the second oil crisis, in 1979, the production of hydrated alcohol and the development of vehicles powered by 100% ethanol had started (AMORIM et al., 2011; CRUZ et al., 2016). In the first phase of Proálcool, ethanol production took place in distilleries attached to the sugar refineries. With the increasing demand for ethanol, during the second phase of the program distilleries, dedicated exclusively to ethanol production, were established (CRUZ et al., 2016).

The program led to the successful technological development of sectors related to ethanol production, such as sugarcane breeding, modernization of the industrial process, and improvement of agricultural management (MATSUOKA; FERRO; ARRUDA, 2009). With incentive policies, government subsidies and technological advances, in the early 1980s, most of the Brazilian fleet of cars and small trucks were running with ethanol as fuel. Nonetheless, after the oil prices had dropped, increasing sugar prices in international markets, political and economic factors culminated in the shortage of fuel that led to a huge decline in ethanol-powered cars (LOPES et al., 2016). There was a stagnation of Brazilian ethanol production until the early 2000s with the development and commercialization of cars with flex-fuel technology. Those cars were designed to run on pure ethanol, gasoline, or a mixture of both fuels in any

proportion, providing drivers the flexibility to choose the fuel in the gas station (MATSUOKA; FERRO; ARRUDA, 2009; CRUZ et al., 2016). This new scenario restarted the interest in ethanol consumption and production which stimulated investments in the industrial sector resulting in a significant increase in Brazilian ethanol production (MATSUOKA; FERRO; ARRUDA, 2009).

A new biofuel policy was launched by the Brazilian government in December 2017, aiming to promote and develop technologies to produce ethanol from sugarcane bagasse and other cellulosic feedstocks, known as second-generation ethanol (E2G) (CHANDEL et al., 2019). The production of energy through cellulosic biomass would reduce the need for more agricultural lands for crop expansion, decreasing environmental impacts, and considerably increasing the supply of ethanol (BUCKERIDGE et al., 2012). Although important progress has been made, the E2G production is not economically competitive yet (CHANDEL et al., 2019). In addition to the development of new technologies, an improvement in ethanol productivity can be achieved through the optimization of current processes (DOS SANTOS E SILVA; BOMTEMPO; ALVES, 2019). Thereby, this work will only address ethanol production derived from sucrose, known as first-generation ethanol.

The Brazilian National Association of Automotive Vehicle Manufacturers (ANFAVEA) reported in 2019 that over 87% of Brazilian licensed fleet was composed of flex-fuel cars. It is important to highlight that in Brazil no car runs exclusively on gasoline. Under Brazilian law it is mandatory to add 27% of anhydrous ethanol to commercialized gasoline - PORTARIA MAPA No. 75, of 05-03-2015. According to the Union of Sugarcane Industry (UNICA) database, since 2015 there has been an increase in hydrous ethanol and a decrease in anhydrous ethanol consumption in Brazil, which suggests an increase in ethanol use by drivers.

During the 2018/2019 sugarcane harvest, there was an increase in the production of Brazilian hydrated ethanol, influenced by the decrease in sugar prices in the international market and increased ethanol fuel consumption in Brazil (CONAB, 2019b). This scenario led the industry to prioritize the fermentation of sugarcane in detriment of sugar manufacturing. Nevertheless, it was necessary to import ethanol to supply domestic consumption, and according to UNICA, most of this imported ethanol was produced in the USA.

1.3 The Industrial Ethanol Fermentation Process in Brazil

The main route for ethanol production worldwide is the sugar fermentation by yeast, in particular, *Saccharomyces cerevisiae* (LOPES et al., 2016). In Brazil, ethanolic fermentation in fuel production is mainly derived from sugarcane. However, it is important to underline the growing production of ethanol from corn substrates in the country. For the first time, this production was computed in an ethanol report from the Brazilian National Supply Company (CONAB, 2019a). The lower production cost, the development of industries capable of handling both feedstocks (sugarcane and corn) and market prospects have contributed to this. However, in this study, it will be only discussed the bioethanol production obtained from sugarcane.

Traditionally first-generation ethanol distilleries are coupled with sugar industries in Brazil. After cane reception, the material pass through a quick cleaning step and the sugarcane juice is extracted. The richest fractions of this juice are destined to sucrose manufacturing while the poor fraction, molasse (a residue of sugar production) or a mixture of both are destined to the ethanol process and used as the fermentation must (OLIVEIRA et al., 2015; BREXÓ; SANT'ANA, 2017).

Brazilian industries generally conduct fermentation in reactors - with a capacity of 0.5 to 3 million liters - operating in fed-batch or in a continuous mode. Fed-batch is characterized by the simultaneous tank filling with yeast and must. The continuous fermentation process is characterized by the continuous entring of must at a certain constant

flow and the withdrawal of fermented wine. This keeps the total volume constant during the fermentation (SCHMIDELL; FACCIOTTI, 2001). The continuous process has a lower cost of implementation and the potential to produce more ethanol than the batch process. However, due to the ethanol industry conditions, distilleries operating with continuous fermentation lines have more problems related to contamination and more difficulties to evaluate the fermentation yield than those operating in batch or fed-batch mode (GODOY et al., 2008).

Regardless of the fermentation system adopted by the distillery, the Brazilian fermentation is based on the process developed by Firmino Boinot in the 1930s. According to Amorim et al. (2011), this process has been improved and nowadays fermentation yields are around 91-93%. The main peculiarity of this process is the cell recycling step to recover the yeast biomass. At the end of the fermentation, the yeast cells are separated from the fermented wine by centrifugation and subjected to a decontamination process with organic acids. (BASSO; BASSO; ROCHA, 2011; BREXÓ; SANT'ANA, 2018). Using this technique, up to 95% of yeast cells can be recovered and reused in another fermentation cycle. Hence less sugar is deviated for biomass formation and ensures that a high concentration of yeast cells are inside the tanks, contributing to a quicker fermentation process (BASSO; BASSO; ROCHA, 2011; OLIVEIRA et al., 2015).

However, besides high ethanol concentration, temperature, biotic and abiotic stress, the yeast cells need to deal with the extra stress caused by the recycling and acid treatment step. To maintain high ethanol yield under these conditions, the selection of resistant *Saccharomyces cerevisiae* strains developed by Basso et al. (2008) was very important for ensuring the success of Brazilian ethanol industries. In this study, two naturally occurring strains of *S. cerevisiae* stood out: CAT-1 and PE-2. Both presented the desired technological properties, good fermentation performance, high dominance, and high persistence during the fermentation season. To date, these strains are still widely used as initial yeast inoculum in Brazilian industries due to their traits and ability to persist throughout the harvest (DELLABIANCA et al., 2013; CARVALHO-NETTO et al., 2015).

1.4 The Contaminants of Industrial Process

Considering the magnitude of ethanol distilleries, industrial fermentation occurs in a non-aseptic environment. Therefore, selected yeast strains should deal with the presence of yeast and bacterial contaminants. As pointed out by Brexó and Sant'ana (2017), microbial contaminants are responsible for reducing the fermentative performance due to the increase of stress conditions and the consequent decrease of yeast cell viability.

Over the last decades, many studies of bacterial contaminants in ethanol industries have been conducted in Brazil. The majority of them pointed out lactic acid bacteria (LAB) as the main contaminant in distilleries, particularly those belonging to the genus *Lactobacillus* (GALLO, 1990; LUCENA et al., 2010; BASSO et al., 2014; COSTA et al., 2015; BONATELLI et al., 2017; QUEIROZ et al., 2019). LAB are more resistant to the fermentative process due to their ability to resist low pH, high ethanol concentration, elevated temperature, and higher osmotic pressure (BASSO; BASSO; ROCHA, 2011). Besides *Lactobacillus*, other genera have already been reported in distilleries such as *Acetobacter*, *Acinetobacter*, *Anaeosporobacter*, *Anoxybacillus*, *Bacillus*, *Enterobacter*, *Enterococcus*, *Geobacillus*, *Leuconostoc*, *Pedococcus*, *Streptococcus*, *Weissella*, *Zymomonas* among others (COSTA et al., 2015; BONATELLI et al., 2017; QUEIROZ et al., 2019). To control the bacterial populations, ethanol distilleries use antibiotics, hop products, and chemical bacterial biocides. Even so, bacterial contamination may persist bringing damage to yeast cells (AMORIM et al., 2011).

Contaminant yeasts can be divided into two groups: *non-Saccharomyces* and *Saccharomyces*. Belonging to the first one, we can highlight *Dekkera bruxellensis* as the main contaminant, but yeasts belonging to the genus *Candida*, *Pichia*, *Meyerozyma*, *Hansenula*, *Kloeckera*, *Cryptococcus*, *Rhodotorula*, *Torulopsis*, *Trichosporon* and *Schizosaccharomyces* have

already been reported (LIBERAL et al., 2007; COSTA et al., 2015; LOPES et al., 2016; BREXÓ; SANT'ANA, 2017). Wild *Saccharomyces* strains frequently present undesirable characteristics such as flocculation properties, pseudohyphae development, and produce foam in excess (LOPES et al., 2016). Contamination by wild *Saccharomyces* species is more difficult to control since their metabolism is similar to industrial yeast strains (AMORIM et al., 2011).

Yeast flocculation is one of the major concerns of Brazilian ethanol distilleries. This phenomenon decreases ethanol yield, slow down fermentation kinetics, and impairs the centrifugation step for cell recovery (AMORIM et al., 2011; CARVALHO-NETTO et al., 2015). Flocculation is a reversible process where yeast cells come together forming flakes. It is a calcium-dependent process, mediated by specific cell wall proteins and sugars residues (LI et al., 2013). Although the selected yeast strains cannot flocculate, this phenomenon is frequently observed. This is often caused by co-aggregation between yeast cells and some bacterial contaminants, leading to a flocculation-like phenotype. Some *Lactobacillus* strains, such as *L. fermentum*, *L. vini* and *L. plantarum*, have been observed in co-aggregation with *S. cerevisiae* cells during fermentation (CARVALHO-NETTO et al., 2015).

Besides impairing the process and lessening productivity, microbial contamination also brings financial issues due to the need of increasing the consumption of antibiotics/biocides, acids and antifoam agents (LOPES et al., 2016).

1.5 New Technologies Applied to Ethanol Fermentation

DNA metabarcoding is a term used to define the technique of multispecies identification through the amplification of specific sequences, which just became popular through the emergence of next-generation sequencing (NGS) platforms (TABERLET et al., 2012). With this technique, the characterization of microbial communities using fragments of the 16S region of ribosomal DNA has become highly popular. Considered as a simple approach with a few experimental steps, cheap and with relative fast processing, several complex microbiomes were investigated with metabarcoding (DONG et al., 2017).

Recently this methodology was applied in some studies of microbial contaminants of Brazilian ethanol industries. The biggest advantage over previous studies is the possibility of identifying microorganisms that are difficult to cultivate. Although *Lactobacillus* was the dominant group in the fermentation tanks, it was possible to verify the change in the microbiota during the stages of the process and differences among distilleries (COSTA et al., 2015; BONATELLI et al., 2017; QUEIROZ et al., 2019). Investigations of these nature are useful for a better understanding of the microorganisms' dynamics during the fermentation process and can help the ethanol industry to make decisions.

Besides, to know about the microbiota and its impact on the fermentation process, the industry needs to have the ability to detect signs of potential alterations in yield of the fermentation process that might be occurring. Concerning the detection of biological problems, metabolomics had emerged as the “omics” tool with the greatest potential to solve them. Considering that any change in cell physiology will be amplified through the hierarchy - genome, transcriptome, proteome and metabolome – the detection of metabolites can be more sensitive and accurate than other molecules. (KELL et al., 2005; ALONSO; MARSAL; JULIÃO, 2015). The mass spectrometry (MS) associated with a separation method is the main approach for metabolites identification (REGIANI et al., 2013).

The combination of MS with gas chromatography (GC-MS) is recognized for its higher sensitivity and reproducibility for low molecular weight compounds. In this analysis, the electron impact (EI) is the ionization source, allowing the establishment of a common spectral library worldwide. The existence of a spectrum library for EI fragmentation is an advantage over other MS-based technologies since it allows direct identification of compounds (ZHANG et al., 2020). Thus, GC-MS has been a good method for diagnostics, identification of biomarkers, and also

the elucidation of biological variations (PASIKANTI; HO; CHAN, 2008). GC-MS has been pointed out as a mature chromatographic-MS technology with the advantages of being relatively inexpensive, easy to operate, and with satisfactory stability (ZHANG et al., 2020). Due to its characteristics, this methodology has been applied to biomarkers researches, investigation of biomolecules, evaluation of systemic responses, the identification of metabolic pathways and also for food authentications (JOHNSON; IVANISEVIC; SIUZDAK, 2016; BÖHME et al., 2019; DEMARQUE et al., 2020). In fact, this technique has also been applied in studies of microorganisms' physiology and metabolism during ethanol fermentation (BONATELLI, 2016).

In the next sections, we report our efforts to applied these new approaches to characterize the contaminant bacterial community and explore the overall metabolic profile of an ethanol distillery in normal and low ethanol yield conditions.

References

- ALONSO, A.; MARSAL, S.; JULIÀ, A. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. **Frontiers in Bioengineering and Biotechnology**, v. 3, n. March, p. 1–20, 2015.
- AMORIM, H. V. et al. Scientific challenges of bioethanol production in Brazil. **Applied Microbiology and Biotechnology**, v. 91, n. 5, p. 1267–1275, 2011.
- ANFAVEA. **Associação Nacional dos Fabricantes de Veículos Automotores**. Available in: <<http://www.anfavea.com.br/estatisticas.html>>. Acesso em: 14 jan. 2020.
- BASSO, L. C. et al. Yeast selection for fuel ethanol production in Brazil. **FEMS Yeast Research**, v. 8, n. 7, p. 1155–1163, 2008.
- BASSO, L. C.; BASSO, T. O.; ROCHA, S. N. Ethanol Production in Brazil: The Industrial Process and Its Impact on Yeast Fermentation. In: BERNARDES, M. A. DOS S. (Ed.). **Biofuel Production-Recent Developments and Prospects**. [s.l.] IntechOpen, 2011. 1530.
- BASSO, T. O. et al. Homo- and heterofermentative lactobacilli differently affect sugarcane-based fuel ethanol fermentation. **Antonie van Leeuwenhoek**, v. 105, p. 169–177, 2014.
- BÖHME, K. et al. Recent applications of omics-based technologies to main topics in food authentication. **TrAC - Trends in Analytical Chemistry**, v. 110, p. 221–232, 2019.
- BONATELLI, M. L. **Caracterização da comunidade bacteriana contaminante do processo fermentativo para produção de etanol e o impacto no metaboloma da fermentação**. 2016. Universidade de São Paulo - Escola Superior de Agricultura “Luiz de Queiroz”, 2016.
- BONATELLI, M. L. et al. Characterization of the contaminant bacterial communities in sugarcane first-generation industrial ethanol production. **FEMS Microbiology Letters**, v. 364, n. 17, 2017.
- BREXÓ, R. P.; SANT’ANA, A. de S. Microbial interactions during sugar cane must fermentation for bioethanol production : does quorum sensing play a role ? **Critical Reviews in Biotechnology**, v. 38, n. 2, p. 231–244, 2018.
- BREXÓ, R. P.; SANT’ANA, A. S. Impact and significance of microbial contamination during fermentation for bioethanol production. **Renewable and Sustainable Energy Reviews**, v. 43, p. 423–434, 2017.
- BUCKERIDGE, M. S. et al. Ethanol from sugarcane in Brazil: A “midway” strategy for increasing ethanol production while maximizing environmental benefits. **GCB Bioenergy**, v. 4, n. 2, p. 119–126, 2012.
- CARVALHO-NETTO, O. V et al. Saccharomyces cerevisiae transcriptional reprogramming due to bacterial contamination during industrial scale bioethanol production. **Microbial Cell Factories**, v. 14, n. 1, p. 13, 2015.
- CHANDEL, A. K. et al. Comparative analysis of key technologies for cellulosic ethanol production from Brazilian sugarcane bagasse at a commercial scale. **Biofuels, Bioproducts and Biorefining**, v. 13, n. 4, p. 994–1014, 2019.

- CONAB. **Acompanhamento de safra brasileira - cana-de-açúcar - V.6 - safra 2019/2020 n.1 – Primeiro Levantamento**, 2019a.
- CONAB, C. N. de A. **Acompanhamento da safra brasileira de cana-de-açúcar - V.5 - SAFRA 2018/19 - Quarto levantamento**, 2019b.
- COSTA, O. Y. A. et al. Microbial diversity in sugarcane ethanol production in a Brazilian distillery using a culture-independent method. **Journal of Industrial Microbiology and Biotechnology**, v. 42, n. 1, p. 73–84, 2015.
- CRUZ, C. H. de B. et al. **Universidades e empresas: 40 anos de ciência e tecnologia para o etanol brasileiro**. Editora Edgard Blücher, 2016.
- DELLA-BIANCA, B. E. et al. What do we know about the yeast strains from the Brazilian fuel ethanol industry? **Applied Microbiology and Biotechnology**, v. 97, n. 3, p. 979–991, 2013.
- DEMARQUE, D. P. et al. Mass spectrometry-based metabolomics approach in the isolation of bioactive natural products. **Scientific Reports**, v. 10, n. 1, p. 1051, 2020.
- DONG, X. et al. Fast and simple analysis of MiSeq amplicon sequencing data with MetaAmp. **Frontiers in Microbiology**, v. 8, n. AUG, p. 1–8, 2017.
- DOS SANTOS E SILVA, D. F.; BOMTEMPO, J. V.; ALVES, F. C. Innovation opportunities in the Brazilian sugar-energy sector. **Journal of Cleaner Production**, v. 218, p. 871–879, 2019.
- GALLO, C. R. **Determinação da microbiota bacteriana de mosto e de dornas na fermentação alcoólica**. 1990. University of Campinas, 1990.
- GODOY, A. et al. Continuous and batch fermentation processes: Advantages and disadvantages of these processes in the Brazilian ethanol production. **International Sugar Journal**, v. 110, n. 1311, p. 175–181, 2008.
- IPCC. Climate Change 2014 Synthesis Report Summary Chapter for Policymakers. **Ipcc**, p. 31, 2014.
- IPCC. **Global Warming of 1.5°C**. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change. **Ipcc**, p. 175–311, 2018.
- JOHNSON, C. H.; IVANISEVIC, J.; SIUZDAK, G. Nrm.2016.25. v. 17, n. July, p. 451–459, 2016.
- KELL, D. B. et al. Metabolic footprinting and systems biology: The medium is the message. **Nature Reviews Microbiology**, v. 3, n. 7, p. 557–565, 2005.
- LI, E. et al. Deletion of Intragenic Tandem Repeats in Unit C of FLO1 of *Saccharomyces cerevisiae* Increases the Conformational Stability of Flocculin under Acidic and Alkaline Conditions. **Plos one**, 2013.
- LIBERAL, A. T. de S. et al. Identification of *Dekkera bruxellensis* as a major contaminant yeast in continuous fuel ethanol fermentation. **Journal of Applied Microbiology**, v. 102, n. 2, 2007.
- LOPES, M. L. et al. Ethanol production in Brazil: a bridge between science and industry. **Brazilian Journal of Microbiology**, v. 47, p. 64–76, 2016.
- LUCENA, B. T. L. et al. Diversity of lactic acid bacteria of the bioethanol process. **BMC Microbiology**, v. 10, n. 1, p. 298, 2010.
- MATSUOKA, S.; FERRO, J.; ARRUDA, P. The Brazilian experience of sugarcane ethanol industry. **In Vitro Cellular and Developmental Biology - Plant**, v. 45, n. 3, p. 372–381, 2009.
- OLIVEIRA, M. et al. Sugarcane processing for ethanol and sugar in Brazil. **Environmental Development**, v. 15, p. 35–51, 2015.
- PASIKANTI, K. K.; HO, P. C.; CHAN, E. C. Y. Gas chromatography / mass spectrometry in metabolic profiling of biological fluids. **Journal of Chromatography B**, v. 871, p. 202–211, 2008.
- QUEIROZ, L. L. et al. Dynamics of microbial contaminants is driven by selection during ethanol production. **Brazilian Journal of Microbiology**, 2019.
- REGIANI, T. et al. Eletroforese, Cromatografia e Espectrometria de Massas. In: BORÉM, A.; FRITSCHÉ-NETO, R. (Ed.). **ÔMICAS 360: Aplicações e Estratégias para o Melhoramento de Plantas**.

- RFA, R. F. A. **2019 Ethanol Industry Outlook**. Available in: <<https://ethanolrfa.org/wp-content/uploads/2019/02/RFA2019Outlook.pdf>>.
- SCHMIDELL, W.; FACCIO*TTI, M. C. R. Biorreatores e o Processo fermentativo. In: BORZANI, W. et al. (Ed.). **Biotechnologia Industrial Vol. 2 - Engenharia Bioquímica**.
- TABERLET, P. et al. Towards next-generation biodiversity assessment using DNA metabarcoding. **Molecular Ecology**, v. 21, n. 8, p. 2045–2050, 2012.
- UNICA. **Unicadata - União da Indústria de Cana-de-Açúcar**. Available in: <unica.com.br/unicadata>. Accessed in: Jan. 5 2020.
- ZHANG, X. et al. Mass spectrometry-based metabolomics in health and medical science : a systematic review. **The Royal Society of Chemistry**, v. 10, p. 3092–3104, 2020.

2. AIMS

To expand the knowledge about the impact of contaminants in ethanol production, this work aimed to investigate the bacterial community profile and explored the overall metabolomic profile of a first-generation ethanol distillery upon normal and low fermentative yield. The following approaches were used in this study:

- Bacterial community characterization through cultivation-independent methodologies (metabarcoding);
- Comparison of the bacterial community structure in normal and low fermentative yield;
- Exploratory analysis of metabolites dispersed in the medium and communities' intracellular metabolic profile, through untargeted gas chromatography coupled with mass spectrometry (GC-MS);
- Search for potential metabolic markers through untargeted gas chromatography coupled with mass spectrometry, associated with normal and low fermentative ethanol yield.

3. ASSESSMENT OF BACTERIAL COMMUNITIES PROFILE IN A FIRST-GENERATION ETHANOL DISTILLERY ON NORMAL AND LOW FERMENTATIVE YIELD

Abstract

The presence of bacterial contaminants is common in all stages of Brazilian ethanol production and can lead to a decrease in ethanol production. Great efforts have been made to identify those contaminants and recent studies have accessed the bacterial community through 16S rDNA gene. However, no published work has attempted to use this methodology to explore the relation between bacterial contaminants and ethanol productivity.

In this work, we aimed to evaluate the relation between the bacterial community and fermentative yield, in the industrial environment. Samples from must and fermentation tanks were collected from an ethanol distillery in the Piracicaba region (SP-Brasil) and identified through metabarcoding of the 16S rDNA gene. In the fermentation tanks, as expected, we observed a predominance of *Lactobacillus*. However, we were able to observe some differences in the must bacterial community, especially related to the *Lactobacillus* and *Geobacillus* genera, from samples representing low and normal fermentative yield.

Keywords: Ethanol fermentation, bacterial contamination, distillery, 16S rDNA.

3.1 Introduction

According to the American Renewable Fuels Association (RFA, 2019), Brazil is the second-largest ethanol producer in the world, only surpassed by the United States of America. According to the Brazilian National Supply Company report (CONAB, 2019), 33.14 billion liters of ethanol were produced in the 2018/2019 sugarcane harvest. Brazilian first-generation ethanol production from sugarcane is a well-consolidated process. The combination of sugarcane as feedstock and yeast cell recycling contributes to a fast process with a high fermentative ethanol yield. One advantage of cell recycling is to ensure a higher number of yeast cells, decreasing the sugar deviation for biomass formation, saving it to ethanol formation (AMORIM et al., 2011; BASSO; BASSO; ROCHA, 2011).

The selection of specific strains of *Saccharomyces cerevisiae* to be used as the inoculum was extremely important to ensure that the yeast cell recycling approach was successful (BASSO et al., 2008). However, the industrial process does not happen in aseptic conditions, from the beginning when the fermentation substrate is being prepared to acid treatment after the recycling step, microbial contaminants are not effectively eliminated. As a result contaminants, such as bacteria and wild yeast, can enter and persist in the system, and may interfere directly or indirectly with the fermentation process (BASSO; OLIVEIRA, 2018).

Due to the industrial relevance of this challenge, many efforts have previously been made to identify these contaminants. Wild yeast - *Saccharomyces* and non-*Saccharomyces* species - and gram-positive bacteria - such as *Lactobacillus* and *Bacillus* genera - have been reported as the main industrial contaminants. They are responsible for economic losses due to the decrease of ethanol yield, reduction of yeast viability, excessive foam formation, and induction of flocculation events in the distillery (AMORIM et al., 2011; BREXÓ; SANT'ANA, 2017). In addition, organic acids production by lactic acid bacteria (LAB) adds one more stressful factor to the yeast fermentation, and this has made them the largest target for bacterial control. For example, the use of broad-spectrum antibiotics to control LABs is

common in distilleries, although this naturally increases the production costs and, the environmental impact, and may also induce bacterial resistance (MUTHAIYAN; LIMAYEM; RICKE, 2011).

The majority of previous bacterial contaminant studies have been conducted through culture-dependent techniques, which often underestimate the true community diversity (STREIT; SCHMITZ, 2004). More recently, studies have begun to apply culture-independent methodologies to assess this bacterial community, and they have confirmed that *Lactobacillus* is the most abundant genus in industrial fermentation tanks across different distilleries and Brazilian regions (COSTA et al., 2015; BONATELLI et al., 2017; QUEIROZ et al., 2019). Furthermore, these authors demonstrated differences in the microbiota throughout the sugarcane pre-process and fermentation steps (COSTA et al., 2015; QUEIROZ et al., 2019). This methodology is based on the taxa identification of multiple species through the sequencing of environmental DNA targeting fragments and it is defined as 'metabarcoding' or 'amplicon-based sequencing' (TABERLET et al., 2012).

Although the use of metabarcoding techniques to study the microbial community of the Brazilian ethanol production process has begun to shed light on some aspects of the bacterial contaminant community, data linking the contaminant communities to ethanol productivity is still missing. Therefore in this study, we used 16S metabarcoding to study the bacterial communities of a Brazilian distillery, between August and November 2017, aiming to explore the differences in the bacterial contaminant profile under distinctive ethanol productivity conditions, in order to improve our understanding of how population dynamics can interfere with ethanol productivity.

3.2 Methods

3.2.1 Sampling

Between August and November 2017, samples were collected from an ethanol fermentation distillery in the Piracicaba region in São Paulo state (22° 43' 30" S, 47° 38' 51" W), that operates a continuous line process, running with 3 consecutive tanks connected (**Figure S1**). On each collection day, samples of the must and from fermentative tanks 1 and 3 were collected and named according to their location and collection day (**Figure S1** and **Table S1**). It is worth mentioning that up to the middle of September the distillery used two simultaneous fermentation lines. For this reason, samples C.8, C.9, C.10 and C.11 showed variations A and B. For operational reasons, this strategy was abandoned for the distillery.

The must samples were collected by a tap in the pipe between the heater exchange and the first fermentation tank. Tank samples were collected from the bottom tap of fermentation tanks (**Figure S1**). All sample tubes were immediately frozen using dry ice. Samples were subsequently transported to our laboratory at "Luiz de Queiroz" College of Agriculture/ University of São Paulo, and kept at -80°C. Further processing was performed in sterile conditions.

3.2.2 DNA extraction

DNA extraction was performed on 1ml of each sample using the DNeasy PowerSoil Kit (Qiagen) following the manufacturer's protocol. After DNA extraction, the material was lyophilized then transported to the GLOBE Institute, University of Copenhagen, Denmark, where the subsequent metabarcoding protocol was performed.

3.2.3 Metabarcoding PCR, library preparation and sequencing

The V4 region of 16S rDNA gene was PCR amplified with a single PCR step using the primers 515F (PARADA; NEEDHAM; FUHRMAN, 2016) and 806R (AMY APPRILL et al., 2015) fused with 5' 8 nucleotide tags

(Table S2), as proposed by Murray et al. (2015). Each sample was amplified in triplicates with different tag combinations to avoid possible tagged primer biases and cross-contamination. Prior to metabarcoding PCRs, a quantitative PCR was performed to determine the average number of cycles of PCR amplification for the metabarcoding. Each qPCR reaction consisted of 1µl of diluted sample DNA (1:5) and 19µl of a master mix composed of 2µl Gold PCR Buffer (Applied Biosystems), 2.5 mM Magnesium chloride (Applied Biosystems), 0.2 mM DNTP, 0.5 mg/mL of Bovine Serum Albumin (Invitrogen), 0.6 µM of forward primer, 0.6 µM of reverse primer, 1µl SYBR Green/ROX solution, 0.75 units of AmpliTaq Gold DNA polymerase (Applied Biosystems) and AccuGene molecular biology water (Lonza) up to 20µl. Reactions were run in an Agilent qPCR machine model MX3005 with the following conditions: 95°C for 10 minutes, 45 cycles composed of 95°C for 15 seconds, 20 seconds at 55°C and 72°C for 40 seconds.

Each metabarcoding PCR was performed in a 20µl of reaction using 1µl of DNA template and 19µl of PCR master mix. The PCR master mix was the same as that used in qPCR excluding the SYBR Green/ROX solution. Negative controls were also used, where the DNA was replaced by molecular biology grade water. PCR products were analyzed by electrophoresis on 2% agarose gels to control size and amplification level. PCR products with different tag combination were pooled, thus enabling backtracking of the correct PCR product to each replicate.

Libraries were prepared on pooled amplicons according to CARØE et al. (2020). Following the library purification and qPCR based absolute quantification, libraries were sent to Danish National High- Throughput DNA Sequencing Centre for sequencing with 2x250bp chemistry on an Illumina MiSeq platform.

3.2.4 Sequence analysis

AdapterRemoval v2 tool (SCHUBERT; LINDGREEN; ORLANDO, 2016) was used to process the initial sequencing data. Using this tool the sequences were demultiplexed, primers were trimmed and forward and reverse reads merged with default parameters, except for: minimum overlap 100bp, removal of reads shorter than 50bp and 5% mismatch. After sorting amplicons by tag combinations, we used a modified version of DAME (ZEPEDA-MENDOZA et al., 2016) to retain only sequences that were detected in at least two of the three PCR replicates per sample, and had a minimum sequence length of 150bp. Sequences were filtered and clustered with SUMACLUSt (MERCIER et al., 2013) using a 99% similarity threshold. Blastn (ALTSCHUL et al., 1997) was used to compare the OTUs against SILVA_132_SSURef_tax_silva database (QUAST et al., 2013) and an OTU table was created.

Metabarcoding data visualization and statistical analyses were performed in the R Studio interface (version 1.2.13) (TEAM, 2015), using phyloseq (MCMURDIE; HOLMES, 2013), vegan (OKSANEN et al., 2019) and microbiome (LAHTI; SHETTY,) R packages, following proposed pipelines from Lee (2019) and Mak (2018). For Bray-Curtis dissimilarity and Jaccard index analyses, the original counts were subjected to variance stabilization transformation with fittype 'mean' followed by min-max scaling as recommended by Mc Murdie and Holmes (MCMURDIE; HOLMES, 2014). Statistical differences were tested using PERMANOVA and ADONIS tests (ANDERSON, 2017; LEE, 2019). To visualize the must community composition, original sample counts were transformed in relative abundance data.

3.2.5 Fermentative yield parameters

The ethanol fermentation distillery where the sampling was performed, works with a daily ethanol production target of approximately 600 m³ per day. Due to the constant input of feedstock and output of the fermented

wine in the continuous fermentation line, the processing monitoring is temporal. Thus, a punctual fermentative yield estimative is needed to detect changes and guide industry decision-making. The ratio between the amount of ethanol produced and the amount of expected ethanol produced is then utilized.

The calculation of the expected amount of ethanol is expressed by the following formula:

$$Exp\ Ethanol = \frac{\left\{ \left(\frac{must_{TRS}}{100} \right) * 0.6475 * 0.89 * must_{flow_{rate}} + \left(\frac{eth_{yeast}}{100} \right) * yeast_{flow_{rate}} \right\}}{(must_{flow_{rate}} + yeast_{flow_{rate}})}$$

It takes into account the total reducing sugars present in must ($must_{TRS}$); the stoichiometric sugar to ethanol conversion factor 0.6475 - which equates to the production of 0.6475 liters of 100% ethanol for each kg of total reducing sugars; a theoretical efficiency of fermentation at 89%; the must flow rate; the percentage of ethanol detected in yeast cells in the acid treatment and the flow rate of yeast cells in the first fermentative tank.

In normal conditions, the value of this ratio should be close to 1. Measurements for ratio calculations are usually taken three times a day. Considering that small measurement deviation is expected, during this sampling period daily ratio values up to 0.93 were considered as normal fermentative yield by the distillery.

To facilitate understanding, in this study we named "Normal productivity" ratio values up to 0.93 and "Low productivity" below 0.93. It is important to highlight that ratio values were provided by the ethanol distillery where samples were collected (**Table S1**).

3.3 Results

3.3.1 The metabarcoding sequence of microorganisms present in the must and in the fermentation tanks

After the trimming and pre-processing steps, 3,558,856 metabarcoding reads with an average length of 280bp were retained. These reads were clustered into OTUs, and all OTUs containing fewer than 10 reads were removed (OLIVER et al., 2015). Since the objective of this work was to evaluate the bacterial community, the sequences belonging to other Kingdom and those assigned to mitochondrial sequences were also removed for the following analyzes. In total 635 OTUs were retained, of which nearly 82 % were affiliated to Firmicutes, the predominant phylum across all samples (**Table S3**). It was also possible to identify smaller proportions of sequences belonging to the phyla *Actinobacteria*, *Aquificae*, *Bacteroidetes*, *Cyanobacteria*, *Deinococcus-Thermus*, *Epsilonbacteraeota*, and *Proteobacteria* (**Table S3**).

In must samples, nearly 74% of the OTUs were assigned to the Firmicutes, whereas in tank samples this percentage increase to almost 93%. The percentage of OTUs assigned to Firmicutes also varied with indicators of productivity. Firmicutes represented 77% (must) and 92% (tank) of the OTUs from samples collected during Low productivity fermentation, while these percentages changed to approximately 70% (must) and 96% (tanks), in samples collected during normal productivity fermentation (**Table S4, S5**). We subsequently explored the communities of the must and tanks in more detail, in light of finding differences in the resulting productivity.

3.3.2 Must communities

In order to verify differences between bacterial communities of the must in Low and Normal productivity situations, β -diversity analyzes were applied. The communities' structures were investigated with a PCoA considering Bray-Curtis and Jaccard index as distance metrics. A PERMANOVA test was run to partition the variation and check

for statistical evidence with the chosen measure. As **Figure 1** illustrates, samples from the Low and Normal productivity fermentations groups showed a different profile. It was possible to observe the grouping of samples based on Bray-Curtis dissimilarity (PERMANOVA: $p = 0.013$) and presence/absence indicators - Jaccard index (PERMANOVA: $p = 0.009$). This strongly suggests that there are differences in bacterial communities between different productivity conditions in must samples.

Figure 1: β -diversity of must bacterial communities separated by different productivity conditions. (1A) PCoA plot from β -diversity of Bray-Curtis Dissimilarity (PERMANOVA $p = 0.013$). (1B) PCoA plot from β -diversity of Jaccard index (PERMANOVA $p = 0.009$). Red and blue dots represent Low and Normal productivity samples, respectively.

The must community global profiles were visualized through the relative abundance of each genus across Normal and Low productivity groups (**Figure 2A**). As suggested in previous analyses, the communities showed different taxa compositions. In the Normal fermentative group, *Geobacillus* was the genus with the highest relative abundance (56%) followed by *Lactobacillus* (10%). Meanwhile, in Low productivity group, the genus *Lactobacillus* prevailed accounting for 72% of relative abundance. We also compared the individual profile of the must samples (**Figure 2B**). It was possible to observe the predominance of the genus *Lactobacillus* in Low productivity samples and a low abundance of *Geobacillus* in the samples M.8, M.9 and M.16. In the Normal productivity group, samples presented a greater abundance of *Geobacillus* than samples of Low productivity group and, except for sample M.13, presented a low relative abundance of *Lactobacillus*.

Figure 2: Relative abundance of the main genera found in must samples. (2A) The must samples are classified according to their productivity condition. (1B) Each must sample is presented separately. Samples M.11, M.12, M.13, M.14 and M.15 belong to the Normal productivity group while samples M.16, M.21, M.22, M.8 and M.9 belong to to the Low productivity group.

3.3.3 Tanks communities

Since this study was conducted in an ethanol distillery operating in continuous mode, in an attempt to find differences in the microbiota of the tanks that could explain the reduction in ethanol production, the first and the last tank of the fermentation line were chosen for sampling. However, it was not possible to separate tank 1 from tank 3 in terms of microorganisms' communities using the Bray-Curtis dissimilarity, Jaccard index, or PERMANOVA analysis (Figure 3). For this reason, samples from tanks 1 and 3 were analyzed together.

Figure 3: β -diversity of bacterial communities from of fermentative tanks, separated by different productivity conditions. (3A) PCoA plot from β -diversity of Bray-Curtis Dissimilarity. (3B) PCoA plot from β -diversity of Jaccard index. Dots correspond to Tank 1 samples and the triangles represent samples from Tank 3. Red and blue characters represent Low and Normal productivity samples, respectively.

In contrast to the must community results, the tank community profiles could not be clearly distinguished accordingly to productivity groups (Figure 4A). To assist the comprehension of this data and the communities' composition observation, the relative abundance of each genus in tanks samples are represented in Figure 4. The genus *Lactobacillus* was dominant in nearly all samples (Figure 4B). Please note that sample names are composed by the tank number in which the sample was collected, followed by the fermentative line and the sampling code. Thus, sample 1A.10 belongs to tank 1, fermentative line A and sampling code 10, while sample 3A.10 belongs to tank 3, fermentative line A and sampling code 10. Both samples were collected on the same sampling day. Interestingly, it was possible to observe that on the same collection day, samples from tanks 1 and 3 presented a similar genera composition. This supports the lack of differences in β -diversity analysis between tank1 and 3.

Figure 4: Relative abundance of the main genera found in tank samples. (4A) Samples are classified according to their productivity condition (4B) The profile of each tankl sample is represented. Samples between 1A.10 and 1B.9 belong to Tank 1 while samples between 3A.10 and 3B.9 belong to Tank 3.

3.3.4 Overall community profile

Intending to have an overview of the genus and OTU distribution across the fermentation process in the distillery, the 30 most abundant OTUs in this study were selected (**Figure 5**), which represented more than 90% of total reads distribution. The most abundant genera in our samples were *Anoxybacillus*, *Bacillus*, *Geobacillus*, *Lactobacillus*, *Megasphaera*, *Saccharococcus*, *Staphylococcus* and *Streptococcus*. Despite *Lactobacillus* being the most abundant genus in the tank samples (**Figure 5A**), when the OTU profile was explored in more detail, it was possible to observe that some specific OTUs, and the general OTU proportions differed between Normal and Low fermentative productivity groups (**Figure 5B**). It is possible to highlight the OTU4 – assigned to *Lactobacillus* genus – that was well represented in the must and tank samples of the Low fermentative productivity group but miss-represented in the Normal group. Notably, this pattern was also seen in the must samples, with the same OTU present in the must of the Low productivity samples, but absent in the must of the Normal group.

Our results also indicated that in fermentative tanks of both productivity groups the OTUs 1 and 3 had a representative relative abundance, while in the must they accounting for a smaller proportion of the communities. OTU 1 and 3 were also assigned to the *Lactobacillus* genus. It was also apparent that the high relative abundance of *Geobacillus* in the must of the Normal productivity group – as represented by OTUs 5, 6, and 14 - was reduced in the fermentative tanks.

Figure 5: Relative abundance of the 30 most abundant OTUs across all the collection sites. (5A) Relative abundance profile of the assigned genera for must and fermentative tanks (5B) Relative abundance profile of each OTU in must and fermentative tanks

3.4 Discussion

It is common knowledge that Brazilian ethanol production is undertaken in non-aseptic environments and consequently, the existing microbial contaminants can affect the ethanol yield and result in economic losses (BREXÓ; SANT'ANA, 2017). Previous studies have suggested that fermentation feedstock is an important source of contamination and that microbial abundance in the substrate may be related to the sugarcane plant, soil conditions,

weather, and industrial pre-processing steps (LUCENA et al., 2010; QUEIROZ et al., 2019). Upon entering the fermentation process, contaminants may interfere with yeast growth through competition for nutrients, production of toxic substances (eg. organic acids) and/or consuming trace nutrients required for yeast growth (BAYROCK; INGLEDEW, 2004; COSTA et al., 2015). Thus, it is expected that feedstock contaminants directly impact the fermentation yield.

Interestingly, β -diversity analysis showed that the must community structure, belonging to two different productivity conditions, differed. Previous work accessed the Brazilian ethanol must microbial community using 16s metabarcoding (COSTA et al., 2015; QUEIROZ et al., 2019), but they did not correlate these data with ethanol yield parameters. Evaluating the relative abundance data, we reported different predominant genera in each productivity group. In the must Low productivity group, the predominant genus was *Lactobacillus* while in Normal group *Geobacillus* showed a higher relative abundance. Interestingly in sample M.13 we observed an expressive relative abundance of the genera *Lactobacillus* and *Geobacillus*. Even with a greater abundance of *Lactobacillus*, this sample belonged to the normal productivity group. Perhaps the presence of *Geobacillus* has some controlling effect on *Lactobacillus* population.

In Costa et al. (2015) study, *Tatumella* was identified as the predominant genus in must samples. In Queiroz et al. (2019) study, where must samples were collected before and after passing through the heat exchange, *Acinetobacter* and *Zymomonas* were identified as the most abundant genera in must samples collected before and after the must pass by the heat exchange, respectively. No bacteria belonging to the genus *Tatumella* have been reported in our analysis, but *Acinetobacter* and *Zymomonas* were detected in must samples belonging to both productivities group.

Lucena et al. (2010) using ARDRA fingerprinting and pheS and 16S rRNA gene sequencing techniques, reported *Lactobacillus* as the main LAB genus in the fermentation substrates and highlighted how different *Lactobacillus* species in the substrate may affect the selected yeast strain metabolism during fermentation, leading to alterations in ethanol yield. It is worth mentioning that the chemical composition of must varies widely and also plays a role in the fermentation yield (BASSO; BASSO; ROCHA, 2011; SOUZA et al., 2015). This is probably related to alterations to yeast cell metabolism and the favoring growth conditions of certain contaminants. When the must enter the fermentation tank, a shift in the microbial community is expected due to the changes in the environmental conditions.

In fermentation tanks, we found that the bacterial populations presented more homogeneity. *Lactobacillus* was the genus with the highest relative abundance, regardless of productivity. This data corroborated with the previously published literature on contaminants in the Brazilian industrial fermentative tanks (LUCENA et al., 2010; COSTA et al., 2015; BONATELLI et al., 2017; BREXÓ; SANT'ANA, 2017; QUEIROZ et al., 2019), and this observation can be explained by the great ability of *Lactobacillus* to survive in the fermentative tank environment. It is recognized that the low pH, high ethanol content, high osmotic pressure and high temperatures that are typical of the tanks are well-tolerated conditions by this group of bacteria (BASSO; BASSO; ROCHA, 2011a). Interactions between *Lactobacillus* and yeast cells in fermentative tanks may change the yeast metabolism and lead to undesirable phenomena such as fermentation inhibition, increased gum formation and flocculation induction (CARVALHO-NETTO et al., 2015).

In order to have a better overview of group community compositions across sampling points, the 30 most abundant OTUs were selected for evaluation. The majority of genera assigned to those OTUs had already been described in previous studies on the fuel ethanol industry. In Brazilian distilleries, the genera *Lactobacillus*, *Bacillus*, *Anoxybacillus*, *Geobacillus*, and *Streptococcus* had already been identified through studies using 16S approaches (COSTA et al., 2015; BONATELLI et al., 2017; QUEIROZ et al., 2019). The genus *Staphylococcus* had been reported by Skinner-Nemec et al (2007) as a biofilm component in fermentation tanks of a continuous ethanol industry. Although we are

not aware of any work reporting *Saccharococcus* bacteria in the ethanol fermentation industry, it is known that some soils thermophilic strains of *Saccharococcus* species were reclassified as belonging to the genus *Geobacillus* (FORTINA et al., 2001). In our analysis, this genus was only detected in samples belonging to the normal productivity group, following a pattern similar to *Geobacillus*. For this reason, we believe that this OTU may also be related to the genus *Geobacillus*. We had also identified a OTU assigned to the *Megasphaera* genus. This genus is commonly reported in breweries as beer spoilage bacteria, and it is composed of strictly anaerobic gram-negative bacteria (SAKAMOTO; KONINGS, 2003; JUVONEN, 2015).

When evaluating the OTUs distribution, it is important to highlight the difference in OTU4 representativeness in the must and tanks from both groups. This OTU, assigned to the *Lactobacillus* genus, showed a high abundance only in the Low fermentative productivity group. It is possible to observe that OTU4 was present in must samples and remained up to the end of the fermentative process. Although we cannot classify the species of *Lactobacillus* using only fragments of the 16S rRNA gene (NASER et al., 2007), it is tempting to speculate that different OTUs belong to different species that might have distinct roles in the fermentation process. It has previously been reported that different *Lactobacillus* species affect yeast metabolism in various ways (BASSO et al., 2014). According to their sugar metabolism pathway, *Lactobacillus* species are classified into homofermentative or heterofermentative. Both groups are commonly found in industrial ethanol fermentation tanks, but as previously reported, when simulating Brazilian industry fermentation conditions, heterofermentative species are more detrimental to ethanol yield (COSTA et al., 2008; BASSO et al., 2014). Perhaps the prevalence of a certain group in the must affect directly the proportion of homo and heterofermentative *Lactobacillus* populations in the tanks, resulting in a more prominent reduction of fermentation yield.

Other interesting group found in our work was *Geobacillus* and although they were presented in high abundance in must samples, its relative abundance was highly reduced in the tanks. As mentioned earlier, the conditions in the tanks favor *Lactobacillus*, but perhaps the presence of *Geobacillus* in the must could play a role in the composition of the microbial tank community by influencing the profile of contaminants in the must. Originally classified as representatives of *Bacillus* genus, *Geobacillus* genus was established in 2001. It is composed of thermophiles gram-positive bacteria, which are found in a variety of environments and have the ability to ferment a wide range of substrates such as complex polysaccharides of plant biomass (ZEIGLER, 2014; NOVIK; SAVICH; MEEROVSKAYA, 2019). Another important characteristic of this group is the ability to produce bacteriocins and antimicrobial peptides, which are capable of inhibiting some pathogenic bacteria (POKUSAEVA et al., 2009). However, sequence similarity between the *Bacillus* and *Geobacillus* genera implies that some species classifications need to be revised (NOVIK; SAVICH; MEEROVSKAYA, 2019).

Recently Saunders et al (2019) published on the ability of a substance produced by *Bacillus* strains, to inhibit biofilm formation in the ethanol fuel industry by isolates of *Lactobacillus*, due to the persistence of *Lactobacillus* infections been associated with its ability to form a biofilm. In fact, Dellias et al. (2018) isolated fifty lactobacilli from industrial ethanol fermentation tanks in Brazil and pointed out that more than 87% of them were able to produce biofilms in pure culture. Saunders et al. (2019) also highlighted that the substance produced by *Bacillus* strains acts specifically in *Lactobacillus* strains, without any prejudice to *S. cerevisiae* growth. Although that study was conducted with *Bacillus*, due to the high similarity between both groups, this enables us to hypothesize about a similar potential role of *Geobacillus* in the must. Also, we can speculate that *Geobacillus* might compete with more harmful microorganisms in the fermentation process when both of them are presented in the must, preventing them from prevailing in fermentation tanks.

The distillery conditions favor the presence of microbial contaminants throughout the ethanol fermentative process. Therefore methods of control and monitoring are needed to achieve the predicted fermentation yield. When evaluating the bacterial profile of contaminants in different productivity conditions, we observed differences in the population composition, particularly of the must samples, especially related to the *Lactobacillus* and *Geobacillus* genera. A better understanding of contaminants' dynamics can assist in the development of more accurate control, monitoring, and combat strategies focused on detecting potentially more harmful bacterial groups.

REFERENCES

- ALTSCHUL, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997.
- AMORIM, H. V. et al. Scientific challenges of bioethanol production in Brazil. **Applied Microbiology and Biotechnology**, v. 91, n. 5, p. 1267–1275, 2011.
- AMY APPRILL et al. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. **Aquatic Microbial Ecology**, v. 75, n. 2, p. 129–137, 2015.
- ANDERSON, M. J. Permutational Multivariate Analysis of Variance (PERMANOVA). **Wiley StatsRef: Statistics Reference Online**, p. 1–15, 2017.
- BASSO, L. C. et al. Yeast selection for fuel ethanol production in Brazil. **FEMS Yeast Research**, v. 8, n. 7, p. 1155–1163, 2008.
- BASSO, L. C.; BASSO, T. O.; ROCHA, S. N. Ethanol Production in Brazil: The Industrial Process and Its Impact on Yeast Fermentation. In: BERNARDES, M. A. DOS S. (Ed.). **Biofuel Production-Recent Developments and Prospects**. IntechOpen, 2011. 1530.
- BASSO, T. O. et al. Homo- and heterofermentative lactobacilli differently affect sugarcane-based fuel ethanol fermentation. **Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology**, v. 105, n. 1, p. 169–177, 2014.
- BASSO, T. O.; OLIVEIRA, F. S. de. Clash of Kingdoms: How Do Bacterial Contaminants Thrive in and Interact with Yeasts during Ethanol Production? In: **Fuel Ethanol Production from Sugarcane**. IntechOpen, 2018.
- BAYROCK, D. P.; INGLEDEW, W. M. Inhibition of yeast by lactic acid bacteria in continuous culture: Nutrient depletion and/or acid toxicity? **Journal of Industrial Microbiology and Biotechnology**, v. 31, n. 8, p. 362–368, 2004.
- BONATELLI, M. L. et al. Characterization of the contaminant bacterial communities in sugarcane first-generation industrial ethanol production. **FEMS Microbiology Letters**, v. 364, n. 17, 2017.
- BREXÓ, R. P.; SANT'ANA, A. S. Impact and significance of microbial contamination during fermentation for bioethanol production. **Renewable and Sustainable Energy Reviews**, v. 43, p. 423–434, 2017.
- CARØE, C. et al. Tagsteady: a metabarcoding library preparation protocol to avoid false assignment of sequences to samples. **bioRxiv 2020.01.22.915009**, p. 1–20, 2020.
- CARVALHO-NETTO, O. V et al. Saccharomyces cerevisiae transcriptional reprogramming due to bacterial contamination during industrial scale bioethanol production. **Microbial Cell Factories**, v. 14, n. 1, p. 13, 2015.
- CONAB, C. N. de A. **Acompanhamento da safra brasileira de cana-de-açúcar - V.5 - SAFRA 2018/19 - Quarto levantamento**. 2019
- COSTA, O. Y. A. et al. Microbial diversity in sugarcane ethanol production in a Brazilian distillery using a culture-independent method. **Journal of Industrial Microbiology and Biotechnology**, v. 42, n. 1, p. 73–84, 2015.
- COSTA, V. M. et al. Production of acetic acid, ethanol and optical isomers of lactic acid by *Lactobacillus* strains isolated from industrial ethanol fermentations. **Ciência e Agrotecnologia**, v. 32, n. 2, p. 503–509, 2008.
- DELLIAS, M. de T. F. et al. Biofilm formation and antimicrobial sensitivity of lactobacilli contaminants from

- sugarcane-based fuel ethanol fermentation. **Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology**, v. 111, n. 9, p. 1631–1644, 2018.
- FORTINA, M. G. et al. Reclassification of *Saccharococcus caldxylosilyticus* as *Geobacillus caldxylosilyticus* (Ahmad et al. 2000) comb. nov. **International Journal of Systematic and Evolutionary Microbiology**, v. 51, n. 6, p. 2063–2071, 2001.
- JUVONEN, R. Strictly anaerobic beer-spoilage bacteria. In: **Brewing Microbiology: Managing Microbes, Ensuring Quality and Valorising Waste**. Elsevier Ltd, 2015. p. 195–218.
- LAHTI, L.; SHETTY, S. **Microbiome R package**, [s.d.]. Disponível em: <<http://microbiome.github.io>>.
- LEE, M. D. Happy Belly Bioinformatics: an open-source resource dedicated to helping biologists utilize bioinformatics. **Journal of Open Source Education**. v. 2, p. 1–2, 2019.
- Zepeda-Mendoza. et al. DAME : a toolkit for the initial processing of datasets with PCR replicates of double - tagged amplicons for DNA metabarcoding analyses. **BMC Research Notes**, v. 9, n. 255, p. 1–13, 2016.
- LUCENA, B. T. L. et al. Diversity of lactic acid bacteria of the bioethanol process. **BMC Microbiology**, v.10, n. 1, p. 298, 2010.
- MAK, S. T. (Sarah). **Riesling, ferments, microbes - The development and application of NGS tolls in the context of winemaking**. 2018. University of Copenhagen, 2018.
- MCMURDIE, P. J.; HOLMES, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. **PLoS ONE**, v. 8, n. 4, 2013.
- MCMURDIE, P. J.; HOLMES, S. Waste not, want not: why rarefying microbiome data is inadmissible. **PLoS computational biology**, v. 10, n. 4, p. e1003531, 2014.
- MERCIER, C. et al. **SUMACLUSt: fast and exact comparison and clustering of sequences**, 2013. Available in : <http://metabarcoding.org/sumatra>.
- MURRAY, D. C.; COGHLAN, M. L.; BUNCE, M. From Benchtop to Desktop : Important Considerations when Designing Amplicon Sequencing Workflows. p. 1–21, 2015.
- MUTHAIYAN, A.; LIMAYEM, A.; RICKE, S. C. Antimicrobial strategies for limiting bacterial contaminants in fuel bioethanol fermentations. **Progress in Energy and Combustion Science**, v. 37, n. 3, p. 351–370, 2011.
- NASER, S. M. et al. Identification of lactobacilli by pheS and rpoA gene sequence analyses. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, n. 12, p. 2777–2789, 2007.
- NOVIK, G.; SAVICH, V.; MEEROVSKAYA, O. *Geobacillus* Bacteria: Potential Commercial Applications in Industry, Bioremediation, and Bioenergy Production. In: MISHRA, M. (Ed.). **Growing and Handling of Bacterial Cultures**. [s.l.] IntechOpen, 2019.
- OKSANEN, J. et al. **Vegan: Community Ecology Package**, 2019. . Disponível em: <<https://cran.r-project.org/web/packages/vegan/index.html>>.
- OLIVER, A. K. et al. Polymerase matters: Non-proofreading enzymes inflate fungal community richness estimates by up to 15%. **Fungal Ecology**, v. 15, p. 86–89, 2015.
- PARADA, A. E.; NEEDHAM, D. M.; FUHRMAN, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. **Environmental Microbiology**, v. 18, n. 5, p. 1403–1414, 2016.
- POKUSAEVA, K. et al. Novel bacteriocins produced by *Geobacillus stearothermophilus*. **Central European Journal of Biology**, v. 4, n. 2, p. 196–203, 2009.
- QUAST, C. et al. The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. **Nucleic Acids Research**, v. 41, n. November 2012, p. 590–596, 2013.
- QUEIROZ, L. L. et al. Dynamics of microbial contaminants is driven by selection during ethanol production. **Brazilian Journal of Microbiology**, 2019.
- RFA, R. F. A. **2019 Ethanol Industry Outlook**. Available in <https://ethanolrfa.org/wp-content/uploads/2019/02/RFA2019Outlook.pdf>.

- SAKAMOTO, K.; KONINGS, W. N. Beer spoilage bacteria and hop resistance. **International Journal of Food Microbiology**, v. 89, n. 2–3, p. 105–124, 2003.
- SAUNDERS, L. P. et al. Inhibition of Lactobacillus biofilm growth in fuel ethanol fermentations by Bacillus. **Bioresource Technology**, v. 272, n. October 2018, p. 156–161, 2019.
- SCHUBERT, M.; LINDGREEN, S.; ORLANDO, L. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. **BMC Research Notes**, v. 9, n. 1, p. 1–7, 2016.
- SKINNER-NEMEC, K. A.; NICHOLS, N. N.; LEATHERS, T. D. Biofilm formation by bacterial contaminants of fuel ethanol production. **Biotechnology Letters**, v. 29, n. 3, p. 379–383, 2007.
- SOUZA, R. B. De et al. Mineral Composition of the Sugarcane Juice and Its Influence on the Ethanol Fermentation. **Applied Biochemistry and Biotechnology**, v. 175, p. 209–222, 2015.
- STREIT, W. R.; SCHMITZ, R. A. Metagenomics - The key to the uncultured microbes. **Current Opinion in Microbiology**, v. 7, n. 5, p. 492–498, 2004.
- TABERLET, P. et al. Towards next-generation biodiversity assessment using DNA metabarcoding. **Molecular Ecology**, v. 21, n. 8, p. 2045–2050, 2012.
- TEAM, Rs. **RStudio: Integrated Development for R**. RStudio Inc., Boston, MA, 2015. Available in: <http://www.rstudio.com/>.
- ZEIGLER, D. R. The Geobacillus paradox : why is a thermophilic bacterial genus so prevalent on a mesophilic planet. **Microbiology**, v.160, p. 1–11, 2014.

SUPPLEMENTARY

Supplementary Figure 1

Figure S1: Ethanol distillery's scheme and sampling sites. In the distillery where the samplings were carried out, the fermentation was conducted in a continuous system with 3 connected tanks. Samples from must, between the heater exchange and the first fermentation tank, and from the first and third fermentation tank were collected – represented by the brown circles. Figure adapted from Schmidell & Facciotti 2001

Supplementary Table 1

Table S1 - Sample information. The collection code is in the first column, note that samples with the same code were collected on the same day. Following columns presents the information of the collection site, followed by the name of the sample and the productivity group to which the sample belongs. The last column shows the value of the ethanol ratio (produced / expected)

Type of Sample	Fermentation Line	Collection Code	Sample Name	Productivity Group	Produced/Expected Ratio
Tank 1	A	C8	1A.8	Low	0.91
Tank 1	B	C8	1B.8	Low	0.85
Tank 3	A	C8	3A.8	Low	0.91
Tank 3	B	C8	3B.8	Low	0.85
Must		C8	M.8	Low	0.88
Tank 1	A	C9	1A.9	Low	0.82
Tank 1	B	C9	1B.9	Low	0.83
Tank 3	A	C9	3A.9	Low	0.82
Tank 3	B	C9	3B.9	Low	0.83
Must		C9	M.9	Low	0.83
Tank 1	B	C10	1B.10	Low	0.67
Tank 3	B	C10	3B.10	Low	0.67
Tank 1	A	C10	1A.10	Normal	1.03
Tank 3	A	C10	3A.10	Normal	1.03
Tank 1	A	C11	1A.11	Normal	1
Tank 1	B	C11	1B.11	Normal	1
Tank 3	A	C11	3A.11	Normal	1
Tank 3	B	C11	3B.11	Normal	1.02
Must		C11	M.11	Normal	1.01
Tank 1	A	C12	1A.12	Normal	1.08
Tank 3	A	C12	3A.12	Normal	1.08
Must		C12	M.12	Normal	1.08
Tank 1	A	C13	1A.13	Normal	1.03
Tank 3	A	C13	3A.13	Normal	1.03
Must		C13	M.13	Normal	1.03
Tank 1	A	C14	1A.14	Normal	0.97
Tank 3	A	C14	3A.14	Normal	0.97
Must		C14	M.14	Normal	0.97
Tank 1	A	C15	1A.15	Normal	1.04
Tank 3	A	C15	3A.15	Normal	1.04
Must		C15	M.15	Normal	1.04
Tank 1	B	C16	1B.16	Low	0.8
Tank 3	B	C16	3B.16	Low	0.8
Must		C16	M.16	Low	0.8
Tank 1	A	C21	1A.21	Low	0.82
Tank 3	A	C21	3A.21	Low	0.82
Must		C21	M.21	Low	0.83
Tank 1	B	C22	1B.22	Low	0.76
Tank 3	B	C22	3B.22	Low	0.76
Must		C22	M.22	Low	0.76

Supplementary Table 2

Table S2 – Sequence of primers used for PCRs

Primer name	Sequence
Reverse Primers	
806R_1	TAGTCCAGGACTACNVGGGTWTCTAAT
806R_2	GTAACGTGGGACTACNVGGGTWTCTAAT
806R_3	TAAGACAGGACTACNVGGGTWTCTAAT
806R_4	ATAGATGGGACTACNVGGGTWTCTAAT
806R_5	GTAAGCAGGGACTACNVGGGTWTCTAAT
806R_6	ATTGTCAGGACTACNVGGGTWTCTAAT
806R_7	TCTGTTGGGACTACNVGGGTWTCTAAT
806R_8	ACCTACAGGACTACNVGGGTWTCTAAT
806R_9	GTGTATAGGACTACNVGGGTWTCTAAT
806R_10	CATACTAGGACTACNVGGGTWTCTAAT
Forward Primers	
515F_1	TAGTCCAGTGYCAGCMGCCGCGGTAA
515F_2	GTAACGTGGTGYCAGCMGCCGCGGTAA
515F_3	TAAGACAGTGYCAGCMGCCGCGGTAA
515F_4	ATAGATGGTGYCAGCMGCCGCGGTAA
515F_5	GTAAGCAGGTGYCAGCMGCCGCGGTAA
515F_6	ATTGTCAGTGYCAGCMGCCGCGGTAA
515F_7	TCTGTTGGTGYCAGCMGCCGCGGTAA
515F_8	ACCTACAGTGYCAGCMGCCGCGGTAA
515F_9	GTGTATAGGTGYCAGCMGCCGCGGTAA
515F_10	CATACTAGGTGYCAGCMGCCGCGGTAA

Supplementary Table 3

Table S3 – OTUs classification and number of reads in each sample

[illegible]

to be continued

[illegible]

to be continued

[illegible]

to be continued

Supplementary Table 4

Table S4 – OTUs classification and number of reads in must samples from belonging to Low and Normal productivity group

OTU	Kingdom	Phylum	Class	Order	Family	Genus	Productivity group	
							Low	Normal
OTU1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	46732	2120
OTU3	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	8146	135
OTU4	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	54562	1139
OTU5	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	1728	42936
OTU6	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	29	43252
OTU7	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megaspheara	340	7
OTU8	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	10440	54
OTU9	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	1304	125
OTU10	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	6885	3543
OTU11	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	784	5
OTU12	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	2534	100
OTU13	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	348	0
OTU14	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	478	10178
OTU15	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	233	18382
OTU16	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	27	6704
OTU17	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	Triticum	NA	1693	3349
OTU18	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	523	0
OTU19	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	473	24
OTU20	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	6402	0
OTU22	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus	0	4046
OTU23	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	3859	899
OTU24	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	282	22
OTU25	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	223	0
OTU26	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	178	0
OTU27	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	87	0
OTU28	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	228	0
OTU29	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	198	0
OTU30	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	89	0
OTU31	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	663	0
OTU32	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	4	3162
OTU33	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	892	1119
OTU34	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	Pantoea	0	2200
OTU35	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	17	0
OTU36	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	19	0
OTU37	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	138	0
OTU38	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	58	0
OTU39	Bacteria	Deinococcus-Thermus	Deinococci	Thermales	Thermaceae	Thermus	628	110
OTU40	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	22	33
OTU41	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	16	0
OTU42	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	123	0
OTU43	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	812	0
OTU44	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	37	0
OTU45	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	9	0
OTU46	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Acetobacter	1170	0
OTU48	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	43	0
OTU50	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	Enterobacter	180	520
OTU52	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	702	0
OTU53	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	24	0
OTU54	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	16	0
OTU55	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	14	0
OTU56	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	573	0
OTU58	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	23	0
OTU60	Bacteria	Actinobacteria	Actinobacteria	Corynebacterales	Corynebacteriaceae	Corynebacterium	406	490
OTU61	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	37	183
OTU62	Bacteria	Bacteroidetes	Bacteroidia	Sphingobacteriales	Sphingobacteriaceae	Sphingobacterium	933	0
OTU64	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	4	0
OTU65	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus	0	10
OTU67	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	15	0
OTU69	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	34	0
OTU70	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	579	0
OTU72	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	8	29
OTU73	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	441	0
OTU74	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	5	34
OTU76	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	114	10
OTU77	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	175	0
OTU78	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	54	0
OTU79	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	708	15
OTU80	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	126	0
OTU83	Bacteria	Aquificae	Aquificae	Hydrogenothermales	Hydrogenothermaceae	Sulfurihydrogenibium	517	0
OTU84	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	29	0
OTU85	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	43	0
OTU87	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	32	0
OTU88	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	30	0
OTU89	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	159	0
OTU90	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	7	0
OTU91	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	135	0
OTU92	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	670	0
OTU93	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Pseudoxanthomonas	477	0
OTU95	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	79	0
OTU97	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	18	40
OTU98	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	23	6
OTU99	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	6	0
OTU101	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	14	0
OTU102	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	134	118
OTU103	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	24	371

to be continued

OTU	Kingdom	Phylum	Class	Taxonomy			Productivity group	
				Order	Family	Genus	Low	Normal
OTU105	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	8	0
OTU106	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	74	0
OTU107	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	434	0
OTU109	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	118	0
OTU110	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	28	81
OTU111	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	589	0
OTU112	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	112	0
OTU113	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	4	0
OTU114	Bacteria	Firmicutes	Bacilli	Bacillales	NA	NA	379	0
OTU116	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	19	0
OTU117	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	894
OTU118	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	89	0
OTU120	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	83	0
OTU124	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	162	0
OTU126	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	463	0
OTU127	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	Actinidia	NA	322	0
OTU128	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	177	43
OTU129	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	20	0
OTU131	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	0	24
OTU132	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	199	0
OTU133	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	106	0
OTU135	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Enhydrobacter	285	0
OTU137	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	8	0
OTU138	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Diaphrobacter	0	53
OTU139	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	26	0
OTU143	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas	254	0
OTU144	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	80	25
OTU145	Bacteria	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	Dolosigranulum	31	216
OTU146	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	155
OTU147	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	252
OTU149	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	265	33
OTU150	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	233	0
OTU151	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	296
OTU155	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	Elodea	NA	0	401
OTU157	Bacteria	Firmicutes	Clostridia	Thermoanaerobacteriales	NA	NA	6	180
OTU160	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	74	0
OTU162	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	9	0
OTU163	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus	0	7
OTU164	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	147	0
OTU166	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0	363
OTU167	Bacteria	Firmicutes	Clostridia	Thermoanaerobacteriales	NA	NA	108	86
OTU172	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	81	105
OTU174	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	6	0
OTU176	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Zymomonas	124	52
OTU179	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	Dechloromonas	164	0
OTU180	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	12	0
OTU181	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	5	33
OTU183	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	384	0
OTU184	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	93	93
OTU185	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	20	0
OTU187	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	341
OTU189	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	123	0
OTU193	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	Cloacibacterium	60	68
OTU194	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	16	0
OTU195	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	20	0
OTU196	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	0	134
OTU197	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	8	30
OTU200	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	7	123
OTU201	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	139	0
OTU203	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	12	0
OTU204	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	Chryseobacterium	0	73
OTU205	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	9	97
OTU206	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	102	0
OTU207	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	158
OTU211	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	7	11
OTU214	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	174	0
OTU218	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	Triticum	NA	31	8
OTU220	Bacteria	Firmicutes	Bacilli	Bacillales	Alicyclobacillaceae	Alicyclobacillus	21	0
OTU221	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	0	36
OTU222	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	114	0
OTU223	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	108	0
OTU225	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	32	0
OTU226	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Idonella	5	41
OTU227	Bacteria	Proteobacteria	Alphaproteobacteria	Micavibrionales	uncultured	uncultured	109	0
OTU228	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	8	0
OTU230	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	4	0
OTU232	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	42	0
OTU234	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	179
OTU236	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	15	0
OTU238	Bacteria	Bacteroidetes	Bacteroidia	Sphingobacteriales	Sphingobacteriaceae	Sphingobacterium	101	0
OTU239	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	9	0
OTU241	Bacteria	Proteobacteria	Deltaproteobacteria	Desulfobivirionales	Desulfobivriionaceae	Desulfobivrio	100	0
OTU242	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	34	0
OTU243	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobaillus	4	0
OTU244	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	209
OTU245	Bacteria	Firmicutes	Clostridia	Thermoanaerobacteriales	NA	NA	94	0
OTU247	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	293

to be continued

OTU	Kingdom	Phylum	Class	Order	Family	Genus	Productivity group	
							Low	Normal
OTU249	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	C39	90	0
OTU250	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Pseudoxanthomonas	0	83
OTU251	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	9	0
OTU252	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	118	0
OTU253	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Catellibacillus	0	8
OTU254	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	111
OTU256	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	6	0
OTU257	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	130
OTU258	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococaceae	Micrococcus	38	48
OTU260	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	125
OTU261	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	78	0
OTU263	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	135
OTU265	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	104
OTU266	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	23	0
OTU269	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	49	0
OTU270	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	12	0
OTU272	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	14
OTU273	Bacteria	Firmicutes	Bacilli	Lactobacillales	Leuconostocaceae	Weissella	71	0
OTU274	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	19	0
OTU275	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	19	0
OTU276	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	94
OTU279	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	13	24
OTU280	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	19	24
OTU281	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	20	0
OTU283	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	17	0
OTU284	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	14
OTU286	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	24
OTU287	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Massilia	47	42
OTU288	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	71
OTU291	Bacteria	Firmicutes	Clostridia	Clostridiales	NA	NA	0	70
OTU292	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	60	0
OTU294	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	36	0
OTU297	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	52	0
OTU299	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	12	0
OTU301	Bacteria	Firmicutes	Clostridia	Thermoanaerobacteriales	NA	NA	53	0
OTU303	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Xanthomonas	0	67
OTU305	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	64	0
OTU307	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	30	0
OTU308	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Tannerellaceae	Macellibacteroides	64	0
OTU310	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	10	0
OTU314	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	99	0
OTU316	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	62	29
OTU318	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	71
OTU321	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Kosakonia	0	59
OTU322	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	58	0
OTU323	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	uncultured	5	0
OTU325	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0	58
OTU327	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weekseleaceae	Chryseobacterium	0	57
OTU330	Bacteria	Actinobacteria	Corynebacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	56	0
OTU331	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	76	0
OTU332	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	16	12
OTU334	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	140	0
OTU335	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	55	0
OTU338	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	0	54
OTU339	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	39	0
OTU340	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	uncultured	NA	30	0
OTU341	Bacteria	Proteobacteria	Gammaproteobacteria	Aeromonadales	Aeromonadaceae	Tolumonas	71	0
OTU342	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	43	0
OTU343	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	28	4
OTU349	Bacteria	Firmicutes	Bacilli	Lactobacillales	Leuconostocaceae	Leuconostoc	51	0
OTU354	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	7	41
OTU357	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	7	0
OTU361	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Filimonas	0	46
OTU363	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0	52
OTU372	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Citrobacter	0	61
OTU373	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	57
OTU375	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	53
OTU377	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	31	0
OTU378	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	73
OTU380	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Filimonas	0	41
OTU384	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	58	0
OTU386	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	41
OTU390	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	8
OTU391	Bacteria	Firmicutes	Clostridia	Thermoanaerobacteriales	Thermoanaerobacteraceae	Caldanaerobius	0	38
OTU393	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Vulcanibacterium	37	0
OTU394	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	37	0
OTU397	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	14	46
OTU399	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Lactococcus	36	0
OTU404	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobacter	0	35
OTU406	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	27	0
OTU408	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	4	31
OTU409	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	34	0
OTU410	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	34	0
OTU412	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	52	0
OTU414	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	34	0
OTU415	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	23	0
OTU418	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Sporichthyaceae	hgd	22	0

to be continued

OTU	Kingdom	Phylum	Class	Order	Family	Genus	Productivity group	
							Low	Normal
OTU420	Bacteria	Proteobacteria	Gammaproteobacteria	Pasteurallales	Pasteuraceae	Haemophilus	32	0
OTU423	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	7	0
OTU425	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0	31
OTU426	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	35
OTU427	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	31
OTU428	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	48
OTU432	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae	Solibacillus	31	0
OTU433	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	8	5
OTU437	Bacteria	Bacteroidetes	Bacteroidia	Sphingobacteriales	Sphingobacteriaceae	Sphingobacterium	30	0
OTU439	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	22	7
OTU440	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Paludibacteraceae	Paludibacter	29	0
OTU442	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	29	0
OTU448	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	22	0
OTU449	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	8	51
OTU450	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacteriales	Caulobacteraceae	uncultured	27	0
OTU451	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	32	0
OTU452	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	29	0
OTU455	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	27	0
OTU457	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Alcanivoracaceae	Alcanivorax	26	0
OTU458	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacteriales	Acetobacteraceae	Glucobacter	0	26
OTU463	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	36
OTU470	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	24
OTU473	Bacteria	Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	Alloiococcus	24	0
OTU475	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Dermatophilaceae	uncultured	0	12
OTU485	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	14
OTU491	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	15
OTU493	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	22
OTU495	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	22	0
OTU497	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	18	0
OTU498	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	22
OTU499	Bacteria	Epilobacteriota	Campylobacteria	Campylobacteriales	Sulfurospirillaceae	Sulfurospirillum	21	0
OTU500	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Pantoea	0	21
OTU502	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Burkholderia	0	21
OTU505	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	27
OTU506	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	21
OTU507	Bacteria	Firmicutes	Clostridia	Clostridiales	NA	NA	21	0
OTU508	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas	0	20
OTU509	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	0	20
OTU510	Bacteria	Proteobacteria	Gammaproteobacteria	Aeromonadales	Aeromonadaceae	Aeromonas	20	0
OTU511	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	uncultured	0	20
OTU518	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacteriales	Acetobacteraceae	Acetobacter	19	0
OTU519	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas	0	19
OTU520	Bacteria	Deinococcus-Thermus	Deinococcus	Thermales	Thermaceae	Thermus	19	0
OTU521	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	4
OTU528	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	31	0
OTU529	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Filimonas	0	18
OTU530	Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomycetes	18	0
OTU535	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	54
OTU538	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	12	0
OTU540	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Acidaminococcaceae	uncultured	18	0
OTU543	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	18
OTU545	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	0	17
OTU546	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia-Shigella	0	17
OTU547	Bacteria	Proteobacteria	Deltaproteobacteria	Desulfobacteriales	Desulfobulbaceae	Desulfobulbus	17	0
OTU549	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	uncultured	NA	0	17
OTU550	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	34
OTU551	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Microbacteriaceae	Curtobacterium	0	17
OTU552	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Methylophilaceae	Methylophilus	17	0
OTU553	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	28	0
OTU555	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	14	0
OTU562	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA	0	24
OTU564	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia-Shigella	0	16
OTU565	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia-Shigella	0	16
OTU570	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	16
OTU571	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	14
OTU573	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	8	0
OTU574	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	15
OTU575	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	15
OTU576	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Lautropia	15	0
OTU577	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Variovorax	15	0
OTU578	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	15	0
OTU581	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	15	0
OTU584	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	15	0
OTU592	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Pantoea	0	14
OTU593	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Enterobacter	0	14
OTU595	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	24
OTU597	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	C39	10	0
OTU598	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	C39	14	0
OTU604	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	14	0
OTU609	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	14	0
OTU613	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	14	0
OTU617	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	14
OTU618	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacteriales	Acetobacteraceae	Acetobacter	19	0
OTU619	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotellaceae	13	0
OTU620	Bacteria	Deinococcus-Thermus	Deinococcus	Thermales	Thermaceae	Thermus	13	0
OTU622	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	13	0
OTU624	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	4

to be continued

OTU	Kingdom	Phylum	Class	Taxonomy			Genus	Productivity group	
				Order	Family			Low	Normal
OTU628	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae		unculture.d	0	13
OTU629	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	13	0
OTU630	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	13	0
OTU631	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae		NA	0	29
OTU639	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	12	0
OTU640	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	13	0
OTU654	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae		Geobaillus	6	5
OTU657	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	12	0
OTU658	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	11	0
OTU660	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	11	0
OTU664	Bacteria	Deinococcus-Thermus	Deinococci	Thermales	Thermaceae		Thermus	6	0
OTU666	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	11	0
OTU667	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	11	0
OTU669	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae		NA	0	11
OTU670	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae		NA	0	16
OTU676	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae		Bacillus	0	13
OTU680	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae		Geobaillus	0	37
OTU688	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae		Lactobaillus	16	0
OTU696	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae		NA	0	17
OTU756	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae		Escherichia-Shigella	0	15
OTU757	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae		NA	0	13
OTU785	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae		Anoxybaillus	0	11

Supplementary Table 5

Table S5 – OTUs classification and number of reads in tank samples from belonging to Low and Normal productivity group

OTU	Kingdom	Phylum	Class	Taxonomy				Productivity group	
				Order	Family	Genus		Low	Normal
OTU1	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		420591	472899
OTU3	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		247664	195316
OTU4	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		231908	30937
OTU5	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		1060	33317
OTU6	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	24472
OTU7	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		42292	0
OTU8	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus		794	365
OTU9	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		22485	13610
OTU10	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		742	906
OTU11	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		22563	14003
OTU12	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		27376	13382
OTU13	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		16390	10377
OTU14	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		6	6055
OTU15	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	NA		0	8
OTU16	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	5663
OTU17	Bacteria	Cyanobacteria	Oxyphotobacteria	Chloroplast	Tricicum	NA		28	8
OTU18	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		9926	1440
OTU19	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		6050	2400
OTU20	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		101	0
OTU22	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus		0	4636
OTU23	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		702	36
OTU24	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		4270	2667
OTU25	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		6349	6281
OTU26	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		2763	1696
OTU27	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		3145	3183
OTU28	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		2942	1030
OTU29	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1732	1298
OTU30	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		2750	2730
OTU31	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		893	2168
OTU32	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	1623
OTU34	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Pantoea		7	7
OTU35	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1838	2017
OTU36	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		3457	3000
OTU37	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		2336	285
OTU38	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		2441	2680
OTU39	Bacteria	Deinococcus-Thermus	Deinococci	Thermales	Thermaceae	Thermus		58	760
OTU40	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		58	4821
OTU41	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		2062	1673
OTU42	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1717	148
OTU43	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		217	75
OTU44	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		997	1184
OTU45	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1682	1610
OTU46	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacteriales	Acetobacteraceae	Acetobacter		0	93
OTU48	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1208	655
OTU49	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		163	2406
OTU50	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Enterobacter		579	12
OTU51	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1228	856
OTU53	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1303	1334
OTU54	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1274	430
OTU55	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		975	975
OTU56	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		217	128
OTU57	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		36	1468
OTU58	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		828	553
OTU59	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		17	2174
OTU61	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	1722
OTU63	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1582	1103
OTU64	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1044	923
OTU65	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus		0	2173
OTU66	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1849	582
OTU67	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		884	843
OTU68	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1210	4
OTU69	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1607	662
OTU70	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		175	0
OTU71	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		461	639
OTU72	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		5	1173
OTU74	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		10	832
OTU75	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		765	194
OTU76	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		458	843
OTU77	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		641	287
OTU78	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		434	264
OTU80	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		467	67
OTU82	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		876	0
OTU83	Bacteria	Aquificae	Aquificae	Hydrogenothermales	Hydrogenothermaceae	Sulfurihydrogenibium		82	0
OTU84	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		784	608
OTU85	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		165	726
OTU86	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		1002	0
OTU87	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		682	0
OTU88	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		684	581
OTU89	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		302	514
OTU90	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		606	304
OTU91	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		379	69
OTU94	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	512
OTU95	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		340	242
OTU96	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		28	631
OTU97	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		17	583
OTU98	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	1251
OTU99	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1122	130

to be continued

OTU	Kingdom	Phylum	Class	Taxonomy				Productivity group	
				Order	Family	Genus		Low	Normal
OTU101	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		425	78
OTU103	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	7
OTU105	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		531	494
OTU106	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		396	238
OTU107	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		4	4
OTU108	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		537	133
OTU110	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	502
OTU113	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		721	373
OTU115	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		274	392
OTU116	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		917	518
OTU117	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	362
OTU119	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		590	107
OTU121	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		518	0
OTU122	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		1053	231
OTU123	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Virgibacillus		0	845
OTU125	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		372	0
OTU128	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		191	161
OTU129	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		437	0
OTU130	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		352	341
OTU131	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		7	425
OTU132	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus		108	0
OTU136	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		690	0
OTU137	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		405	108
OTU138	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Diaphorobacter		19	91
OTU139	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		62	216
OTU140	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		476	420
OTU141	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		347	341
OTU144	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	177
OTU146	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	687
OTU147	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	98
OTU148	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		351	49
OTU149	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		4	107
OTU151	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	95
OTU152	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		393	0
OTU153	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		4	582
OTU154	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Spirosomaceae	Pseudarcobacter		105	0
OTU156	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		5	1068
OTU158	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		292	230
OTU159	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	753
OTU161	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		5	552
OTU162	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		121	146
OTU163	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus		0	235
OTU165	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	337
OTU168	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		271	44
OTU169	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		415	37
OTU170	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		4	242
OTU171	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		228	0
OTU174	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		172	0
OTU175	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		100	165
OTU177	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		415	16
OTU180	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		20	189
OTU181	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	146
OTU182	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	374
OTU183	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		43	0
OTU185	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		84	137
OTU186	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		187	178
OTU187	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	120
OTU188	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		48	120
OTU189	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		29	0
OTU191	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Catellibacillus		0	236
OTU192	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		131	212
OTU193	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	Cloacibacterium		4	0
OTU194	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	198
OTU195	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		18	243
OTU197	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	218
OTU198	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		179	0
OTU199	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		275	0
OTU202	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	383
OTU203	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		186	11
OTU205	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	6
OTU207	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	217
OTU209	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	263
OTU210	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		234	244
OTU211	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		4	104
OTU212	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		185	322
OTU213	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		142	83
OTU215	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		122	16
OTU216	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		245	170
OTU217	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	362
OTU219	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		246	165
OTU221	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	106
OTU224	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus		0	316
OTU225	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		144	0
OTU228	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		244	0
OTU229	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		25	0
OTU230	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		7	166
OTU231	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	344
OTU233	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	uncultured		0	307

to be continued

OTU	Kingdom	Phylum	Class	Order	Family	Genus	Productivity group	
							Low	Normal
OTU235	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	163	41
OTU237	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	126	0
OTU239	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	103	13
OTU240	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	5	143
OTU242	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	59	0
OTU243	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	117	97
OTU244	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	44
OTU247	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	41
OTU248	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	255	44
OTU251	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	81	0
OTU253	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Catellibacillus	0	197
OTU256	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	119
OTU259	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	172
OTU260	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	76
OTU262	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	149	28
OTU266	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	68	0
OTU267	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera	99	0
OTU268	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	62	27
OTU270	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	74	27
OTU271	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	130	0
OTU272	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	165
OTU273	Bacteria	Firmicutes	Bacilli	Lactobacillales	Leuconostocaceae	Weissella	0	6
OTU278	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	113
OTU279	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	11
OTU280	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	62	17
OTU281	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	67
OTU282	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera	62	0
OTU283	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	55	0
OTU284	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	128
OTU286	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	60
OTU287	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Massilia	29	0
OTU289	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	172
OTU290	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	254
OTU295	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	98	0
OTU296	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	76	0
OTU297	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	21	53
OTU298	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	56	27
OTU299	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	40	0
OTU302	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	21	52
OTU306	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	161
OTU307	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	5	12
OTU311	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	132
OTU312	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	138
OTU313	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	68	38
OTU315	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	131	0
OTU317	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	31	75
OTU319	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	72
OTU320	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	9	56
OTU323	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	uncultured	0	208
OTU324	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera	68	0
OTU326	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	44	68
OTU328	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	133
OTU329	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	130
OTU332	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	44
OTU333	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	76	0
OTU337	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	43	11
OTU343	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	42	0
OTU345	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus	0	108
OTU346	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	101	0
OTU347	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	161	0
OTU348	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	71	0
OTU350	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Dialister	50	0
OTU352	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	76	0
OTU353	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	59
OTU355	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	75	43
OTU357	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	52	0
OTU358	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Catellibacillus	0	77
OTU362	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	57	0
OTU364	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus	0	110
OTU365	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	97	7
OTU366	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	40	69
OTU367	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	28	85
OTU368	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	45	0
OTU369	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	20	32
OTU370	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	70
OTU371	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	65	8
OTU374	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	51	0
OTU377	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	11	0
OTU381	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	87
OTU382	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus	0	78
OTU383	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	84	0
OTU385	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0	132
OTU387	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	60	0
OTU388	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	9	32
OTU389	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	81
OTU390	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus	0	90
OTU392	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	101	0
OTU395	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0	81
OTU396	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	5	71

to be continued

OTU	Kingdom	Phylum	Taxonomy				Genus	Productivity group	
			Class	Order	Family	Genus		Low	Normal
OTU587	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	13
OTU588	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	23
OTU589	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		24	0
OTU590	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		12	12
OTU591	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		15	4
OTU594	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	14
OTU596	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	19
OTU600	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus		14	0
OTU606	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		24	4
OTU610	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		14	0
OTU611	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		29	0
OTU612	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	14
OTU614	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	23
OTU616	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		14	0
OTU618	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacteriales	Acetobacteraceae	Acetobacter		0	5
OTU621	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Catellibacillus		0	13
OTU624	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	13
OTU625	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	13
OTU626	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		9	14
OTU627	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	13
OTU634	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	16
OTU635	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	12
OTU637	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Comamonas		12	0
OTU638	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	12
OTU640	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		5	0
OTU641	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	12
OTU642	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Diaphorobacter		0	12
OTU643	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		18	0
OTU644	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	14
OTU647	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	18
OTU648	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		12	0
OTU649	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		12	0
OTU650	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	12
OTU651	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		7	16
OTU652	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		5	7
OTU653	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Saccharococcus		0	20
OTU659	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	11
OTU661	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		17	0
OTU662	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		17	0
OTU663	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	11
OTU664	Bacteria	Deinococcus-Thermus	Deinococci	Thermales	Thermaceae	Thermus		5	0
OTU665	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		11	0
OTU668	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Spirosomaceae	Pseudocircella		11	0
OTU677	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	15
OTU678	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Geobacillus		0	14
OTU686	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		14	0
OTU693	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	uncultured		0	16
OTU694	Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Megasphaera		25	0
OTU710	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		20	0
OTU711	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		16	0
OTU712	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	17
OTU717	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		9	6
OTU725	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		5	9
OTU729	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	15
OTU737	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		14	0
OTU739	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	16
OTU742	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	15
OTU746	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	20
OTU750	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	15
OTU752	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		15	0
OTU765	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	12
OTU766	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		0	12
OTU789	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	12
OTU790	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anoxybacillus		0	17
OTU795	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus		0	15
OTU809	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Catellibacillus		0	11
OTU817	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		11	0
OTU819	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		6	8
OTU826	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		6	6
OTU841	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		17	0
OTU892	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		10	4
OTU942	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus		12	0

4. UNTARGETED METABOLOMICS OF FIRST GENERATION ETHANOL: COMPARING THE METABOLIC PROFILE OF NORMAL AND LOW ETHANOL YIELDS

Abstract

The great complexity of the large-scale industrial ethanol production leads to difficulty monitoring and consequently to possible productivity losses. This work aimed to explore the differences in the metabolic profile of an ethanol distillery under normal and low fermentative yield. Samples from must, fermentative tanks and acid treatment were collected in a Brazilian first-generation ethanol distillery. Statistical analysis of the gas chromatography–mass spectrometry (GC-MS) data resulted in the identification of 141 differently abundant metabolites. We can highlight the metabolite 5 beta-Dihydrotestosterone, which was detected in the must and in the fermentative tanks samples only under normal ethanol yield. By evaluating must samples, we can also highlight the potential of productivity monitoring of Oleic acid, Tagatose and 3-O-methyl-Glucose metabolites. Given this, we encourage future works to focus on the identification and quantification of those specific metabolites. With this work, we hope to contribute to the future development of new tools for quick and simple detection of selected metabolites to help distilleries' on fermentative monitoring and decision making.

Keywords: GC-MS, ethanol fermentation, monitoring markers

4.1 Introduction

Brazil is the second larger ethanol producer in the world, behind only the United States of America (RFA, 2019). According to the latest CONAB report, the country produced in the 2018/2019 sugarcane harvest season, more than 33 billion liters of ethanol, of which 1.82 billion liters were exported (CONAB, 2019). Brazilian ethanol production is essentially derived from sugarcane fermentation by selected strains of *Saccharomyces cerevisiae*. However, distilleries are non-aseptic environment and, consequently, microbial contaminants are expected, and even tolerated, to a certain extent (COSTA et al., 2015).

Bacteria and wild yeast are the main microbial contaminants of ethanol distilleries. They compete for space, sugar, and nutrients with selected yeast strains. Contaminants also produce compounds that inhibit the growth of selected yeast strains and can lead to flocculation events. (BAYROCK; INGLEDEW, 2004; CARVALHO-NETTO et al., 2015; BREXÓ; SANT'ANA, 2017). In the fermentative tanks, flocculation is responsible for ethanol yield reduction, the increase of gum and foam production, and the impairment in the centrifugation step at the end of fermentation. (BASSO; BASSO; ROCHA, 2011; CARVALHO-NETTO et al., 2015). When assessing contamination in a distillery that produces approximately 1 million liters of ethanol, Amorim et al (2011) estimated that contamination can lead to daily losses of up to 30,000 liters. In addition to contamination, another factor that can directly impact the fermentative performance is the feedstock composition. As reviewed by Basso and Lino (2018), the must can contain different proportions of nutrients, inhibitors, potentially toxic metals, organic acids, pH, and salts concentrations. Taking this into account, the development of efficient and faster mechanisms for the monitoring changes in the expected fermentation pattern has the potential to aid distilleries in making decisions to control the situation and consequently reducing losses.

Due to the quick and simple sample preparation, metabolomics detection methods may offer a solution for contaminants and inhibitory compounds monitoring. Metabolomics can be defined as a multidisciplinary science that aims to analyze the set of metabolites within a biological matrix (TRIVEDI; HOLLYWOOD; GOODACRE, 2017).

Due to its potential for evaluating hundreds of metabolites in complex biological samples at a time, metabolomics is currently used in several research fields, including the search of diagnostic and prognostic markers, investigation of biomolecules, evaluation of systemic responses, the *in situ* characterization of interactions between different microorganisms and their environment, the identification of metabolic pathways and also for food authentications (JOHNSON; IVANISEVIC; SIUZDAK, 2016a; BÖHME et al., 2019; DEMARQUE et al., 2020). Actually, some ethanol distilleries quantify specific metabolites in the fermented wine, such as sugars, organic acids, mannitol, and glycerol, to monitoring contaminants, yeast biomass formation and estimate the fermentation yield (BORGES et al., 2015; LOPES et al., 2016). This methodology is known as targeted metabolomics and can be applied for analytical validation, specific metabolites detection, and quantification (JOHNSON; IVANISEVIC; SIUZDAK, 2016). However, these metabolites are only quantified at the end of the process, reflecting the final fermentation result. To effectively assist in the ethanol yield monitoring, it is necessary to identify indicators able to detect important alterations before or throughout the fermentative process.

Untargeted metabolomics approaches are recommended for exploratory analysis and impartial evaluation of metabolites interactions, since its aim to acquire as many metabolite data as possible (JOHNSON; IVANISEVIC; SIUZDAK, 2016). Therefore, for a prospective analysis of differences in a metabolic pattern, an untargeted approach must be initially considered. The promising obtained results can then be used in further studies with targeted approaches and culminate in the development of fast metabolite detection technologies. An example of easily metabolite detector's, which we believe that can be applied for the ethanol industrial environment, are paper-based sensors. Paper-based biosensors can be applied to many areas including diagnostics, monitoring, compounds detection and quality control in health, industrial and environmental sciences (SINGH et al., 2018).

A commonly employed methodology for exploratory analysis and biomarker discovery is the gas chromatography coupled to mass spectrometry (GC-MS) with electron impact (EI) as ionization method (ZHANG et al., 2015; BEALE et al., 2017). This technique is robust, easy to operate, relatively cheap, and shows high sensitivity, specificity and reproducibility in metabolite profile evaluations (SUE et al., 2011; BEALE et al., 2017; ZHANG et al., 2020). Furthermore, the existence of libraries for EI-based GC-MS simplifies the identification of metabolites (PASIKANTI; HO; CHAN, 2008; ZHANG et al., 2020).

The present work aimed to explore the metabolic profile of a fermentation line from an ethanol distillery, under different fermentative yields. Sampling was conducted in a ethanol first-generation distillery, operating in a continuous fermentation line, between August and November 2017. The extracellular metabolites from must, fermentative tanks and yeast cells tanks were submitted to an untargeted GC-MS analysis.

4.2 Methods

4.2.1 Sampling

Sampling was performed between August and November 2017, in a continuous ethanol distillery in the southeast region of Brazil, Piracicaba - SP (22° 43' 30" S, 47° 38' 51" W). Specifically, a set of samples was collected on six days of Low productivity, and six days of Normal productivity. On each collection day, samples of must, fermentation tanks and acid treatment were collected (Figure S1).

Samples from must were collected by a tap in the pipe between the heat exchange and the first fermentation tank. Tank samples were collected from the bottom tap of each fermentation tank. Acid treatment samples were collected directly from the surface of the container. We followed Bonatelli et al (2017b) protocol where the collected samples were immediately cooled in ice and temperature was maintained at 4°C.

4.2.2 Sample preparation for GC-MS analysis

According to Bonatelli et al. (2017) a higher number of metabolites were reported in the polar fraction of fermented ethanol wine, so we focused specifically on the polar fractions of samples' metabolites. Upon arriving in the laboratory samples were centrifuged at 4000g (Allegra 64R - Beckman centrifuge) for ten minutes at 4°C. The metabolites were extracted following Bonatelli (2016), using 200 µL of the sample, 500 µL of a methanol: water solution (1:1 v/v) and 500 µL of dichloromethane. Along with solvents isotopic compounds were added as internal standard - succinic acid (D4, 98% - DLM 584-5), myristic acid (1, 2, 3 - 13C3, 99% - CLM 3665- 0.5) and palmitic acid (1,2,3,4 - 13C4), prepared at a concentration of 1 mg/mL. After extractions, samples were filtered through a 0.22 µm PVDF filter and stored at -80°C.

Prior to the derivatization procedure, 200µL of each sample was lyophilized, following by a derivatization step, according to the protocol employed by Budzinski et al (2019). Methoxyamination consists in an overnight reaction - addition 30 µL of methoxyamine hydrochloride (15 mg mL⁻¹) in pyridine - followed by the silylation step – the addition of 30 µL of N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) containing 1% trimethylchlorosilane (TMCS) – and after one hour the reaction was stopped with the addition of 30 µL of heptane.

4.2.3 GC- MS data analysis

Samples were analyzed with two technical replicates following Budzinski et al. (2019), alongside with extraction blanks and a series of n-alkanes (C12–C40), allowing the retention index calculation as proposed by Jonsson et al. (2005). One microliter of each derivatized sample was injected in the splitless mode in a 7890A gas chromatograph (Agilent Technologies, Santa Clara, USA), coupled to an automatic Comb-Xt processor (Leap Technologies, Carrboro, USA). The injection temperature was 280 °C and the septum purge flow rate was 20 mL min⁻¹ lasting for 60 seconds. The Helium flow rate through the column was 1 mL min⁻¹. The column temperature was maintained at 80°C for two minutes and then increased by 15 ° C every minute until reaching a temperature of 305°C, which was held for another 10 minutes. The column effluent was inserted into an ionization source of a gas chromatography with time-of-flight mass spectrometry (Pegasus 4D, Leco Corp., St. Joseph, USA) composed by two fused silica columns - first dimension column (Agilent DB- 5) 20 m long (0.18 mm inner diameter x 0.18 µm film); second dimension column 0.96 m (RXT- 170.10 mm inner diameter x 0.10 µm film). The temperature of the transfer line and ion source were 280 and 250 °C, respectively. An electron beam of 70 eV generated ions in an ionization current of 2.0 mA, and 10 spectra s⁻¹ were recorded in the mass range 45-800 m/z. The ChromaTOF v.4.51 (Leco Corp., St. Joseph, USA) software was used to perform baseline correction and export all MS files in NetCDF format.

The R package Target Search (CUADROS-INOSTROZA et al., 2009; R, 2013) was used for peak detection, retention time alignment and library matching. For metabolites identification, their retention indexes and spectra similarity (score greater than 600) were compared with compounds of the Golm Metabolome Database (KOPKA et al., 2005). After this step samples were grouped according to their collection site and productivity. For a metabolite to be considered valid, it should present at least three mass fragments at the correct RI, and present in at least 50% of samples from one group. For the following analyses, metabolite intensities were normalized using total ion chromatogram (TIC), and technical replicates were merged by average. The resulting matrix (Table S1) was Pareto scaled and log-transformed prior to statistical analysis. Statistical multivariate analysis (Principal Component Analysis - PCA) and univariate (Student t-test, False discovery rate - FDR - adjusted $p \leq 0.05$) analysis were done in MetaboAnalyst 4.0 software (XIA et al., 2009; CHONG et al., 2018).

4.2.4 Fermentative yield parameters

The ethanol fermentation distillery where the sampling was performed, works with a daily ethanol production target of approximately 600 m³ per day. Due to the constant input of feedstock and output of the fermented wine in the continuous fermentation line, the processing monitoring is temporal. Thus, a punctual fermentative yield estimative is needed to detect changes and guide industry decision-making. The ratio between the amount of ethanol produced and the amount of expected ethanol produced is then utilized.

The calculation of the expected amount of ethanol is expressed by the following formula:

$$Exp\ Ethanol = \frac{\left\{ \left(\frac{must_{TRS}}{100} \right) * 0.6475 * 0.89 * must_{flowrate} + \left(\frac{eth_{yeast}}{100} \right) * yeast_{flowrate} \right\}}{(must_{flowrate} + yeast_{flowrate})}$$

It takes into account the total reducing sugars present in must (*must*_{TRS}); the stoichiometric sugar to ethanol conversion factor 0.6475 - which equates to the production of 0.6475 liters of 100% ethanol for each kg of total reducing sugars; a theoretical efficiency of fermentation at 89%; the must flow rate; the percentage of ethanol detected in yeast cells in the acid treatment and the flow rate of yeast cells in the first fermentative tank.

In normal conditions, the value of this ratio should be close to 1. Measurements for ratio calculations are usually taken three times a day. Considering that small measurement deviation is expected, during this sampling period daily ratio values up to 0.93 were considered as normal fermentative yield by the distillery.

To facilitate understanding, in this study we named "Normal productivity" ratio values up to 0.93 and "Low productivity" below 0.93. It is important to highlight that ratio values were provided by the ethanol distillery where samples were collected.

4.3 Results

After data processing, a total of 213 metabolites were identified in the samples (Table S1). However, unidentified metabolites corresponded to 48.8% of the metabolites. According to Krumsiek et al (2012), unknown metabolites can be defined as small molecules whose detection and quantification are reproducible nonetheless with its chemical identity not yet understood. Since the main purpose of this work was to explore the differences in the relative abundance of metabolites under distinct productivity conditions and may contribute with future studies of productivity monitoring, the unknown metabolites were maintained in the following analyzes. The identified metabolites were assigned to 24 chemical classes, being "Carbohydrates and carbohydrate conjugates" and "Fatty Acyls" the main reported classes (Figure S2).

4.3.1 Investigation of differences in metabolic profile across the continuous fermentation line

To compare the overall metabolic profile and visualize if each sampling point had its distinct metabolic profile, a Venn diagram and exploratory principal component analysis (PCA) were performed with all collected samples, regardless of productivity (Figure 1). Of the 213 total, 92 were present in at least one sample from all the sampling points. (Figure 1A and Table S1). The group of samples that showed the highest number of exclusive metabolites was the must.

PCA showed that the must samples were grouped, indicating greater homogeneity in the group (Figure 1B). However, it was not possible to observe a similar pattern between any other sample type. Based on this visualization, it was not possible to identify a metabolic profile characteristic of each sampling point in the fermentation line, suggesting that the metabolite flow is intense throughout the continuous fermentation line.

Figure 1- Metabolic profile of each collection site. (1A) Venn diagram showing the of common and unique metabolites in each collection site. (1.B) PCA plot of extracellular metabolites profile. In the PCA the red dots represent acid treatment samples, green dots to must samples, purple dots to tank 1 samples, blue dots to tank 2 samples, and pink dots to tank 3 samples.

4.3.2 Metabolic profile of Low and Normal productivity

To check for global differences in the Normal and Low ethanol productivity metabolic profile, samples were first separated into groups considering only productivity parameters. Nonetheless, by PCA it was not possible to visualize any group segregation (Figure S3). Thus, for the following analyses, we considered the productivity profile within each different sampling site.

The visualization of differences in the metabolic profile at different productivity situations was based on PCA. It was possible to observe the separation of samples belonging to Normal and Low productivity groups across all the sampling points (Figure 2A-E). To investigate the differentially abundant metabolites, a Fold Change analyzes followed by a Student's t-test was applied. For statistical analyzes were considered the minimum fold change threshold 2 (Normal/Low) and t-tests threshold 0.05 (FDR corrected).

In the must samples (Table 1) 52 metabolites were reported of which 19 were more abundant in Low productivity conditions - presented as negative values for $\log_2(\text{FC})$ - and 33 in Normal productivity situations. In fermentation tanks during Normal productivity situations, 18, 25, and 16 metabolites were reported as differently abundant in tanks 1, 2, and 3, respectively (Table 2-4). It is possible to notice that the metabolites 5 beta-Dihydrotestosterone (ID1948) and Melibiose (ID 2147) were more abundant in all fermentative tanks. It is worth mentioning that in the must sample the metabolite 5 beta-Dihydrotestosterone was also reported as more abundant in situations of Normal productivity. On the other hand, the metabolite Melibiose was pointed out as more abundant in the must sample belonging to the low productivity group (Table 1).

Assessing differentially abundant metabolites in fermentative tanks 1, 2 and 3 during Low productivity situations, 7, 5, and 17 metabolites were identified, respectively (Table 2-4). It is possible to notice that the metabolite Androstendione (ID 2064) was detected as more abundant in Low productivity conditions in tank 3, however, in the other tanks its higher abundance is related to Normal productivity situations. Ultimately acid treatment analysis revealed 10 statistically relevant metabolites (Table 5), of which 5 were found to be more abundant in the Low and 5 in the Normal productivity group.

Figure 2- PCA of exometabolome in low and normal productivity. PCA plot of must (2.A) , tank 1 (2.B), tank 2 (2.C), tank 3 (2.D) and acid treatment (2.E). Low and Normal productivity samples are represented by red and green dots, respectively.

Table1: Metabolites differentially abundant in the Must samples belonging to Low and Normal productivity groups. This analysis considered the fold change with threshold 2 (Normal/Low) and the t-tests threshold 0.05 with FDR corrected.

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
2147	Melibiose	3.82E-06	-17.997	7.1E-12	11.149
1221	2-amino-2-deoxy-Glucose	9.87E-05	-13.307	0.0058152	2.2354
704	NA	0.000142	-12.784	0.044187	1.3547
1358	NA	0.000219	-12.159	0.0092923	2.0319
1823	N-Dansylaziridine	0.000332	-11.558	0.044187	1.3547
2532	NA	0.000495	-10.981	0.0058152	2.2354
2462	NA	0.000608	-10.683	0.044187	1.3547
1600	Octadecadienoic acid	0.000687	-10.508	0.0058152	2.2354
1364	NA	0.000847	-10.205	0.044187	1.3547
1597	Oleic acid	0.001443	-9.4372	5.2165E-08	7.2826
839	Tricarballic acid	0.001574	-9.3117	0.044187	1.3547
2503	NA	0.002717	-8.5239	0.0058152	2.2354
1751	Cystine	0.002768	-8.4969	0.0058152	2.2354
1391	Palmitic acid	0.002987	-8.3871	0.048383	1.3153
910	L-Dihydroorotic acid	0.003484	-8.1652	0.0058152	2.2354
173	Sarcosineamide	0.003607	-8.1151	0.047474	1.3235
2527	NA	0.004626	-7.756	0.044187	1.3547
270	Citraconic acid	0.010828	-6.5291	0.0093042	2.0313
2507	NA	0.1496	-2.7408	0.047474	1.3235
224	NA	66.532	6.056	0.047474	1.3235
1156	Meta-Tyrosine	276.34	8.1103	0.044187	1.3547
1749	Senecionine	748.6	9.5481	0.047474	1.3235
1067	Tagatose	875.21	9.7735	5.3898E-08	7.2684
1234	NA	899.59	9.8131	0.044187	1.3547
1394	NA	1051.9	10.039	0.028227	1.5493
1186	5-(4-hydroxybutyl)-Hydantoin	1411.3	10.463	0.010739	1.969
113	Phenylacetaldehyde	1586.5	10.632	0.047474	1.3235
1626	Stearic acid	1923.1	10.909	0.016811	1.7744
1594	6-(Z)-Octadecenoic acid	1993.6	10.961	0.016293	1.788
962	NA	2006	10.97	1.4124E-08	7.85
1071	NA	2770.9	11.436	0.0058152	2.2354
221	2-Ketohexanoic acid	2931.7	11.518	0.047474	1.3235
1340	NA	2997.5	11.55	0.044187	1.3547
1176	Galactose	3065.4	11.582	0.046609	1.3315
2562	NA	3668.6	11.841	0.047474	1.3235
1050	Methyl-beta-D-galactoside	3850.3	11.911	0.0058152	2.2354
252	NA	3914.5	11.935	0.046609	1.3315
197	NA	4550.6	12.152	0.047474	1.3235
225	3,6-Dimethyl-2,5-piperazinedione	4639.2	12.18	0.044187	1.3547
550	5-methyl-Cytosine	5832.5	12.51	0.0060744	2.2165
1585	3-Hydroxyhexadecanoic acid	8625.8	13.074	0.016293	1.788
982	3-O-methyl-Glucose	9714.1	13.246	1.163E-08	7.9344
453	2-(methylaminomethyl)Tartronic acid	12222	13.577	0.044187	1.3547
1982	Turanose	12809	13.645	0.0058152	2.2354
1045	NA	13080	13.675	0.0058152	2.2354
1947	alpha-D-Mannopyranosyl-(1,2)-D-mannopyranose	15461	13.916	0.044557	1.3511
1948	5 beta-Dihydrotestosterone	17416	14.088	2.815E-08	7.5505
593	L-3-Phenyllactic acid	23509	14.521	0.0058152	2.2354
935	NA	32869	15.004	0.044187	1.3547
2090	n-Octacosane	76803	16.229	0.044187	1.3547
58	2-Oxovaleric acid	92191	16.492	0.044187	1.3547
2115	NA	109120	16.736	0.047474	1.3235

Table2: Metabolites differentially abundant in the fermentative Tank 1 samples belonging to Low and Normal productivity groups.
 This analysis considered the fold change with threshold 2 (Normal/Low) and the t-tests threshold 0.05 with FDR corrected.

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
41	Hydroxylamine	0.00039548	-11.304	0.024391	1.6128
281	Alanine	0.000691	-10.499	0.0059863	2.2228
2003	NA	0.00093461	-10.063	0.0059863	2.2228
1939	beta-D-Fructofuranosyl-(2,1)-beta-D-Fructofuranose	0.00097693	-9.9995	0.024391	1.6128
1050	Methyl-beta-D-galactoside	0.0010159	-9.9431	0.024391	1.6128
2354	Cholestane, 3beta-hydroxy-, 5alpha-	0.0011066	-9.8196	0.024391	1.6128
1031	2-amino-3-ureido-Propanoic acid	0.0014728	-9.4072	0.024391	1.6128
1751	Cystine	395.14	8.6262	0.024391	1.6128
58	2-Oxovaleric acid	473.98	8.8887	0.024391	1.6128
375	NA	541.43	9.0806	0.024391	1.6128
1594	6-(Z)-Octadecenoic acid	763.31	9.5761	0.024391	1.6128
1847	1-O-Benzoyl-beta-D-glucoside	771.17	9.5909	0.024391	1.6128
1340	NA	954.01	9.8979	0.024391	1.6128
2147	Melibiose	1071.9	10.066	0.024391	1.6128
388	Erythrose	1111.7	10.119	0.0071113	2.1481
37	Valine	1148.4	10.165	0.0071113	2.1481
1149	Sebacic acid	1255.8	10.294	0.024391	1.6128
2064	Androstendione	1271.1	10.312	0.024391	1.6128
1156	Meta-Tyrosine	1436.8	10.489	0.024391	1.6128
1948	5 beta-Dihydrotestosterone	1558	10.606	0.024391	1.6128
1150	NA	2227.8	11.121	0.0059863	2.2228
270	Citraconic acid	2285.6	11.158	0.024391	1.6128
1045	NA	2561.8	11.323	0.024391	1.6128
102	Succinic acid semialdehyde	2845.9	11.475	0.024391	1.6128
453	2-(methylaminomethyl)Tartronic acid	4283.3	12.065	0.024391	1.6128

Table3: Metabolites differentially abundant in the fermentative Tank 2 samples belonging to Low and Normal productivity groups.
This analysis considered the fold change with threshold 2 (Normal/Low) and the t-tests threshold 0.05 with FDR corrected..

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
2354	Cholestane, 3beta-hydroxy-, 5alpha-	0.000539	-10.858	0.021869	1.6602
2560	NA	0.000628	-10.636	1.68E-10	9.7741
2187	Melibiose	0.000684	-10.513	0.021869	1.6602
1269	NA	0.001048	-9.8976	0.021869	1.6602
886	Lewulinic acid	0.003234	-8.2723	0.02285	1.6411
1948	5 beta-Dihydrotestosterone	217.03	7.7618	0.021869	1.6602
1921	Behenic acid	221.65	7.7922	0.021869	1.6602
1358	NA	384.12	8.5854	0.021869	1.6602
1617	NA	416.44	8.702	0.021869	1.6602
935	NA	455.19	8.8303	0.021869	1.6602
818	Methyl jasmonate	473.71	8.8879	0.021869	1.6602
2314	Quinic acid, 3-caffeoyl-, E-	524.08	9.0336	0.021869	1.6602
962	NA	653.06	9.3511	0.004633	2.3342
25	Lactic acid	668.79	9.3854	0.021869	1.6602
988	3-Hydroxydodecanoic acid	677.23	9.4035	0.021869	1.6602
37	Valine	728.79	9.5094	0.02285	1.6411
1772	6-Phosphogluconic acid	732.6	9.5169	1.33E-11	10.876
2196	NA	789.84	9.6254	2.4E-16	15.62
911	NA	870.36	9.7655	0.025314	1.5966
1810	1D-myo-inositol 2-phosphate	881.11	9.7832	1.68E-10	9.7741
2147	Melibiose	1002.3	9.9691	8.23E-12	11.085
388	Erythrose	1075.2	10.07	0.001895	2.7224
2064	Androstendione	1195.2	10.223	9.05E-08	7.0436
1208	Mannitol	1209.3	10.24	0.021869	1.6602
698	NA	1212.6	10.244	0.029711	1.5271
853	N-Acetylputrescine	1397.2	10.448	8.23E-12	11.085
836	N-Hexanoyl- homoserine lactone	1484.3	10.536	0.028431	1.5462
281	Alanine	1489.4	10.54	0.001895	2.7224
377	NA	1531	10.58	3.84E-10	9.4162
319	Tartronic acid	2546.9	11.315	0.001895	2.7224
276	NA	4470.9	12.126	0.001895	2.7224

Table4: Metabolites differentially abundant in the fermentative Tank 3 samples belonging to Low and Normal productivity groups.
This analysis considered the fold change with threshold 2 (Normal/Low) and the t-tests threshold 0.05 with FDR corrected.

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
2560	NA	0.000173	-12.496	1.5E-05	4.8233
841	NA	0.000208	-12.234	0.022326	1.6512
25	Lactic acid	0.000233	-12.07	0.003958	2.4025
1358	NA	0.00025	-11.966	0.021025	1.6773
2003	NA	0.000254	-11.942	0.021025	1.6773
1585	3-Hydroxyhexadecanoic acid	0.000272	-11.843	0.021025	1.6773
58	2-Oxovaleric acid	0.000351	-11.478	0.021025	1.6773
2064	Androstendione	0.000523	-10.901	0.021025	1.6773
482	NA	0.000563	-10.794	0.022326	1.6512
1176	Galactose	0.000601	-10.701	0.021025	1.6773
155	NA	0.000624	-10.647	0.021025	1.6773
127	Dihydroxyacetone	0.00063	-10.632	0.022398	1.6498
1285	NA	0.000633	-10.625	0.002848	2.5455
2401	Corticosterone	0.000636	-10.619	0.021025	1.6773
990	3,4-Dihydroxyphenylacetic acid	0.000821	-10.25	0.011014	1.9581
1030	N-heptanoyl-homoserine lactone	0.000902	-10.115	0.021025	1.6773
1234	NA	1274.9	10.316	0.021025	1.6773
927	Arabinonic acid	1302.2	10.347	0.021025	1.6773
1970	Lactose	1362.1	10.412	0.021025	1.6773
1948	5 beta-Dihydrotestosterone	1513.5	10.564	0.002848	2.5455
2147	Melibiose	1776.5	10.795	0.021025	1.6773
617	Alanylalanine	1785.3	10.802	0.021025	1.6773
1160	NA	2003.6	10.968	0.021025	1.6773
2187	Melibiose	2118	11.048	0.021025	1.6773
852	N-gamma-ethyl-Glutamine	2156.2	11.074	0.021031	1.6771
836	N-Hexanoyl- homoserine lactone	2209.5	11.11	0.022326	1.6512
1810	1D-myo-inositol 2-phosphate	2282.2	11.156	0.021025	1.6773
250	Glyceric acid	2451.2	11.259	0.002848	2.5455
2061	Androstendione	2476.9	11.274	0.002848	2.5455
1340	NA	3461.7	11.757	0.002848	2.5455
853	N-Acetylputrescine	3777.9	11.883	0.003371	2.4723
698	NA	7863	12.941	0.005638	2.2489
276	NA	10619	13.374	0.005497	2.2599

Table 5: Metabolites differentially abundant in the acid treatment samples belonging to Low and Normal productivity groups.
This analysis considered the fold change with threshold 2 (Normal/Low) and the t-tests threshold 0.05 with FDR corrected..

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
49	NA	0.000499	-10.968	0.022932	1.6396
988	3-Hydroxydodecanoic acid	0.00057	-10.777	0.022907	1.64
1045	NA	0.00068	-10.522	4.2572E-11	10.371
250	Glyceric acid	0.000816	-10.259	0.022907	1.64
1031	2-amino-3-ureido-Propanoic acid	0.002409	-8.6972	0.022907	1.64
927	Arabinonic acid	1641.2	10.681	0.022907	1.64
901	NA	1643.9	10.683	5.6327E-09	8.2493
911	NA	1738.9	10.764	0.042429	1.3723
1970	Lactose	3713.5	11.859	0.022907	1.64
2003	NA	18727	14.193	0.042429	1.3723

4.4 Discussion

The analysis of extracellular metabolites comprises the metabolites present in the medium and those secreted by cells. This technique allows for collecting information from cellular alterations in response to environmental changes using simple sample processing methodologies. Therefore, this approach can be easily applied to bioprocess monitoring, biomarkers prospection, metabolism alterations detections, and other research fields (PINU; VILLAS-BOAS, 2017). Untargeted metabolomics approaches allow the exploration of metabolites set and the identification of changes in samples metabolic profile, without prior knowledge of its metabolome (ZHANG et al., 2015; JOHNSON; IVANISEVIC; SIUZDAK, 2016). However, relating the identified metabolites to their biological role is the main challenge of metabolomics. This requires a clear understanding of the studied organism, the knowledge of the compound function in the affected metabolic pathway, determine which metabolic pathways are perturbed by the studied condition, and the metabolite's interactions (JOHNSON; IVANISEVIC; SIUZDAK, 2016). We are aware that changes in ethanol productivity may involve several biotic and abiotic factors. Due to the complexity of our collected samples, no inferences about changes in must composition, microbial metabolism or population dynamics could be done. In the following discussion, we presented the metabolites that showed a relevant and consistent relative abundance in their sampling group. We also had hypothesized about the possible dynamics of certain microbial groups based on the differences in metabolite patterns.

Aiming to assess the ethanol distillery overall metabolic profile, firstly the extracellular metabolites of sampling sites were analyzed. Of the 213 metabolites, 43% were detected in all evaluated sectors. We believe that this could be a consequence of the intense metabolic flux since in the continuous fermentative line the first tank is fed with the must and yeast cells and the other tanks are fed with effluent from the previous tank (FACCIOTTI, 2001). So it was not a surprise to identify a more homogeneous profile in must samples since the must is the system input.

It is common knowledge that the must composition can impact yeast metabolism and consequently the fermentative yield (BASSO; OLIVEIRA, 2018; REIS et al., 2018). Thus, we believe that the evaluation of differentially abundant metabolites in the must can be useful in the future for the development of productivity monitoring tools to check the feedstock and help in the decision making before the fermentation process starts. Among the most abundant metabolites in the must Low productivity samples, we can highlight Melibiose and Oleic acid. None of these metabolites were detected in samples belonging to the Normal productivity group, suggesting that these may be potential Low fermentation productivity indicators. Melibiose is a disaccharide composed of galactose and glucose and an intermediate of raffinose catabolism, a common component of plant tissues and molasse (MONIRUZZAMAN et al., 1997). Interestingly, in the beer industry melibiose is a marker sugar to identify yeast strains, since not all *Saccharomyces cerevisiae* yeasts can hydrolyze this sugar (BOKULICH; BAMFORTH, 2013). Besides some *Saccharomyces* strains, other microorganisms can metabolize this sugar. Thus, further studies may clarify if the presence of this compound is related to must composition or with differences in the must microbiota of Low and Normal productivity samples. Oleic acid derivatives are known to be growth factors for acid lactic bacteria and its antibacterial activity against *Staphylococcus*, *Streptococcus*, *Mycobacteria*, *Helicobacter*, *Campylobacter*, *Enterococcus*, *Listeria*, and *Bacillus* (HINTON A.; INGRAM, 2000; PARTANEN; MARTINEN; ALATOSSAVA, 2001; ZHENG et al., 2005). We can speculate that the presence of oleic acid in the must favor the growth of *Lactobacillus* while inhibiting the growth of other bacterial genera. This metabolite is also one of the major unsaturated fatty acids of *Saccharomyces cerevisiae* and is related to ethanol tolerance in yeast (YOU; ROSENFELD; KNIPPLE, 2003), its presence may also be related to the presence of wild yeast strains in the must.

Evaluating the relative abundance of metabolites in the must samples belonging to the Normal productivity group, we verified that the metabolites 5 beta-Dihydrotestosterone, 3-O-methyl-Glucose, Unidentified 962, and Tagatose were exclusively detected in this group. 5 beta-Dihydrotestosterone is a steroid and curiously has been described as a testosterone-related metabolite produced by coryneform bacteria in humans (NIXON et al., 1984). Since coryneform bacteria is a phylogenetically diverse group formed by several bacterial genera occurring in diverse environments (KIM; REBOLI, 2015), its detection may be related to the presence of a certain group of bacteria in must. This should be further investigated in order to clarify which bacteria may be involved with 5 beta-Dihydrotestosterone productions in this environment. Tagatose is an intermediate of lactose and galactose degradation and it is known that some strains of *Lactobacillus*, *Leuconostoc*, *Enterococcus*, *Streptococcus*, and *Lactococcus* are able to ferment this sugar (HANS BERTELSEN, HANS ANDERSEN, MICH, 2001). Admitting that all of these bacterial genera have already been reported in Brazilian distilleries (COSTA et al., 2015; QUEIROZ et al., 2019), we can speculate that the greater relative abundance of tagatose is related to differences in the contaminants metabolism, or in their distribution under distinct productivity situations. Although no biological supposition could be made about the role of unclassified metabolite and 3-O-Methyl glucose, those metabolites should be considered as potential markers of Normal productivity.

Since in the continuous ethanol fermentative line all the tanks are connected, the fermentative tanks metabolic profile were jointly assessed. Since the main purpose of this analysis was to detect differences in the relative abundance and draw attention to potential markers, here only the metabolites that followed the same pattern in all tanks were discussed. Interestingly, the metabolites beta-Dihydrotestosterone and Melibiose (ID 2147) were more abundant in Normal conditions of productivity in all the fermentative tanks, dissimilar to the abundance pattern of those metabolites in the must. However, the metabolite reported with the ID number 2187 also corresponds to Melibiose. Melibiose (2187) was detected as differentially abundant in tank 2 and tank 3 and presented a different abundance pattern from Melibiose (2147) in tank 2. According to GOLM database, the difference between both metabolites is the carbonyl group induced by the methoximation reaction. The lack of pattern in FC analysis of this disaccharide and the difference in detection regarding derivatization misgiving the efficiency of this metabolite as a marker. Accessing the metabolome of the acid treatment, only two unidentified metabolites (ID 1045 and 901) were exclusively detected in all samples belonging to the Low and Normal productivity group, respectively.

Owing to the complexity of the large-scale industrial process, the lack of knowledge of the organisms presented in each sample, and the absence of a controlled environment, no inferences about changes in metabolic pathways or microbial community were possible to be made. However, the identification of patterns in this diverse environment, suggests that in the future it is feasible to search for metabolic markers and compounds for fermentative efficiency monitoring to help in decision making and may reducing productivity losses. We are aware that our results are preliminary and the search for metabolic markers involves stages of validation and quantification. Since the must is the input of the fermentation system and showed a distinct metabolic profile in both productivity conditions, we suggest that future studies on metabolic productivity indicators focus on the identification of must compounds.

This work aimed to explore the metabolic differences in an ethanol distillery under distinct fermentation yields through untargeted GC-MS approaches. To our knowledge, this was the first study that correlates exploratory GC-MS data with the ethanol fermentative productivity in a Brazilian distillery. Even in a complex environment, it was possible to identify relevant differences in the metabolic profile of both conditions. We believe that future studies may investigate the use of some metabolites as productivity indicators.

References

- AMORIM, H. V. et al. Scientific challenges of bioethanol production in Brazil. **Applied Microbiology and Biotechnology**, v. 91, n. 5, p. 1267–1275, 2011.
- BASSO, L. C.; BASSO, T. O.; ROCHA, S. N. Ethanol Production in Brazil: The Industrial Process and Its Impact on Yeast Fermentation. In: BERNARDES, M. A. DOS S. (Ed.). **Biofuel Production-Recent Developments and Prospects**. IntechOpen, 2011. 1530.
- BASSO, T. O.; OLIVEIRA, F. S. de. Clash of Kingdoms: How Do Bacterial Contaminants Thrive in and Interact with Yeasts during Ethanol Production? In: **Fuel Ethanol Production from Sugarcane**. [s.l.] IntechOpen, 2018.
- BAYROCK, D. P.; INGLEDEW, W. M. Inhibition of yeast by lactic acid bacteria in continuous culture: Nutrient depletion and/or acid toxicity? **Journal of Industrial Microbiology and Biotechnology**, v. 31, n. 8, p. 362–368, 2004.
- BEALE, D. J. et al. **Review of recent developments in GC–MS approaches to metabolomics-based research**. [s.l.] Springer US, 2017. v. 14
- BÖHME, K. et al. Recent applications of omics-based technologies to main topics in food authentication. **TrAC - Trends in Analytical Chemistry**, v. 110, p. 221–232, 2019.
- BOKULICH, N. A.; BAMFORTH, C. W. The Microbiology of Malting and Brewing. **Microbiology and Molecular Biology Reviews**, v. 77, n. 2, p. 157–172, 2013.
- BONATELLI, M. L. **Caracterização da comunidade bacteriana contaminante do processo fermentativo para produção de etanol e o impacto no metaboloma da fermentação**. 2016. Universidade de São Paulo - Escola Superior de Agricultura “Luiz de Queiroz”, 2016.
- BONATELLI, M. L. et al. First-generation fermentation metabolomics by GC-MS: establishment of analysis methodology. In: XXI SIMPÓSIO NACIONAL DE BIOPROCESSOS ; XII SIMPÓSIO DE HIDRÓLISE ENZIMÁTICA DE BIOMASSA, Aracajú. **Anais...** Aracajú: 2017.
- BORGES, E. P. et al. The benefits of applied research: 37 years of discoveries, adaptations and solutions. **Sugar Industry**, n. April, p. 209–216, 2015.
- BREXÓ, R. P.; SANT'ANA, A. S. Impact and significance of microbial contamination during fermentation for bioethanol production. **Renewable and Sustainable Energy Reviews**, v. 43, p. 423–434, 2017.
- BUDZINSKI, I. G. F. et al. Integrated analysis of gene expression from carbon metabolism, proteome and metabolome, reveals altered primary metabolism in *Eucalyptus grandis* bark, in response to seasonal variation. **BMC Plant Biology**, v. 16, n. 1, p. 1–15, 2016.
- BUDZINSKI, I. G. F. et al. Network Analyses and Data Integration of Proteomics and Metabolomics From Leaves of Two Contrasting Varieties of Sugarcane in Response to Drought. **Frontiers in Plant Science**, v. 10, n. November, p. 1–19, 2019.
- CARVALHO-NETTO, O. V et al. *Saccharomyces cerevisiae* transcriptional reprogramming due to bacterial contamination during industrial scale bioethanol production. **Microbial cell factories**, v. 14, n. 1, p. 13, 2015.
- CONAB, C. N. de A. **Acompanhamento da safra brasileira de cana-de-açúcar - V.5 - SAFRA 2018/19 - Quarto levantamento**.
- COSTA, O. Y. A. et al. Microbial diversity in sugarcane ethanol production in a Brazilian distillery using a culture-independent method. **Journal of Industrial Microbiology and Biotechnology**, v. 42, n. 1, p. 73–84, 2015.
- CUADROS-INOSTROZA, Á. et al. TargetSearch - a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. **BMC Bioinformatics**, v. 10, n. 1, p. 428, 2009.
- DEMARQUE, D. P. et al. Mass spectrometry-based metabolomics approach in the isolation of bioactive natural products. **Scientific Reports**, v. 10, n. 1, p. 1051, 2020.
- FACCIO'TTI, M. C. R. Fermentação Contínua. In: SCHMIDELL, W. et al. (Ed.). **Biotecnologia Industrial Vol. 2 - Engenharia Bioquímica**. 1. ed. [s.l.] Blucher, 2001. p. 223–246.

- HANS BERTELSEN, HANS ANDERSEN, MICH. Fermentation of D-Tagatose by Human Intestinal Bacteria and Dairy Lactic Acid Bacteria. **Microbial Ecology in Health and Disease**, v. 13, n. 2, p. 87–95, 2001.
- HINTON A., J.; INGRAM, K. D. Use of oleic acid to reduce the population of the bacterial flora of poultry skin. **Journal of Food Protection**, v. 63, n. 9, p. 1282–1286, 2000.
- JOHNSON, C. H.; IVANISEVIC, J.; SIUZDAK, G. Metabolomics: beyond biomarkers and towards mechanisms. **Nature Reviews Molecular Cell Biology**, v. 17, n. July, p. 451–459, 2016.
- JONSSON, P. et al. High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. **Analytical Chemistry**, v. 77, n. 17, p. 5635–5642, 2005.
- KIM, R.; REBOLI, A. C. Other Coryneform Bacteria and Rhodococci. In: BENNETT, J. E.; DOLIN, R.; BLASER, M. J. (Ed.). **Principles and Practice of Infectious Diseases**. Saunders, 2015. p. 4320.
- KOPKA, J. et al. GMD@CSB.DB: The Golm metabolome database. **Bioinformatics**, v. 21, n. 8, p. 1635–1638, 2005.
- KRUMSIEK, J. et al. Mining the Unknown : A Systems Approach to Metabolite Identification Combining Genetic and Metabolic Information. **PLoS Genetics**, v. 8, n. 10, 2012.
- LOPES, M. L. et al. Ethanol production in Brazil: a bridge between science and industry. **Brazilian Journal of Microbiology**, v. 47, p. 64–76, 2016.
- MONIRUZZAMAN, M. et al. Extracellular melibiose and fructose are intermediates in raffinose catabolism during fermentation to ethanol by engineered enteric bacteria. **Journal of Bacteriology**, v. 179, n. 6, p. 1880–1886, 1997.
- NIXON, A. et al. Production of 5a and 5b-dihydrotestosterone by isolated human axillary bacteria. **FEMS Microbiology Letters**, v. 25, n. 2–3, p. 153–157, 1984.
- PARTANEN, L.; MARTINEN, N.; ALATOSSAVA, T. Fats and Fatty Acids as Growth Factors for *Lactobacillus delbrueckii*. **System. Appl. Microbiol.**, v. 24, p. 500–506, 2001.
- PASIKANTI, K. K.; HO, P. C.; CHAN, E. C. Y. Gas chromatography / mass spectrometry in metabolic profiling of biological fluids. **Journal of Chromatography B**, v. 871, p. 202–211, 2008.
- PINU, F. R.; VILLAS-BOAS, S. G. Extracellular microbial metabolomics: The state of the art. **Metabolites**, v. 7, n. 3, 2017.
- QUEIROZ, L. L. et al. Dynamics of microbial contaminants is driven by selection during ethanol production. **Brazilian Journal of Microbiology**, 2019.
- R. R: **A language and environment for statistical computing**. VienaR Core Team, , 2013.
- REIS, V. R. et al. Effects of feedstock and co-culture of *Lactobacillus fermentum* and wild *Saccharomyces cerevisiae* strain during fuel ethanol fermentation by the industrial yeast strain PE-2. **AMB Express**, v. 8, n. 1, 2018.
- RFA, R. F. A. 2019 **Ethanol Industry Outlook**. Available in: <https://ethanolrfa.org/wpcontent/uploads/2019/02/RFA2019Outlook.pdf>
- SINGH, A. T. et al. Paper-based sensors: Emerging themes and applications. **Sensors**, v. 18, n. 9, p. 1–22, 2018.
- SUE, T. et al. An exometabolomics approach to monitoring microbial contamination in microalgal fermentation processes by using metabolic footprint analysis. **Applied and Environmental Microbiology**, v. 77, n. 21, p. 7605–7610, 2011.
- TRIVEDI, D. K.; HOLLYWOOD, K. A.; GOODACRE, R. Metabolomics for the masses: The future of metabolomics in a personalized world. **New Horizons in Translational Medicine**, v. 3, n. 6, p. 294–305, 2017.
- YOU, K. M.; ROSENFELD, C.; KNIPPLE, D. C. Ethanol Tolerance in the Yeast *Saccharomyces cerevisiae* is Dependent on Cellular Oleic Acid Content. **Applied and Environmental Microbiology**, v. 69, n. 3, p. 1499–1503, 2003.
- ZHANG, A. et al. Metabolomics for Biomarker Discovery: Moving to the Clinic. **BioMed Research International**, v. 2015, 2015.
- ZHANG, X. et al. Mass spectrometry-based metabolomics in health and medical science : a systematic review. **The Royal Society of Chemistry**, v. 10, p. 3092–3104, 2020.

ZHENG, C. J. et al. Fatty acid synthesis is a target for antibacterial activity of unsaturated fatty acids. **FEBS Letters**, v. 579, n. 23, p. 5157–5162, 2005.

SUPPLEMENTARY

Supplementary Figure 1

Figure S1: Ethanol distillery's scheme and sampling sites. In the distillery where the samplings were carried out, the fermentation was conducted in a continuous system with 3 connected tanks. Samples from the tank where yeast cells stay after the acid treatment, from the must - between the heater exchange and the first fermentation tank - from the first, second and third fermentation tank were collected. The sampling sites are represented by the brown circles. Figure adapted from Schmidell & Facciotti 2001

Supplementary Figure 2

Figure S2- Reported metabolites chemical classification.

Supplementary Figure 3

Figure S3- PCA of global Low and Normal productivity profile. PCA plot of extracellular metabolic profile, regardless sampling site. Low and Normal productivity samples are represented by red and green dots, respectively.

Table S1- Metabolites intensities of extracellular metabolites

[illegible]

to be continued

Sample	Sampling site	Group	313	319	322	375	377	388	392	453	482	509	532	550	593	595	606	611	617	619	628	652	653	661	679	698
L 1.1	Must	Low Must	0	0	0	0.475035	0	0.395544	0	0	0	1.791311	0.001974	1.659005	0.378878	0	1.881561	0.698523	0.15008	0	0	0	0	0	0	
L 1.1A	Tank 1 A	Low T1	0	0	0	0.300185	0	0.300185	0	0	0	2.202942	7.022072	2.477881	2.510382	0	1.449334	0.592109	1.612178	0.565811	0	0	0	0	0	
L 1.1B	Tank 1 B	Low T1	0	0	0	0.400761	0	0	0	0	0	1.831531	0.010946	1.13459	0.73305	1.404469	10.6501	6.092109	1.612178	0.565811	0	0	0	0	0	
L 1.2	Tank 2 A	Low T2	0	0	0	0.655854	0	0	0	0	0	2.818599	6.674551	2.090703	3.689462	2.651233	2.651233	2.651233	2.651233	2.651233	0	0	0	0	0	
L 1.2A	Tank 2 A	Low T2	0	0	0	0.413991	28.92941	0	0	0	0	2.767507	1.674551	15.56492	20.20603	7.38409	10.8142	1.841156	2.37004	2.12707	0	0	0	0	0	
L 1.3	Tank 3 A	Low T3	0	0	0	90.16152	0	0	0	0	0	2.509054	0.169498	0	0.205931	1.737859	8.60601	0	0	0	0	0	0	0	0	
L 1.3A	Tank 3 A	Low T3	0	0	0	13.34691	0	0	0	0	0	7.196334	3.95002	10.86132	8.16138	10.86679	3.44379	0	0	0	0	0	0	0	0	
L 1.3B	Tank 3 B	Low T3	0	0	0	17.77114	0	0	0	0	0	1.184607	0.443432	7.99507	5.746466	7.99507	10.12586	10.12586	10.12586	10.12586	0	0	0	0	0	
L 1.F	Acid treatment	Low Acid treatment	0	0	0	6.888957	0	20.3352	0	0	0	8.44409	4.44372	7.98221	13.79507	5.746466	7.99507	10.12586	10.12586	10.12586	10.12586	0	0	0	0	
L 2	Must	Low Must	0	0	0	0.99819	0	0.205811	0	0	0	6.36283	0.26948	5.952166	0.067011	0	0	0	0	0	0	0	0	0	0	
L 2.1A	Tank 1 A	Low T1	1.862953	0	1.744881	0	5.5155	0	2.151774	0	0	1.044002	0.919517	1.21479	1.167602	1.21479	0.741773	2.560592	2.780197	2.95586	0	0	0	0	0	
L 2.1B	Tank 1 B	Low T1	5.428562	0	4.74453	0	7.392116	0	3.717743	0	0	1.11406	0.919517	3.27933	1.345192	1.337617	2.68887	2.68887	2.68887	2.68887	0	0	0	0	0	
L 2.2	Tank 2 A	Low T2	4.785619	0	4.139818	0	4.739574	0	0.997598	0	0	2.461915	3.845032	6.66594	8.241786	2.429587	1.762389	5.06156	8.18721	4.134547	4.134547	0	0	0	0	
L 2.2A	Tank 2 A	Low T2	0	0	0	0.291797	0	0	0	0	0	0.711295	0.201464	2.8012	0	0.738656	0	0	0	0	0	0	0	0	0	
L 2.3	Tank 3 A	Low T3	6.081459	0	5.5386	0.698677	4.202904	0	4.1304	0	0	1.717445	1.74709	5.820495	5.967294	3.579533	0	0	0	0	0	0	0	0	0	
L 2.3A	Tank 3 A	Low T3	2.951907	0	2.858049	0	1.514442	0	1.514442	0	0	0.711295	0.201464	2.8012	0	0.738656	0	0	0	0	0	0	0	0	0	
L 2.B	Acid treatment	Low Acid treatment	2.951907	0	2.858049	0	1.514442	0	1.514442	0	0	0.711295	0.201464	2.8012	0	0.738656	0	0	0	0	0	0	0	0	0	
L 3	Must	Low Must	0	0	0	0.558446	0	0.126132	0	0	0	1.697551	0.310858	4.629511	3.007123	0	1.852331	2.146655	4.629511	3.007123	0	0	0	0	0	
L 3.1	Tank 1 A	Low T1	0	0	0	0.988911	0	0.638417	0.432026	0	0	18.2079	0.788429	13.49705	2.525674	4.197885	0	0	0	0	0	0	0	0	0	
L 3.2	Tank 2 A	Low T2	0	0	0	3.55444	2.921079	0	0.462477	0	0	1.797236	3.797912	4.767062	3.841116	2.561897	0	0	0	0	0	0	0	0	0	
L 3.2A	Tank 2 A	Low T2	0.701988	0	2.502191	1.071386	2.986743	2.326582	1.866331	0.471007	0	1.479236	3.797912	4.767062	3.841116	2.561897	0	0	0	0	0	0	0	0	0	
L 3.3	Tank 3 A	Low T3	13.2418	0	10.21703	0	19.54884	0	0	0	0	3.89521	7.226052	7.226052	3.891587	3.171445	4.987601	6.87216	4.907189	6.297836	0	0	0	0	0	
L 4	Must	Low Must	0	0	0	0	0	0	0	0	0	0.700047	0.015061	0	0.08058	0	0	0	0	0	0	0	0	0	0	
L 4.1A	Tank 1 A	Low T1	8.365889	0	6.995832	0	1.891483	22.15201	0	0	0	1.610407	7.642339	11.82148	29.29558	5.952585	11.6679	6.780755	8.895538	6.50376	7.908622	0	0	0	0	
L 4.1B	Tank 1 B	Low T1	7.795813	0	7.116508	0	0.83207	0	0	0	0	1.021594	10.43631	14.41312	10.38693	9.52688	19.73972	9.210526	11.6679	6.780755	8.895538	6.50376	7.908622	0		
L 4.2A	Tank 2 A	Low T2	4.048312	0	4.425453	0	2.230773	30.95956	2.866079	8.856564	9.899988	0.666998	0.977585	8.049041	11.80354	6.877955	11.11117	5.58727	7.589737	5.205059	0	0	0	0	0	
L 4.3A	Tank 3 A	Low T3	3.33255	0	3.307502	0	0.781377	32.45299	0	0	0	0.568355	1.70629	4.879529	7.342672	0	11.16799	4.786976	5.91295	4.624755	5.20429	0	0	0	0	
L 4.F	Acid treatment	Low Acid treatment	0	0	0	2.500991	0	0.149553	0.366498	0	0	10.47949	0.332313	5.44489	0.214966	0	0	0	0	0	0	0	0	0	0	
L 5	Must	Low Must	0	0	0	0.065383	0	0.26757	0.951799	0	0	1.931213	5.447113	5.57131	7.57618	6.85729	1.697355	2.202621	1.387142	0	0	0	0	0	0	
L 5.1A	Tank 1 A	Low T1	0.270821	0	0.461552	0	3.754546	0	29.06318	0	0	1.931213	5.447113	5.57131	7.57618	6.85729	1.697355	2.202621	1.387142	0	0	0	0	0	0	
L 5.2A	Tank 2 A	Low T2	0.943456	0	1.086438	0	85.27076	0	1.086438	0	0	3.903021	5.447113	5.57131	7.57618	6.85729	1.697355	2.202621	1.387142	0	0	0	0	0	0	
L 5.3A	Tank 3 A	Low T3	4.702978	0	4.677055	0	2.480217	70.55116	2.456958	23.43037	0.976683	11.33622	0.033541	20.78557	31.10377	66.41393	0	0	0	0	0	0	0	0	0	
L 5.F	Acid treatment	Low Acid treatment	4.702978	0	4.677055	0	2.480217	70.55116	2.456958	23.43037	0.976683	11.33622	0.033541	20.78557	31.10377	66.41393	0	0	0	0	0	0	0	0	0	
L 6	Must	Low Must	0	0	0	4.882703	0	0.171732	0.512189	0	0	4.781933	0.061541	0	0	0	0	0	0	0	0	0	0	0	0	
L 6.1B	Tank 1 B	Low T1	4.324669	0	14.7194	0	0.947674	6.615124	0	0	0	1.882532	0	0	9.386211	11.85529	9.1387	10.61361	21.48299	10.2482	9.99962	10.12512	0	0	0	
L 6.2B	Tank 2 B	Low T2	3.096971	0	14.55489	0	0.947674	6.615124	0	0	0	1.882532	0	0	9.386211	11.85529	9.1387	10.61361	21.48299	10.2482	9.99962	10.12512	0	0	0	
L 6.3B	Tank 3 B	Low T3	0	0	0	0.023021	0.856145	0.015846	0.111959	0.138552	0	0	0.432177	0.242199	0.27698	0	1.169468	0.41387	0.165436	0.04444	0	0	0	0	0	
L 6.F	Acid treatment	Low Acid treatment	2.416434	0	1.812755	0	0.49843	0	1.686306	0.561819	0	0	1.050592	1.279999	0.329917	0.155856	1.061272	0.279847	0.15784	0.092037	0.001065	0.564466	0	0	0	
N 1	Must	Normal Must	0	0	0	0.482668	0	0.100931	0.109561	0	0	3.053735	0.054635	2.638848	0.099037	3.420823	0	0	0	0	0	0	0	0	0	
N 1.1	Tank 1 B	Normal T1	0.075089	0	0.076628	0.348887	0.041387	0.32823	0.425748	0.107403	0	1.026151	0	0	1.216422	1.32021	0.66552	0.66552	0.66552	0.66552	0	0	0	0	0	
N 1.2B	Tank 2 B	Normal T2	15.87134	0	13.84079	0	11.4492	0	9.8167	0	0	6.31371	8.937409	4.32875	3.56356	3.877416	9.20467	11.83327	5.24431	3.186138	0	0	0	0	0	
N 1.3B	Tank 3 B	Normal T3	1.128944	0	4.98319	11.89752	5.659674	5.596606	0	0	0	9.43888	12.26031	2.808525	0	1.651579	5.132404	5.87982	3.45535	4.78196	0	0	0	0	0	
N 1.F	Acid treatment	Normal Acid treatment	1.128944	0	4.98319	11.89752	5.659674	5.596606	0	0	0	9.43888	12.26031	2.808525	0	1.651579	5.132404	5.87982	3.45535	4.78196	0	0	0	0	0	
N 2	Must	Normal Must	0	0	0	0.153332	0	0.084122	0.121699	0	0	0.135886	0.167089	0.404398	0.505313	0	2.417382	2.157295	1.687006	1.8257	1.749623	0	0	0	0	
N 2.1B	Tank 1 B	Normal T1	4.283114	0	3.1308	0.844466	0	0	60.74	0	0	0.135886	0.167089	0.404398	0.505313	0	2.417382	2.157295	1.687006	1.8257	1.749623	0	0	0	0	
N 2.1B	Tank 1 B	Normal T1	4.283114	0	3.1308	0.844466	0	0	60.74	0	0	0.135886	0.167089	0.404398	0.505313	0	2.417382	2.157295	1.687006	1.8257	1.749623	0	0	0	0	
N 2.2B	Tank 2 B	Normal T2	8.146911	0	18.7056	25.1891	17.9516	7.82952	5.24604	0	0	4.18419	0.645818	3.712387	4.97759	1.84324	0	0	0	0	0	0	0	0	0	
N 2.3B	Tank 3 B	Normal T3	9.146911	0	21.7820	20.031819	10.69436	8.462146	11.33511	0	0	6.02134	9.684199	2.74548	0	2.705451	0	0	0	0	0	0	0	0	0	
N 2.F	Acid treatment	Normal Acid treatment	25.28782	0	27.98546	18.7035	0	0.38988	0	0	0	1.087556	0.082107	0.1216374	1.277966	0.0728	1.260898	0.127981	0	0	0	0	0	0	0	
N 3	Must	Normal Must	0	0	0	4.47944	0	0.075984	0.170508	0.427771	0.041335	0	0	0.262191	0.306574	0.336249	0	0	0	0	0	0	0	0	0	
N 3.1A	Tank 1 A	Normal T1	11.85145																							

Sample	Sampling site	Group	704	718	733	743	818	824	836	839	841	853	856	858	859	864	866	901	902	908	910	911	927	935	937	957	962		
L 1	Must	Low Must	0	0	0.633603	0	0.469538	0.316144	0	0.085567	0.301200	0.083332	0	0	0.19881	0.168966	0	0.875126	0	0	0	0	0	0	0	52.1176	0		
L 1A	Tank 1A	Low T1	0	0.453927	0	0	0	0	0	0.307497	0	0	0	0.559484	1.025256	0	0	0	4.594259	0	0	0	0	0	1.845639	1.180007	0.7233		
L 1B	Tank 1B	Low T1	0	0.622604	0	0	0	0	0	0	0.314991	0	0	0.314991	0.871607	0	0	0	0	0	0	0	0	0	58.416233	41.83527	0		
L 1.2A	Tank 2 A	Low T2	0	0.695464	0	0	0	0	0	0	0	0	0	1.614419	2.428988	0	0	0	0	0	0	0	0	0	0.205707	1.892842	0		
L 1.2B	Tank 2 B	Low T2	0	0.831582	0	0	0	0	0	0	0	0	0	1.913559	4.349911	0	0	0	0	0	0	0	0	0	11.25622	11.45793	0		
L 1.3A	Tank 3 A	Low T3	1.640903	0	0	0	8.514099	0	0	0	40.0451	0	0	0.4029748	2.033814	0.569374	5.713395	1.583161	2.686841	0	0	0	0	0	5.893528	8.508618	1.161156		
L 1.3B	Tank 3 B	Low T3	0	1.878607	0	0	0	0	0	0	14.68573	0	0	0.895486	3.057103	0	0	0	0	0	0	0	0	0	1.695921	4.078214	1.448908		
L 1F	Acid treatment	Low Acid treatment	9.765381	0.644667	2.05331	0	0.505382	3.294021	0	0.81719	0	0.105912	0	0.849038	0	0	0	0	0	0	0	0	0	0	2.17729	1.11819	0.483741		
L 2	Must	Low Must	0	0	0	0	0	0	0	0	0	0.986793	0	0	0	0.012426	0	0	0	0	0.222872	0	0	0	0	25.04991	0		
L 2.1A	Tank 1 A	Low T1	0	6.03537	0	0	0	0	0	0	4.464113	0	0	1.79811	3.468222	0	0.644314	0	1.51	0	0	0.30524	0	0	0	15.52195	0		
L 2.1B	Tank 1 B	Low T1	0	8.994372	0	0	0	0	0	0	0.648994	0	0	4.763019	2.013555	0	10.17026	5.055053	0	4.56416	0	0	0.914477	6.646666	2.78003	0	16.84154		
L 2.2A	Tank 2 A	Low T2	0	8.927153	0	0	0	0	0	0	16.87504	0	0	0.76692	4.01764	0	10.17026	5.055053	0	0	0	0	0	0	0	2.443614	0		
L 2.2B	Tank 2 B	Low T2	0	8.795531	0	0	0	0	0	0	0.589167	1.1316	0	0.589167	1.1316	0	10.17026	5.055053	0	0	0	0	0	0	0	96.78169	0		
L 2.3A	Tank 3 A	Low T3	0	14.76958	0	0	0	0	0	0	6.992725	0	0	2.687443	4.40846	4.143629	3.391074	0	2.15344	0	0	0	0	0	0	3.950479	20.96194		
L 2.3B	Tank 3 B	Low T3	8.826466	4.446601	0	0	0	0	0	0	21.4248	0	0	6.687443	4.40846	4.143629	3.391074	0	14.55702	0	0	0	0	0	0	41.66208	70.09028		
L 2F	Acid treatment	Low Acid treatment	0	0	0	0	0	0	0	0	6.992725	0	0	31.20443	6.102317	6.535176	0	0	3.78602	0	0	0	0	0	5.170997	0	15.08163		
L 3	Must	Low Must	3.483867	0	0	0	0.345975	2.531867	0	0.619134	0.992656	0	0	0.121713	0.159245	0	0.242255	0	0	1.821229	0	0	0	0	0	18.53507	17.74961		
L 3.1B	Tank 1 B	Low T1	0	0	0	0	0	0	0	0	2.9079	0	0	0.44761	0	0	0.452345	0	15.64587	0	0	0	0	0	0	88.89042	0		
L 3.2B	Tank 2 B	Low T2	0	5.145206	0	0	0	0	0	0	0	0	0	0.248637	0.34237	0	3.143142	0.303082	0	0	0	0	0	0	0	36.7422	52.00685		
L 3.3B	Tank 3 B	Low T3	0	3.092937	0	0	0	0	0	0	5.605904	0	0	15.36295	8.841797	1.862275	3.847793	0	0	0	0	0	0	0	0	8.641621	26.07437		
L 3F	Acid treatment	Low Acid treatment	0	13.29832	0	0	0	0	0	0	18.10719	14.10446	0	35.99908	0	0	4.568464	0	0	0	0	0	0	0	0	0	13.6951		
L 4	Must	Low Must	0	0.173572	0.324454	0	0.019959	0.1177	0	0	0.0228	0.048991	0	0	0.019411	0.07788	0	0	0.565451	0	0	0	0	0	0	1.082436	5.53817		
L 4.1A	Tank 1 A	Low T1	0	0.955165	0	0	0	0.700954	0	0	0.316293	0	0	196.0664	104.8727	0	12.37682	0	4.863818	0	0	0	0	0	0	4.57866	2.665549		
L 4.2A	Tank 2 A	Low T2	0	1.439315	0	0	0	0	0	0	0.087636	0	0	0	237.2215	0	15.767	0	0	0	0	0	0	0	0	5.700609	10.93778		
L 4.3A	Tank 3 A	Low T3	2.024269	0.19431	0	0	0	0	0	0	0.677102	0	0	1.603313	236.5614	132.0192	11.82156	12.391074	0	2.463136	0	0	0	0	0	4.781841	0		
L 4F	Acid treatment	Low Acid treatment	0	0	0	0	0	0	0	0	0.928788	17.99132	0	0	443.4516	0	0	12.76755	0	0.691151	0	0	0	0	0	0	1.621641	0	
L 5	Must	Low Must	10.0871	2.873592	2.246962	0	0.502953	1.864342	0	0.783645	0.707089	0.7133	0	0.100004	0.107899	0	0.107694	0	0.3214357	0	0	0	0	0	0	0.377943	0		
L 5.1A	Tank 1 A	Low T1	0	4.437328	0	0	0	0	0	0	5.260777	0	0	0.89959	2.001018	0	12.89039	0	0	0	0	0	0	0	0	0	0.22857	0	
L 5.2A	Tank 2 A	Low T2	0	7.71037	0	0	0	0	0	0	1.26602	0	0	11.54385	44.46067	0	7.833802	5.13952	0	0	0	0	0	0	0	0	0	0	
L 5.3A	Tank 3 A	Low T3	9.250336	0	0	0	0	0	0	0	20.61366	0	0	15.4665	3.123795	8.264882	52.79803	0	3.653334	0	0	0	0	0	0	0	0	0	
L 5F	Acid treatment	Low Acid treatment	0	1.88167	0	0	0	0	0	0	1.159272	9.071122	0	0	0	0	0	1.29742	0	1.26665	0	0	0	0	0	0	14.03897	26.66021	
L 6	Must	Low Must	2.868521	0.657579	0.42529	0	0.057276	0.72282	0	0	0.095771	0	0	0.111118	0.143043	0	0.32554	0	0	0	0	0	0	0	0	0	1.181089	0	
L 6.1B	Tank 1 B	Low T1	0	1.58894	0	0	0	0	0	0	0.457706	0	0	116.5769	58.93857	0	0.32554	0	0	1.743403	0	0	0	0	0	0	0	0.206484	
L 6.2B	Tank 2 B	Low T2	0	2.007823	0	0	0	0	0	0	0.345492	0	0	214.3526	117.92718	0	0.75808	0.948557	0	0	0	0	0	0	0	0	2.979442	3.627791	
L 6.3B	Tank 3 B	Low T3	0.253773	0.13689	0	0	0	0.24095	0	0	0.056022	0	0	0.308245	0.056269	0.060598	0.048557	0	0	0	0	0	0	0	0	0	0	0	
L 6F	Acid treatment	Low Acid treatment	0	0.1868	0	0	0	0	0	0	0.176747	0.763071	0	0	0.176747	0.763071	0	0.090755	0	0	0	0	0	0	0	0	0	0	0
N 1	Must	Normal Must	0	0	0	0	0.035165	0.174001	0	0	0.937823	0	0	0	0	0	0	0	1.351488	0	0	0	0	0	0	0	12.56619	0.591218	
N 1.1B	Tank 1 B	Normal T1	0	0.560984	0	0	0	0	0	0	0.4665102	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N 1.2B	Tank 2 B	Normal T2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N 1.3B	Tank 3 B	Normal T3	0	1.691914	0	0	0	0	0	0	1.931089	0	0	8.182169	0	0	5.78754	0	0	0	0	0	0	0	0	0	0	0	0
N 1F	Acid treatment	Normal Acid treatment	0	1.979476	0	0	0	0	0	0	12.7583	0	0	8.182169	0	0	5.78754	0	0	0	0	0	0	0	0	0	0	0	0
N 2	Must	Normal Must	0	1.66952	0	0	0.04156	0.09528	0	0	0.020289	0.837794	0	0	5.424549	0	1.98034	4.728195	0	4.71697	0	0	0	0	0	0	0	0	0
N 2.1B	Tank 1 B	Normal T1	0	0.52424	0	0	0	0	0	0	0	0	0	4.454382	1.258245	0	0.031748	0	1.63425	0	0	0	0	0	0	0	0	0	0
N 2.2B	Tank 2 B	Normal T2	0	0.365446	0	0	0	0	0	0	12.78994	0	0	4.14186	2.132426	0	2.162596	0	0	0	0	0	0	0	0	0	0	0	0
N 2.3B	Tank 3 B	Normal T3	0	0.386719	0	0	0	0	0	0	7.82461	9.933528	4.753593	11.18066	6.344444	0	5.841059	5.457074	0	4.04718	0	0	0	0	0	0	0	0	0
N 2F	Acid treatment	Normal Acid treatment	0	0	0	0	0	0.616773	2.779944	0	0	0	0	0.489008	0.069012	0	0.489334	3.472527	0	0.524959	0	0	0	0	0	0	0	0	0
N 3	Must	Normal Must	0	0	0	0	0.071913	0.425747	0.252586	0	0.20558	0	0	0	0	0	0.05315	0	0.6974556	0	0	0	0	0	0	0	0	0	0
N 3.1A	Tank 1 A	Normal T1	0	1.80402	0	0	0	0	0	0	0	0	0	0	0	0	0.380337	0	0	0	0	0	0	0	0	0	0	0	0
N 3.2A	Tank 2 A	Normal T2	0	0	0	0	2.439	0	0.622143	0	0	0	0	0	0	0	0.884324	0	0	0	0	0	0	0	0	0	0	0	0
N 3.3A	Tank 3 A	Normal T3	8.217319	6.083716	0	0	0	0	0.246728	0	0	0	0	0.366403	0.706404	0	4.828558	0.268483	0	0	0	0	0	0	0	0	0	0	0
N 3F	Acid treatment	Normal Acid treatment	0	20.86347	0	0	0	0	0	0	0	0	0	1.821822	0	4.295294	6.408023	0	6.2567	0	0	0	0	0	0	0	0	0	0
N 4	Must	Normal Must	0	7.486793	9.222286	0	64.8622	58.12348	15.34744	0	26.19463	0.690536	0	0	0.86798	5.19093	0	1.254441	0	3.064564	0	0	0	0	0	0	0	0	0
N 4.1A	Tank 1 A	Normal T1																											

Sample	Sampling site	Group	982	988	990	1016	1030	1031	1037	1045	1050	1067	1069	1071	1093	1110	1127	1149	1150	1154	1156	1160	1165	1176	1178	1186	1208	1271	1278
L.1	Low T1	Low T1	0.049373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.1.1	Tank 1 A	Low T1	2.50755	2.23678	1.566052	1.549999	0	2.87102	0.904587	0	2.986662	3.071043	24.42043	45.44711	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.1.2	Tank 1 B	Low T1	113.575	9.277206	0.062869	23.06524	0	1.748235	1.427725	0	1.855365	25.09138	36.4429	234.3007	22.1425	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.1.3	Tank 2 A	Low T2	0.482726	0	0	0.085932	0	1.696588	0	1.778222	0	14.503	13.38669	34.07417	38.13816	37.47367	0.149867	0	0	0	0	0	0	0	0	0	0	0	
L.1.4	Tank 2 B	Low T2	1.328275	0	0	0	0	12.64151	13.11962	0	1.7759	5.67945	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.1.5	Tank 3 A	Low T3	1.282621	0	0	0	0	1.273891	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.1.6	Tank 3 B	Low T3	5.967939	5.633239	5.683331	7.430571	5.390241	4.560939	1.586665	3.019937	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.1.F	Acid treatment	Low Acid treatment	4.802977	3.110362	0	0	1.714382	4.359757	3.75851	1.944246	3.936652	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2	Low T1	Low T1	0.76643	1.153054	0.022446	0	1.964024	0	1.951576	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.1	Tank 1 A	Low T1	3.157555	1.53054	0.864552	0	2.876038	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.2	Tank 1 B	Low T1	0.963735	1.951744	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.3	Tank 2 A	Low T2	0.963735	1.951744	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.4	Tank 2 B	Low T2	0.963735	1.951744	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.5	Tank 3 A	Low T3	0.963735	1.951744	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.6	Tank 3 B	Low T3	0.963735	1.951744	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.2.F	Acid treatment	Low Acid treatment	4.802977	3.110362	0	0	1.714382	4.359757	3.75851	1.944246	3.936652	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3	Low T1	Low T1	0.049373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3.1	Tank 1 A	Low T1	2.50755	2.23678	1.566052	1.549999	0	2.87102	0.904587	0	2.986662	3.071043	24.42043	45.44711	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3.2	Tank 1 B	Low T1	113.575	9.277206	0.062869	23.06524	0	1.748235	1.427725	0	1.855365	25.09138	36.4429	234.3007	22.1425	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3.3	Tank 2 A	Low T2	0.482726	0	0	0.085932	0	1.696588	0	1.778222	0	14.503	13.38669	34.07417	38.13816	37.47367	0.149867	0	0	0	0	0	0	0	0	0	0	0	
L.3.4	Tank 2 B	Low T2	1.328275	0	0	0	0	12.64151	13.11962	0	1.7759	5.67945	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3.5	Tank 3 A	Low T3	5.967939	5.633239	5.683331	7.430571	5.390241	4.560939	1.586665	3.019937	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3.6	Tank 3 B	Low T3	1.282621	0	0	0	0	1.273891	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.3.F	Acid treatment	Low Acid treatment	4.802977	3.110362	0	0	1.714382	4.359757	3.75851	1.944246	3.936652	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4	Low T1	Low T1	0.049373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4.1	Tank 1 A	Low T1	2.50755	2.23678	1.566052	1.549999	0	2.87102	0.904587	0	2.986662	3.071043	24.42043	45.44711	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4.2	Tank 1 B	Low T1	113.575	9.277206	0.062869	23.06524	0	1.748235	1.427725	0	1.855365	25.09138	36.4429	234.3007	22.1425	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4.3	Tank 2 A	Low T2	0.482726	0	0	0.085932	0	1.696588	0	1.778222	0	14.503	13.38669	34.07417	38.13816	37.47367	0.149867	0	0	0	0	0	0	0	0	0	0	0	
L.4.4	Tank 2 B	Low T2	1.328275	0	0	0	0	12.64151	13.11962	0	1.7759	5.67945	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4.5	Tank 3 A	Low T3	5.967939	5.633239	5.683331	7.430571	5.390241	4.560939	1.586665	3.019937	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4.6	Tank 3 B	Low T3	1.282621	0	0	0	0	1.273891	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.4.F	Acid treatment	Low Acid treatment	4.802977	3.110362	0	0	1.714382	4.359757	3.75851	1.944246	3.936652	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5	Low T1	Low T1	0.049373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5.1	Tank 1 A	Low T1	2.50755	2.23678	1.566052	1.549999	0	2.87102	0.904587	0	2.986662	3.071043	24.42043	45.44711	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5.2	Tank 1 B	Low T1	113.575	9.277206	0.062869	23.06524	0	1.748235	1.427725	0	1.855365	25.09138	36.4429	234.3007	22.1425	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5.3	Tank 2 A	Low T2	0.482726	0	0	0.085932	0	1.696588	0	1.778222	0	14.503	13.38669	34.07417	38.13816	37.47367	0.149867	0	0	0	0	0	0	0	0	0	0	0	
L.5.4	Tank 2 B	Low T2	1.328275	0	0	0	0	12.64151	13.11962	0	1.7759	5.67945	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5.5	Tank 3 A	Low T3	5.967939	5.633239	5.683331	7.430571	5.390241	4.560939	1.586665	3.019937	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5.6	Tank 3 B	Low T3	1.282621	0	0	0	0	1.273891	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.5.F	Acid treatment	Low Acid treatment	4.802977	3.110362	0	0	1.714382	4.359757	3.75851	1.944246	3.936652	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6	Low T1	Low T1	0.049373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6.1	Tank 1 A	Low T1	2.50755	2.23678	1.566052	1.549999	0	2.87102	0.904587	0	2.986662	3.071043	24.42043	45.44711	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6.2	Tank 1 B	Low T1	113.575	9.277206	0.062869	23.06524	0	1.748235	1.427725	0	1.855365	25.09138	36.4429	234.3007	22.1425	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6.3	Tank 2 A	Low T2	0.482726	0	0	0.085932	0	1.696588	0	1.778222	0	14.503	13.38669	34.07417	38.13816	37.47367	0.149867	0	0	0	0	0	0	0	0	0	0	0	
L.6.4	Tank 2 B	Low T2	1.328275	0	0	0	0	12.64151	13.11962	0	1.7759	5.67945	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6.5	Tank 3 A	Low T3	5.967939	5.633239	5.683331	7.430571	5.390241	4.560939	1.586665	3.019937	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6.6	Tank 3 B	Low T3	1.282621	0	0	0	0	1.273891	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.6.F	Acid treatment	Low Acid treatment	4.802977	3.110362	0	0	1.714382	4.359757	3.75851	1.944246	3.936652	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.7	Low T1	Low T1	0.049373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.7.1	Tank 1 A	Low T1	2.50755	2.23678	1.566052	1.549999	0	2.87102	0.904587	0	2.986662	3.071043	24.42043	45.44711	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.7.2	Tank 1 B	Low T1	113.575	9.277206	0.062869	23.06524	0	1.748235	1.427725	0	1.855365	25.09138	36.4429	234.3007	22.1425	0	0	0	0	0	0	0	0	0	0	0	0	0	
L.7.3	Tank 2 A	Low T2	0.482726	0	0	0.085932	0	1.696588	0	1.778222	0	14.503	13.38669	34.07417	38.13816	37.47367	0.149867	0	0	0	0	0	0	0	0	0	0	0	
L.7.4	Tank 2 B	Low T2	1.328275	0	0	0	0	12.64151	13.11962</																				

Sample	Sampling site	Group	Low	Mist	1602	1617	1626	1634	1636	1649	1653	1708	1749	1751	1759	1772	1810	1847	1857	1885	1891	1903	1916	1939
L.1	1.09517	2.651018	1.752359	0.279469	0	0	0	0	0	0	0	0	0	0.400241	0.157245	0.00627	0	1.837748	0	0	0	0	0	0
L.1.1	2.788661	2.271655	18.1265	28.40656	0	7.93456	13.92173	10.13448	0	1.171446	0	10.75101	5.583631	0	4.400386	0.208435	0	0	14.97254	9.455467	10.07119	15.3186	10.60731	0
L.1.1B	2.77961	0	6.220498	0	0	0	1.667132	0	0	0	0	0	0	0	0	0.774496	0	0	0	0	0	0	0	0
L.1.2	7.980064	11.47422	10.8633	0	0	0.4177258	13.95163	13.22512	12.93019	15.96088	10.83993	6.088526	0	0	0	9.931486	0	0	8.556364	9.640036	19.54264	15.50117	0	0
L.1.2B	7.980064	5.085754	8.352117	10.255	0	0	0.098658	6.850564	0	0	0	0	0	0	0	6.198754	0	0	0	0	0	0	0	0
L.1.3	4.947182	6.932489	4.697102	6.700724	0	0	2.261417	1.60822	0	2.054228	0	0	0	0	0	0.054068	0	0	0	0	0	0	0	0
L.1.3B	19.42458	4.147822	24.90223	7.380674	0	18.50085	9.402108	10.94393	13.43474	11.86813	8.52851	7.038327	0	0	0	10.34767	10.02208	0	0	0	0	0	0	0
L.1.3B	4.947182	1.908538	4.271968	6.616089	0	6.400312	10.30126	7.01831	4.75119	5.69902	7.36784	6.890162	0	0	0	10.18757	15.11084	15.60253	0	0	0	0	0	0
L.1.4	0.211133	0	0	0	0	22.04402	2.366575	5.78053	9.730266	0	5.69234	4.029146	0	0.208311	0.034836	0.018527	0	0	5.0432	8.89049	7.693448	10.57663	10.27276	0
L.1.4B	8.145199	3.82729	9.952122	0	0	7.274043	6.786701	8.50626	10.36432	10.79281	7.993212	6.177453	0	0	0	2.79013	6.177453	0	0	2.840761	6.15183	7.786613	6.62899	0
L.1.4B	6.373478	6.855242	5.025273	10.62313	0	3.747636	6.4297	7.728919	13.50618	13.50966	6.94566	7.01312	0	0	0	8.010314	13.85167	7.838628	8.97944	10.71815	7.838628	12.48112	0	0
L.2	8.002149	5.849124	1.28275	10.00959	0	0	0.0861	0	0	0	0	0	0	0	0	0.99861	0	0	0	0	0	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.543431	3.983613	0	3.780219	4.654494	9.92146	14.53818	15.55464	5.42058	2.885556	0	0	0	0.123169	0.931877	0	0	1.792244	4.591507	0	0	0
L.2B	8.471109	2.85126	5.																					

to be continued

Sample	Sampling site	Group	1947	1948	1970	1982	2003	2014	2022	2026	2039	2042	2061	2064	2067	2068	2069	2074	2079	2090	2115	2116	2131	2132	2137	2145	2147	2177	
L.1	Must	Low Must	0	0	0	55.84076	261.1646	7.235251	0	24.40043	0	0	0	0	0	133.5593	0	89.5981	18.6727	0	0	0	4.478982	0	0	0	102.9687	0	
L.1.1A	Tank 1.A	Low T1	0	0	0	12.36795	6.913975	1.124924	0	13.77947	10.77842	10.54073	0	15.57786	0	10.12024	12.20544	12.10765	13.20544	24.80612	0	0	0	0	0	18.60764	0	5.50373	
L.1.1B	Tank 1.B	Low T1	0	0	0	0.796095	0.503981	1.490956	0	0.154077	0.296687	1.761527	0	0	0	0.92781	3.37743	0.939539	0.852448	1.317977	0	2.867045	0	0	0	0	2.60426	0	
L.1.2A	Tank 2.A	Low T2	0	0	0	6.888915	5.306596	9.936353	0	17.64814	12.23142	13.40923	0	11.4381	0	7.661095	18.36524	7.73827	8.271893	6.601073	0	9.872258	0	0	0	16.06774	10.31408	0	
L.1.2B	Tank 2.B	Low T2	0	0	0	17.70595	3.462107	5.423113	0	1.858319	9.99963	9.326477	0	4.951869	0	0	0	0	0	0	0	0	0	0	0	0	4.19875	0	
L.1.3A	Tank 3.A	Low T3	0	0	0	4.092651	2.711847	6.21571	0	0.407106	0.787882	2.35454	0	0	0	2.28884	4.47844	1.359716	0.824832	1.114193	7.241675	0	0	0	0	0	21.41683	0	
L.1.3B	Tank 3.B	Low T3	0	0	0	5.050241	2.711847	6.21571	0	0.407106	0.787882	2.35454	0	0	0	2.28884	4.47844	1.359716	0.824832	1.114193	7.241675	0	0	0	0	0	21.41683	0	
L.1.F	Acid treatment	Low Must	4.666262	0	0	62.72613	0	0	0	9.36702	7.5017	6.527944	0	4.168333	0	5.250412	0	5.18535	5.715655	4.775107	0	0	0	0	0	8.582107	0	9.72045	
L.2	Must	Low Must	0	0	0	88.42306	77.36697	0.72743	0	0.375502	12.63279	0	0	0	0	201.9858	0	88.3789	5.493773	0	0	0	0	0	0	0	96.03877	0	
L.2.1A	Tank 1.A	Low T1	22.73542	0	0	0.709739	6.006067	2.09475	7.665637	0.905256	5.711791	5.79792	0	2.348012	0	5.051653	0	6.14207	6.198166	3.61393	0	4.510311	3.048196	0	0	0	8.061284	0	
L.2.1B	Tank 1.B	Low T1	22.71743	0	0	41.74423	7.988239	10.12633	0	12.34925	10.14449	10.31241	0	5.934845	0	8.315262	0	8.602935	9.570882	4.156983	21.6159	8.854385	0	0	0	0	2.21052	0	
L.2.2A	Tank 2.A	Low T2	17.69046	0	0	9.793141	4.448683	10.15548	0	9.78273	2.233685	10.3716	0	7.025169	0	7.90809	22.47395	7.329868	7.910969	7.980614	8.573881	0	0	0	0	6.239958	0	4.20851	
L.2.2B	Tank 2.B	Low T2	4.951569	0	0	12.96907	7.35506	0	0	3.140057	6.711947	10.419	0	12.87368	0	6.638033	26.5146	7.48453	6.771101	6.751204	0	7.943684	9.90884	26.47444	0	0	0	5.22009	
L.2.3A	Tank 3.A	Low T3	24.56297	0	0	4.813745	6.138753	9.650377	0	14.73971	12.38823	11.72755	0	5.881661	0	6.04537	8.21567	8.04517	8.21567	8.04537	24.85915	5.522955	0	0	0	12.53755	0	0	
L.2.3B	Tank 3.B	Low T3	21.3308	0	0	5.826069	0	6.829235	0	14.73971	12.38823	11.72755	0	7.508869	0	6.04537	8.21567	8.04517	8.21567	8.04537	24.85915	5.522955	0	0	0	12.53755	0	0	
L.2.F	Acid treatment	Low Must	8.094531	5.8946	0	4.7842	26.02759	30.307	1.058452	0.044827	2.723494	0	0	3.35281	0	2.465731	8.50489	6.927561	10.0634	8.7466	0.012407	9.148179	0	0	0	7.561053	0	0	
L.3	Must	Low Must	2.048388	0	0	8.39023	2.96545	1.127847	0	0.044827	2.723494	0	0	0.80344	0	2.233464	0	0.945343	10.01879	2.04168	8.44464	0.098227	2.041891	0.908889	0	0	215.8538	0	
L.3.1B	Tank 1.B	Low T1	6.5218	0	0	22.95154	17.50887	0	0	4.614355	4.638541	9.585277	0	15.42848	0	6.972066	13.96533	8.563237	8.525004	2.293552	0	5.268242	19.91243	24.02607	0	0	2.469295	0	
L.3.2B	Tank 2.B	Low T2	13.42551	0	0	9.944209	9.22564	11.56012	0	12.59132	0	0	0	10.79394	0	11.20072	11.91616	14.6019	14.6019	25.92559	0	13.04733	0	0	0	4.44305	0	0	
L.3.F	Acid treatment	Low Must	7.171303	0	0	5.434004	0	0	0	0.492094	3.802907	5.11517	0	5.918548	0	4.869472	0	5.231317	5.940212	2.655404	0	14.14438	0	0	0	6.553372	0	7.045407	
L.4	Must	Low Must	0	0	0	119.6903	361.0939	0	0	0.492094	3.802907	5.11517	0	0	0	268.7626	0	0.80021	0.242904	1.870955	3.519716	0	0	0	0	172.825	0	0	
L.4.1A	Tank 1.A	Low T1	4.503886	0	0	11.24156	6.905575	0	0	5.02672	2.447957	2.43657	0	0	0	2.04003	2.814566	0.40021	0.242904	1.870955	3.519716	0	0	0	0	0	7.792536	0	0
L.4.2A	Tank 2.A	Low T2	5.016449	0	0	7.239755	2.125573	2.417788	0	4.032166	2.049757	2.951762	0	0	0	0.752004	0.248474	0	0.237731	7.445244	4.38884	3.050715	0	0	0	0	0	0	0
L.4.3A	Tank 3.A	Low T3	2.7905	0	0	2.700887	1.39355	1.50603	0	1.950566	1.134092	1.990263	0	0	0	0.7308318	0.850237	0.356475	0.130568	2.278955	0.67554	0.081705	2.435695	0	0	1.93057	0	0	
L.4.F	Acid treatment	Low Must	0.395197	2.838056	0	4.780705	0	1.780833	0	0.268657	0.279584	1.326193	0	0	0	0.97605	0	0.32152	0.154284	1.900675	0	4.160015	0	0	0	0.953153	0	0	
L.5	Must	Low Must	0	0	0	0	0	0	0	0.790381	36.95911	0	0	0	0	0	0	73.23029	8.100935	0	1.747771	0	0	0	0	0	285.1548	0	0
L.5.1A	Tank 1.A	Low T1	3.882173	0	0	16.94343	4.771901	3.828679	3.248543	0.429157	1.875008	3.83812	0	2.131313	0	1.98289	8.683384	3.294936	1.72157	3.46789	0.41224	0.747823	1.70354	4.988317	0	0	5.11654	0	
L.5.2A	Tank 2.A	Low T2	8.967671	0	0	11.02687	9.221359	0	0	6.949291	4.476555	4.768459	0	0	0	3.427242	2.23212	3.132093	3.584881	4.02316	2.228814	0.424517	5.504506	0	0	13.16574	0.501513	0	
L.5.3A	Tank 3.A	Low T3	0.542679	0	0	4.046361	0.571141	0	0	0.188315	0.266183	2.488676	0	0	0	3.373019	0	1.033961	0.546159	8.584476	3.915415	0.501568	6.593555	0	0	5.95881	0	0	
L.5.F	Acid treatment	Low Must	5.580924	5.095128	0	4.568791	0	0	0	0.188315	0.266183	2.488676	0	0	0	2.757881	0	0.896866	0.419715	4.2691	0	1.570594	0	0	0	1.77717	0	0	
L.6	Must	Low Must	0	0	0	64.77447	102.575	1.42879	0	1.07473	1.07702	3.014144	0	0	0	34.00447	0	0.934574	0.700747	2.365339	3.186615	5.46147	6.518543	2.27829	0	0	1.67488	0	0
L.6.1B	Tank 1.B	Low T1	3.583019	0	0	20.34719	14.05168	2.759359	1.51904	1.07473	1.07702	3.014144	0	0	0	1.93259	1.728008	0.9415674	0.700747	2.365339	3.186615	5.46147	6.518543	2.27829	0	0	1.67488	0	0
L.6.2B	Tank 2.B	Low T2	1.724179	0	0	2.739761	1.854792	1.259395	2.760652	0	0	0.140031	0.933921	0	0	0.284784	0	0.371488	0.215588	1.577895	0.325316	3.622727	2.27829	1.160593	0	0	0	0	0
L.6.3B	Tank 3.B	Low T3	0	0	0	6.97758	36.06839	59.7675	0	0	0	29.01884	0	0	0	1.435447	77.38573	0	0	0	0	445.5888	0	0	0	0	0	0	0
L.6.F	Acid treatment	Low Must	0.057158	0.034726	0	8.784826	0	0	0	4.664969	30.57709	28.17299	0	0	0	206.8842	0	74.88951	0	67.73651	0	0	0	0	0	0	42.2822	0	0
N.1	Must	Normal Must	0	0	0	0.382335	0.63296	1.9621	0	1.2613	36.11993	0	0	0	0	311.5588	0	50.36483	6.87616	132.4879	0	0	0	0	0	0	0	0	0
N.1.1B	Tank 1.B	Normal T1	4.45373	26.99138	0	2.979565	0	0	0	0.188315	0.266183	2.488676	0	0	0	3.373019	0	1.033961	0.546159	8.584476	3.915415	0.501568	6.593555	0	0	5.95881	0	0	0
N.1.2B	Tank 2.B	Normal T2	3.814373	2.53382	0	2.354661	0	0	0	0.188315	0.266183	2.488676	0	0	0	3.373019	0	1.033961	0.546159	8.584476	3.915415	0.501568	6.593555	0	0	5.95881	0	0	0
N.1.3B	Tank 3.B	Normal T3	4.081835	0	0	3.405949	0.9767	5.629482	0	12.47311	8.77207	9.65028	0	0	0	6.870779	0	8.04729	8.796348	6.011188	0	305.1693	0	0	0	5.793462	0	0	0
N.1.F	Acid treatment	Normal Must	10.31383	6.680849	0	34.05268	139.9618	143.8278	0	5.843418	10.49173	8.057033	0	0	0	14.33344	0	7.466361	5.635555	1.118899	0	0	0	0	0	5.52775	0	0	0
N.2	Must	Normal Must	0	0	0	41.34413	176.7831	0.72467	0	0.292975	19.46259	0	0	0	0	206.8842	0	5.49516	6.16109	5.80914	0	130.1615	0.512021	0	0	0	38.9194	0	0
N.2.1B	Tank 1.B	Normal T1	3.457112	0	0	2.502269	5.297147	1.92291	2.851971	0	4.314111	4.361356	5.748677	0	0	0	6.057778	11.16998	7.29786	0	6.077008	0	0	0	0	0	1.000844	2.931596	0
N.2.2B	Tank 2.B	Normal T2	4.599042	3.438451	3.538157	4.962666	0	4.707721	0	12.5047	5.78813	6.432383	0	7.36357	0	6.838013	0	6.78005	7.966059	5.62028	0	11.10766	0	0	0	4.689414	0	5.131005	0
N.2.F	Acid treatment	Normal Must																											

UNTARGETED INTRACELLULAR METABOLIC PROFILE OF FERMENTATIVE TANKS FROM A BRAZILIAN ETHANOL DISTILLERY UNDER DISTINCT ETHANOL YIELDS

Abstract

Brazil is the second-largest ethanol producer in the world, and the main route for this production is the fermentation of sugarcane by selected strains of *Saccharomyces cerevisiae*. However, several factors can directly impact fermentative performance and lead to financial losses for the ethanol industry. Aiming to obtain deeper knowledge about metabolic changes during different fermentative yield, exploratory untargeted metabolomics analysis of fermentative tanks was performed. We observed differences in the intracellular profile of the microorganism community, which suggests that differences in metabolic routes are being used in each fermentative situation. A greater abundance of sugars metabolites such as Galactose, Glucose, Glucose-6-phosphate, Lactulose, Maltotriose, Melibiose and Sorbose in cells under normal fermentative yields was detected in normal productivity situation, while a greater abundance of compounds related to the presence of bacterial contaminants and yeast stress response, such as 2-Furoic acid, Glucosaminic acid, Mannitol, Stearic acid and Trehalose, respectively, were detected in the metabolic profile of samples belonging to low fermentative yield group. The study of the fermentation metabolome can help to understand the metabolic alterations on yeast and contaminants that lead to a decrease in ethanol yield. To our knowledge, this was the first exploratory analysis of the metabolic profile of fermentation tanks conducted in a Brazilian ethanol distillery.

Keywords: Ethanol fermentation, GC-MS, intracellular metabolome

5.1 Introduction

Ethanol fermentation is the main industrial route for ethanol production worldwide. In Brazil, this process is usually driven by *Saccharomyces cerevisiae* cells that ferment sugarcane derivatives. Fermentation can be defined as an anaerobic process where sugar molecules are converted into energy, releasing molecules of ethanol and carbon dioxide (LOPES et al., 2016; NELSON; COX, 2018). During the fermentative process, yeast cells encounter various stressing conditions such as osmotic stress, high temperature, ethanol stress, pH, the presence of inhibitory compounds and microbial contaminants. It is known that the usage of selected yeast strains in Brazilian distilleries was fundamental for an increase in the ethanol yield, since these strains may have superior tolerance for stressful conditions (BASSO et al., 2008). However, even selected strains suffer from environmental changes that impact their cellular metabolism (BREXÓ; SANT'ANA, 2017).

As reviewed by Lopes et al (2016), yeast cells respond to stressful conditions through changes in their metabolism to produce protecting metabolites like trehalose, glycogen, and succinic acid. These metabolic changes have a direct impact on ethanol production. The study of these metabolic changes which reflect the interactions between genetic information and environmental changes can be defined as metabolomics (ZHANG et al., 2020). Metabolomics studies provide information about the metabolic profile of a biological matrix usually through a targeted or untargeted approach (JOHNSON; IVANISEVIC; SIUZDAK, 2016). Nowadays metabolites detection on metabolomics studies are commonly based on magnetic resonance (NMR) or mass spectrometry (MS) (ZHANG et al., 2020). MS is usually coupled with a previous separation by Gas or Liquid chromatography.

In targeted metabolomics studies, a specific set of compounds are analyzed, based on predefined selection. On the other hand, untargeted metabolomics approaches aim to analyze all detectable metabolites in a sample (JOHNSON; IVANISEVIC; SIUZDAK, 2016; GERTSMAN; BARSHOP, 2019). Due to their characteristics, untargeted approaches are frequently used as an exploratory tool and are mainly based on the differential metabolites

comparison between groups (GERTSMAN; BARSHOP, 2019). However, the results of the qualitative exploratory analysis cannot provide information on intracellular metabolic rates and pathways (JOHNSON; IVANISEVIC; SIUZDAK, 2016).

Aiming to explore the intracellular metabolic differences in the fermentation tanks of an ethanol distillery under different fermentative yields, an untargeted metabolomics approach was conducted. Sampling was performed in a Brazilian distillery, operating in a continuous fermentation line between August and November 2017. Were analyzed six samples belonging to days with low fermentative yield and six samples from normal fermentative yield days. The intracellular metabolites of all the operating fermentative tanks were submitted to GC-MS analysis.

5.2 Methods

5.2.1 Sampling

Sampling was performed between August and November 2017, in a continuous ethanol distillery operating with 3 fermentative tanks, in the southeast region of Brazil, Piracicaba - SP (22° 43' 30" S, 47° 38' 51" W). Samples from all the fermentative tanks were collected during six days of Low ethanol fermentative yield and six days of Normal fermentative yield. Samples were collected from the bottom tap of each fermentation tank. We followed the leakage-free cold methanol quenching, as described by Canelas et al (2008), although with several modifications. Specifically, samples were transferred to tubes filled with methanol that had been pre-cooled to -40 °C, at a ratio of 1: 5 (sample: methanol). From the sampling site to the laboratory samples were maintained at -40 °C.

5.2.2 Sample preparation for GC-MS analysis

Upon arriving in the laboratory samples were centrifuged at 4000g (Allegra 64R - Beckman centrifuge) for ten minutes at -9°C. After the centrifugation step, the supernatant was discarded and the pellets were stored at -80°C prior to further analysis.

In this study, we privileged the polar fractions of extracted metabolites. Intracellular metabolites were analyzed following Smart et al. (2010) with minor modifications in the sample (250mg) and cold methanol-water solution (750µL) quantity and centrifugation steps. Samples were centrifuged (Centrifuge 5415R - Eppendorf) at 16,000 g at 0°C for 20 min. Also, 10 µL of the isotopically labeled succinic acid (D4, 98% - DLM 584-5), myristic acid (1, 2, 3 - 13C3, 99% - CLM 3665- 0.5) and palmitic acid (1,2,3,4 - 13C4), prepared at a concentration of 1 mg/mL, was used as an internal standard. After extractions, samples were filtered through a 0.22 µm PVDF filter and stored at -80°C.

Before the derivatization procedure, 200µL of each sample was lyophilized. Derivatization and Methoxyamination steps were conducted accordingly to Budzinski et al (2019).

5.2.3 GC- MS data analyses

Samples were analyzed following Budzinski et al. (2019), alongside with extraction blanks and a series of n-alkanes (C12–C40), allowing the retention index calculation as proposed by Jonsson et al. (2005). One microliter of each derivatized sample was injected in the splitless mode in a 7890A gas chromatograph (Agilent Technologies, Santa Clara, USA), coupled to an automatic Comb-Xt processor (Leap Technologies, Carrboro, USA). The injection temperature was 280 °C and the septum purge flow rate was 20 mL min⁻¹ lasting for 60 seconds. The Helium flow rate through the column was 1 mL min⁻¹. The column temperature was maintained at 80°C for two minutes and then increased by 15 °C every minute until reaching a temperature of 305°C, which was held for another 10 minutes. The column effluent was inserted into an ionization source of gas chromatography with time-of-flight mass spectrometry.

(Pegasus 4D, Leco Corp., St. Joseph, USA) composed by two fused silica columns - first dimension column (Agilent DB- 5) 20 m long (0.18 mm inner diameter x 0.18 µm film); second dimension column 0.96 m (RXT-170.10 mm inner diameter x 0.10 µm film). The temperature of the transfer line and ion source were 280 and 250 °C, respectively. An electron beam of 70 eV generated ions in an ionization current of 2.0 mA, and 10 spectra s⁻¹ were recorded in the mass range 45-800 m/z. The ChromaTOF v.4.51 (Leco Corp., St. Joseph, USA) software was used to perform baseline correction and export all MS files in NetCDF format.

The R package Target Search (CUADROS-INOSTROZA et al., 2009; R, 2013) was used for peak detection, retention time alignment and library matching. For metabolites identification, their retention indexes and spectra similarity (score greater than 600) were compared with compounds of the Golm Metabolome Database (KOPKA et al., 2005). After this step samples were grouped according to their collection site and productivity. For a metabolite to be considered valid, it should present at least three mass fragments at the correct RI, and present in at least 50% of samples from one group. For the following analyses, metabolite intensities were normalized using total ion chromatogram (TIC), and technical replicates were merged by average. The resulting matrix (Table S1) was Pareto scaled and log-transformed prior to statistical analysis. Statistical multivariate analysis (Principal Component Analysis - PCA) and univariate (Student t-test, False discovery rate - FDR - adjusted p ≤ 0.05) analysis were done in MetaboAnalyst 4.0 software (XIA et al., 2009; CHONG et al., 2018).

5.2.4 Fermentative Yield parameters

The ethanol fermentation distillery where the sampling was performed, works with a daily ethanol production target of approximately 600 m³ per day. Due to the constant input of feedstock and output of the fermented wine in the continuous fermentation line, the processing monitoring is temporal. Thus, a punctual fermentative yield estimative is needed to detect changes and guide industry decision-making. The ratio between the amount of ethanol produced and the amount of expected ethanol produced is then utilized.

The calculation of the expected amount of ethanol is expressed by the following formula:

$$Exp\ Ethanol = \frac{\left\{ \left(\frac{mustTRS}{100} \right) * 0.6475 * 0.89 * mustflowrate + \left(\frac{eth_{yeast}}{100} \right) * yeastflowrate \right\}}{(mustflowrate + yeastflowrate)}$$

It takes into account the total reducing sugars present in must (mustTRS); the stoichiometric sugar to ethanol conversion factor 0.6475 - which equates to the production of 0.6475 liters of 100% ethanol for each kg of total reducing sugars; a theoretical efficiency of fermentation at 89%; the must flow rate; the percentage of ethanol detected in yeast cells in the acid treatment and the flow rate of yeast cells in the first fermentative tank.

In normal conditions, the value of this ratio should be close to 1. Measurements for ratio calculations are usually taken three times a day. Considering that small measurement deviation is expected, during this sampling period daily ratio values up to 0.93 were considered as normal fermentative yield by the distillery.

To facilitate understanding, in this study we named "Normal productivity" ratio values up to 0.93 and "Low productivity" below 0.93. It is important to highlight that ratio values were provided by the ethanol distillery where samples were collected.

5.2 Results

After data processing, a total of 114 intracellular metabolites were reported (Table S1) of which 37% were unidentified. Unidentified metabolites are molecules whose detection and quantification are reproducible but still lacking a chemical identity assignment (KRUMSIEK et al., 2012). Thus, considering that this is an exploratory analysis and that these compounds can play an important role in the metabolic profile, unidentified metabolites were kept in the profile analyzes.

After verifying that a PCA considering all tank samples did not show differences between Normal and Low productivity groups (figure S1), we evaluated each sampling site separately. In the visualization of the metabolic profile of Tank 1, 2 and 3 at different productivity situations, it was possible to observe the separation between Normal and Low productivity groups (Figure 1A-C). This suggests that there were differences in the metabolic profile in all the fermentative tanks.

Figure 1- PCA of intracellular metabolic profile in low and normal productivity. PCA plot of Tank 1 (A) , Tank 2 (B) and Tank 3 (C). Low and Normal productivity samples are represented by red and green dots, respectively.

To check the differentially abundant metabolites in Low and Normal productivity situations, a Fold Change (FC) analysis followed by a Student's t-test was applied. It was considered the minimum fold change threshold 2 (Normal / Low) and t-tests threshold 0.05 (FDR corrected). Log₂(FC) negative values correspond to metabolites more abundant in Low productivity situations (Table 1-3).

Table 1: Metabolites differentially abundant in the intracellular metabolic profile of Tank 1 from Low and Normal productivity groups. Fold change threshold 2 (Normal/Low); t-tests threshold 0.05; FDR corrected

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
1626	Stearic acid	0,00057698	-10,759	0,0055014	2,2595
2560	NA	0,00060003	-10,703	0,022434	1,6491
1394	NA	0,00074259	-10,395	0,0062353	2,2051
1160	NA	0,00074854	-10,384	0,022434	1,6491
70	2-Furoic acid	0,0010524	-9,8921	0,022434	1,6491
1385	Glucosaminic acid	0,0012806	-9,6089	0,024794	1,6057
377	NA	0,0016415	-9,2507	0,025656	1,5908
37	Valine	245,67	7,9406	0,0068923	2,1616
2543	Maltotriose	314,85	8,2985	0,025656	1,5908
1177	Glucose	609,96	9,2526	0,0064658	2,1894
1662	NA	640,97	9,3241	0,024794	1,6057
846	D-arabinono-1,4-lactone	702,86	9,4571	0,0062353	2,2051
1069	Sorbose	1189,9	10,217	0,025172	1,5991
2090	n-Octacosane	2850,8	11,477	0,024794	1,6057
992	Galactose	4396,5	12,102	0,032227	1,4918
818	Methyl jasmonate	6177,9	12,593	0,0055014	2,2595

Table 2: Metabolites differentially abundant in the intracellular metabolic profile of Tank 2 from Low and Normal productivity groups. Fold change threshold 2 (Normal/Low); t-tests threshold 0.05; FDR corrected

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
837	N-(2-Hydroxyethyl)iminodiacetic acid	0,00144	-9,4393	0,005175	2,2861
1071	NA	0,001766	-9,1452	0,004673	2,3304
130	Glyceraldehyde	0,003244	-8,2682	0,003477	2,4588
1208	Mannitol	0,004192	-7,8982	0,002413	2,6175
2061	Androstendione	0,00572	-7,4497	0,002372	2,6248
2560	NA	0,006142	-7,3471	0,045255	1,3443
37	Valine	0,006986	-7,1613	0,002194	2,6589
1697	NA	0,007398	-7,0786	0,002087	2,6805
106	3-Methyl-2-oxopentanoic-acid	0,008493	-6,8795	0,002548	2,5938
1634	NA	0,00925	-6,7563	0,002404	2,6191
1387	1-Methylhistidine	0,009373	-6,7373	0,002466	2,608
985	Homogentisic acid	0,0098	-6,673	0,013227	1,8785
1650	N-acetyl-Muramic acid	0,010929	-6,5158	0,013871	1,8579
1964	NA	0,011119	-6,4908	0,017164	1,7654
749	Cis-1,2-O-benzylidene-D-glycerol	0,011856	-6,3982	0,016134	1,7923
2519	NA	0,012204	-6,3565	0,012774	1,8937
2011	Trehalose	0,01591	-5,9739	0,013612	1,8661
1932	Spermine	0,01912	-5,7088	0,014204	1,8476
1323	Gluconic acid	0,021506	-5,5391	0,043623	1,3603
1626	Stearic acid	0,42326	-1,2404	0,048469	1,3145
2019	alpha-D-Galactopyranosyl-(1,4)-D-galactopyranoside	74,166	6,2127	0,003845	2,4151
2187	Melibiose	91,734	6,5194	1,98E-05	4,7039
2064	Androstendione	92,361	6,5292	8,52E-05	4,0697
499	Iminodiacetic acid	96,576	6,5936	0,02283	1,6415
1903	Lactulose	98,476	6,6217	0,000358	3,4465
691	NA	109,03	6,7686	0,003044	2,5166
1700	Glucose-6-phosphate	121,55	6,9255	0,00309	2,51
1236	Glucaric acid-1,4-lactone	123,55	6,949	0,000322	3,4919
750	NA	133,84	7,0644	0,004068	2,3907
839	Tricarballic acid	681,52	9,4126	0,000247	3,6074

Table 3: Metabolites differentially abundant in the intracellular metabolic profile of Tank from Low and Normal productivity groups. Fold change threshold 2 (Normal/Low); t-tests threshold 0.05; FDR corrected

ID	Name	FC	log2(FC)	p.adjusted	log10(p)
358	2,4-Dihydroxybutanoic acid	2,15E-05	-15,505	2,45E-07	6,6107
2559	NA	2,42E-05	-15,338	0,000913	3,0393
1653	NA	3,57E-05	-14,774	2,72E-11	10,565
1840	N-tetradecanoyl-Homoserine lactone	3,59E-05	-14,765	0,000913	3,0393
606	NA	4,1E-05	-14,575	0,001308	2,8833
1208	Mannitol	4,24E-05	-14,526	0,000913	3,0393
106	3-Methyl-2-oxopentanoic-acid	6,1E-05	-14,002	0,000913	3,0393
1615	NA	6,37E-05	-13,938	4,52E-11	10,345
2011	Trehalose	8,23E-05	-13,569	0,048475	1,3145
192	Tropinone	8,79E-05	-13,474	0,014442	1,8404
1662	NA	9,01E-05	-13,439	0,048475	1,3145
1387	1-Methylhistidine	0,000101	-13,276	0,048475	1,3145
2075	Sophorose	0,00011	-13,154	0,048475	1,3145
1236	Glucaric acid-1,4-lactone	0,000117	-13,062	0,000913	3,0393
1105	3-Hydroxytetradecanoicacid	0,000232	-12,076	0,048475	1,3145
2056	NA	9885	13,271	0,014442	1,8404
698	NA	11000	13,425	0,014442	1,8404
1442	NA	14574	13,831	0,014442	1,8404
985	Homogentisic acid	15264	13,898	0,024269	1,6149
2019	alpha-D-Galactopyranosyl-(1,4)-D-galactopyranoside	15922	13,959	0,014442	1,8404
1160	NA	17582	14,102	0,000913	3,0393
1903	Lactulose	19843	14,276	0,014442	1,8404
1700	Glucose-6-phosphate	24430	14,576	0,001172	2,931
99	Succinic acid methylester	24543	14,583	0,014442	1,8404
841	NA	34402	15,07	2,45E-07	6,6107
517	NA	36255	15,146	0,014442	1,8404
373	Aspartic acid	38641	15,238	0,014442	1,8404
818	Methyl jasmonate	45759	15,482	0,001172	2,931

In the Tank 1 the metabolites Stearic acid, 2-Furoic acid, Glucosaminic acid and unidentified 2560, 1394, 1160 and 377 were detected in Low productivity group with a higher abundance. On the other hand, the metabolites Valine , Maltotriose, Glucose, D-arabinono-1,4-lactone, Sorbose, n-Octacosane, Galactose, Methyl jasmonate and unidentified 1662 were reported as more abundant in the Normal productivity group (Table 1).

In the Tank 2 the metabolites detected as more abundant in Low productivity group were N-(2-Hydroxyethyl)iminodiacetic acid, Glyceraldehyde, Mannitol, Androstendione, Valine, 3-Methyl-2-oxopentanoic-acid, 1-Methylhistidine, Homogentisic acid, N-acetyl-Muramic acid, Cis-1,2-O-benzylidene-D-glycerol, Trehalose, Spermine, Gluconic acid and unidentified 1071, 2560, 1697, 1964 and 2519. The metabolites alpha-D-Galactopyranosyl-(1,4)-D-galactopyranoside, Melibiose, Androstendione, Iminodiacetic acid, Lactulose, Glucose-6-phosphate, Glucaric acid-1,4-lactone, Tricarballic acid and unidentified 691 and 750 were detected with a higher abundance in the Normal group (Table 2). It is possible to notice that the metabolite Androstendione was identified as more abundant in both productivity situations. In each group, this metabolite was detected with a different library ID (2060 and 2061), which

is not commonly observed. Accordingly to GOLM database, the difference between both metabolites is the carbonyl group induced by the methoximation reaction. In any case, this metabolite was dismissed from our analysis.

In the Low productivity group of Tank, the reported metabolites with a higher abundance were 3,4-Dihydroxybutanoic acid, N-tetradecanoyl-Homoserine lactone, Mannitol, 3-Methyl-2-oxopentanoic-acid, Trehalose, Tropinone, 1-Methylhistidine, Sophorose, Glucaric acid-1,4-lactone, 3-Hydroxytetradecanoic acid and unidentified 2559, 1653, 606, 1615 and 1662. In turn in the Normal productivity group Homogentisic acid, alpha-D-Galactopyranosyl-(1,4)-D, galactopyranoside, Lactulose, Glucose-6-phosphate, Succinic acid methyl ester, Aspartic acid, Methyl jasmonate and unidentified 2056, 698, 1442, 1160, 841 and 517 were detected as more abundant (Table 3).

Although unidentified metabolites play an important role in profile analysis, we cannot discuss them and therefore will be disregarded in the following analysis. Comparing the differentially abundant metabolites reported in all the fermentation tanks (Table 1-3), we can notice that some metabolites showed a different pattern in different fermentative tanks. The metabolite Valine was reported as more abundant in the Tank 1 of Normal productivity group while in the Tank 2 was reported as more abundant in Low productivity group. On the other hand, the metabolites Glucaric acid-1,4-lactone and Homogentisic acid were reported as more abundant in the Low productivity group in Tank 2 and on the Normal group in tank 3. Taking into account that the distillery where the sampling was performed operates in a continuous fermentation line, these differences may be related to specific environmental conditions of each tank.

The metabolite Methyl jasmonate was reported as more abundant in the Normal productivity group in Tank 1 and 3. In the Tank 2 and 3, the metabolites 3-Methyl-2-oxopentanoic-acid, Trehalose, 1-Methylhistidine and Mannitol were more abundant in Low productivity group, while Lactulose, Glucose-6-phosphate, and alpha-D-Galactopyranosyl-(1,4)-D-galactopyranoside were more abundant in the metabolic intracellular profile of Normal productivity group.

5.3 Discussion

The main route of ethanol production worldwide is the microbial fermentation. In Brazil, fermentation is carried out by selected yeast strains that convert sugars from sugarcane into ethanol (AMORIM et al., 2011). It is estimated that sugarcane production is responsible for more than half of ethanol's final cost, so any reduction in the fermentative yield represents a waste of feedstock and a significant financial loss (DELLA-BIANCA et al., 2013). Considering that the ethanol daily production of the distillery where sampling was performed is near to 600m³, a decrease of 7% in the target production means a loss of approximately 42,000 liters of ethanol per day. Environmental conditions, the feedstock composition, the stress caused by frequent cell recycling, and the presence of contaminants are known to affect the fermentative performance of yeast strains and also impact ethanol productivity (BREXÓ; SANT'ANA, 2017). To explore changes in the intracellular metabolic profile of cells in distilleries fermentative tanks on different productivity situations, an untargeted analysis of the intracellular metabolic profile was conducted.

Since untargeted metabolomics is an exploratory analysis of global metabolite profiling, its evaluation is mainly based on the metabolites comparison between groups (GERTSMAN; BARSHOP, 2019), our first goal was to check for differences between Low and Normal productivity groups. Based on the PCA of the detected metabolites, we observed the separation of samples into two groups related to ethanol fermentative yield. So we started to compare the relative abundance of metabolite in each group aiming to observe differences in the intracellular metabolism of fermenting microorganisms. Due to the impossibility of identifying the origin of detected metabolite and the lack of

quantitative and qualitative data on the microorganisms groups present in each sample, we could not assess or address specific changes in metabolic pathways of yeast strains or other contaminants. Thus, in this work, we only evaluated the general changes in the intracellular metabolic profile and hypothesized about the role of each metabolite in the most commonly related microorganisms of ethanol fermentation industries. In this context, a greater relative abundance of specific metabolic compounds in each production situation may be related to the predominance of certain microorganisms capable of producing or metabolizing these compounds, an increase in the production of the metabolite, or to responses of the microbial community to environmental changes.

When observing the metabolites pointed out as more abundant in Tank 1, in the Normal productivity group we could identify a variety of sugar compounds detected. Glucose is the main sugar used for ATP synthesis in cellular metabolism, so the detection of this metabolite in the beginner of the fermentation line may be related to the rapid intake of this metabolite by cells. The lower abundance of this metabolite in cells of the Low productivity group, may be related to a decrease of disponibility of this sugar due to consumption by contaminants in the earlier stages. Commercial yeast strains are known to rapidly ferment glucose even in high sugar concentrations (DELLA-BIANCA; GOMBERT, 2013), which should be relate with the greater relative abundance of this metabolite only in Tank 1 of Normal productivity situations. Although maltotriose is not the preferred sugar of all *Saccharomyces cerevisiae*, some industrial strains and other yeast can use this sugar efficiently (DUVAL et al., 2010), maybe justifying its greater abundance in the intracellular metabolic profile only in the first tank of the fermentative line. Galactose is also probably related to energy production since it can be converted to glucose 6-phosphate in several organisms (NELSON; COX, 2018). Galactose and D-arabinono-1,4-lactone are also involved in the erythroascorbic acid pathway in yeasts (SAUER et al., 2004), an important yeast antioxidant. Sorbose was also relatively more abundant in the Normal productivity group. This monosaccharide is a naturally occurring sugar and can be obtained by the fermentation of sorbitol or D-glucose by some bacteria (SHINTANI, 2019). In The Low productivity group the metabolite 2-Furoic acid was detected as more abundant. This metabolite has already been reported as a by-product of some yeasts and bacteria (BOOPATHY; BOKANG; DANIELS, 1993; NICHOLS et al., 2008). The detection of glucosaminic acid may be related to the presence of *Pseudomonas* bacteria since its production has already been related to this genus (WU et al., 2011). This bacterial genus has already been reported as a contaminant in the ethanol distilleries (BREXÓ; SANT'ANA, 2017), and the detection of those metabolites could be related to its presence.

In the fermentative Tanks 2 and 3, it was possible to observe a similar detection of some intracellular metabolites in both productivity groups. In both tanks, the metabolic profile of Normal productivity showed a greater relative abundance of Glucose-6-phosphate and Lactulose. The detection of Glucose-6-phosphate is not a surprise since this compound is an intermediate of glucose metabolism, involved in several metabolic pathways, including ethanolic fermentation (NELSON; COX, 2018). On the other hand, the presence of the metabolite lactulose in the intracellular metabolic profile may be related to the presence of some bacteria, in the Normal productivity group, which manage to ferment this disaccharide (SCHUSTER-WOLFF-BÜHRING; FISCHER; HINRICHS, 2010).

In Tanks 2 and 3 were possible to detect the metabolites Mannitol, 3-Methyl-2-oxopentanoic-acid, Trehalose and 1-Methylhistidine as more abundant in Low productivity sample group. The presence of mannitol in fermentative tanks is monitored by some distilleries since this metabolite is an indicator of bacterial contamination (EGGLESTON et al., 2007; LOPES et al., 2016), so its greater abundance is probably related to higher contamination. In a study of homo and heterofermentative bacteria, Basso et al (2014) reported that mannitol is primarily produced by heterofermentative bacteria, so its higher abundance in Low productivity group could be related to an increase of heterofermentative bacteria population in this conditions. In *Saccharomyces* cells, trehalose acts as a stress protectant and

the accumulation of this carbohydrate is a common trait of selected yeast strains (BASSO et al., 2008). However, the concentration of this metabolite varies during the fermentation since it is degraded and restored according to sugar levels and stressing factors in the medium (BASSO; BASSO; ROCHA, 2011). Thus, the detection of trehalose in Tank 2 is probably related to an early decrease in sugar levels and/or an increase in stressful conditions in low productivity situations. We are not aware of any relationship between metabolites 3-Methyl-2-oxopentanoic-acid and 1-Methylhistidine with the metabolism of commonly found microorganisms in ethanol distilleries.

Our work aimed to explore differences in the samples' intracellular metabolic profile, from an ethanol distillery, during low and normal fermentative yield. Since it was not possible to identify, quantify or separate each cell type in the sample without affecting the cell metabolism, the global intracellular metabolic profile was evaluated and no correlation with specific metabolic pathways in determinate microorganisms was possible to be explored. Nevertheless, it was possible to relate the greater abundance of some metabolites to possible changes in the fermentation tank environment and to microorganisms commonly found in distilleries. To our knowledge, this was the first exploratory analysis of intracellular metabolic profile using GC-MS, in the distillery environment. This methodology proved to be capable of detecting important metabolic differences in the studied situations. We believe that further studies should be conducted to clarify the role of contaminants in yeast cell metabolism on fermentative tanks. The impact of changes in yeast metabolisms and the decrease in fermentative yield must also be evaluated in the future.

References

- AMORIM, H. V. et al. Scientific challenges of bioethanol production in Brazil. **Applied Microbiology and Biotechnology**, v. 91, n. 5, p. 1267–1275, 2011.
- BASSO, L. C. et al. Yeast selection for fuel ethanol production in Brazil. **FEMS Yeast Research**, v. 8, n. 7, p. 1155–1163, 2008.
- BASSO, L. C.; BASSO, T. O.; ROCHA, S. N. Ethanol Production in Brazil : The Industrial Process and Its Impact on Yeast Fermentation. In: MAS, B. (Ed.). **Biofuel Production-Recent Developments and Prospects**.
- BASSO, T. O. et al. Homo- and heterofermentative lactobacilli differently affect sugarcane-based fuel ethanol fermentation. **Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology**, v. 105, n. 1, p. 169–177, 2014.
- BOOPATHY, R.; BOKANG, H.; DANIELS, L. Biotransformation of furfural and 5-hydroxymethyl furfural by enteric bacteria. **Journal of Industrial Microbiology**, v. 11, n. 3, p. 147–150, 1993.
- BREXÓ, R. P.; SANT'ANA, A. S. Impact and significance of microbial contamination during fermentation for bioethanol production. **Renewable and Sustainable Energy Reviews**, v. 43, p. 423–434, 2017.
- BUDZINSKI, I. G. F. et al. Network Analyses and Data Integration of Proteomics and Metabolomics From Leaves of Two Contrasting Varieties of Sugarcane in Response to Drought. **Frontiers in Plant Science**, v. 10, n. November, p. 1–19, 2019.
- CANELAS, A. B. et al. Leakage-free rapid quenching technique for yeast metabolomics. **Metabolomics**, v. 4, n. 3, p. 226–239, 2008.
- DELLA-BIANCA, B. E. et al. What do we know about the yeast strains from the Brazilian fuel ethanol industry?

- Applied Microbiology and Biotechnology**, v. 97, n. 3, p. 979–991, 2013.
- DELLA-BIANCA, B. E.; GOMBERT, A. K. Stress tolerance and growth physiology of yeast strains from the Brazilian fuel ethanol industry. **Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology**, v. 104, n. 6, p. 1083–1095, 2013.
- DUVAL, E. H. et al. Microarray karyotyping of maltose-fermenting *Saccharomyces* yeasts with differing maltotriose utilization profiles reveals copy number variation in genes involved in maltose and maltotriose utilization. **Journal of Applied Microbiology**, v. 109, n. 1, p. 248–259, 2010.
- EGGLESTON, G. et al. Mannitol as a sensitive indicator of sugarcane deterioration and bacterial contamination in fuel alcohol production. **Zuckerindustrie**, v. 132, n. 1, p. 33–39, 2007.
- GERTSMAN, I.; BARSHOP, B. A. Promises and Pitfalls of Untargeted Metabolomics. **J Inherit Metab Dis**, v. 41, n. 3, p. 355–366, 2019.
- JOHNSON, C. H.; IVANISEVIC, J.; SIUZDAK, G. Metabolomics: beyond biomarkers and towards mechanisms. **Nature Reviews Molecular Cell Biology**, v. 17, n. July, p. 451–459, 2016.
- KRUMSIEK, J. et al. Mining the Unknown : A Systems Approach to Metabolite Identification Combining Genetic and Metabolic Information. **PLoS Genetics**, v. 8, n. 10, 2012.
- LOPES, M. L. et al. Ethanol production in Brazil: a bridge between science and industry. **Brazilian Journal of Microbiology**, v. 47, p. 64–76, 2016.
- NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger**. 7. ed. Artmed Editora, 2018.
- NICHOLS, N. N. et al. Fungal metabolism of fermentation inhibitors present in corn stover dilute acid hydrolysate. **Enzyme and Microbial Technology**, v. 42, n. 7, p. 624–630, 2008.
- SAUER, M. et al. Production of L-Ascorbic acid by metabolically engineered *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*. **APPLIED AND ENVIRONMENTAL MICROBIOLOGY**, v. 70, n. 10, p. 6086–6091, 2004.
- SCHUSTER-WOLFF-BÜHRING, R.; FISCHER, L.; HINRICHS, J. Production and physiological action of the disaccharide lactulose. **International Dairy Journal**, v. 20, n. 11, p. 731–741, 2010.
- SHINTANI, T. Food Industrial Production of Monosaccharides Using Microbial , Enzymatic , and Chemical Methods. **Fermentation**, v. 5, n. 47, 2019.
- SMART, K. F. et al. Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography – mass spectrometry. **NATURE PROTOCOLS**, v. 5, n. 10, p. 1709–1729, 2010.
- WU, B. et al. Efficient production of D-glucosaminic acid from D-glucosamine by *Pseudomonas putida* GNA5. **Biotechnology Progress**, v. 27, n. 1, p. 32–37, 2011.
- ZHANG, X. et al. Mass spectrometry-based metabolomics in health and medical science : a systematic review. **The Royal Society of Chemistry**, v. 10, p. 3092–3104, 2020

SUPPLEMENTARY

Figure S1- PCA of Low and Normal productivity profile. Intracellular metabolic profile regardless sampling site. Low and Normal productivity samples are represented by red and green dots, respectively.

Supplementary Table 1

Sample	Sampling site	Group	Low	1	2	3	31	37	41	70	99	106	130	139	160	180	192	217	238	254	331	337	402	469	512	550	562
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266	0	0.190944	0	0.797382	0	0.080755	0
L1	Trunk 1 A	Low T1	0	0.29247	0	0.30022	0.065415	0	0.0	0	0	0.30228	1.96565	0	0.181176	12.25972	68.9752	77.4442	4.598456	12.27266							

to be continued

[illegible]

to be continued

[illegible]

to be continued

Sample	Sampling site	Group	1482	1440	1415	1406	1384	1350	1363	1362	1695	1697	700	1704	1727	1753	1810	1840	1883	1893	1894	1882	2011	2019	2027	2056	2063
L1	Low Moist	Low Moist	0	0	0	0.60675	0	0	0	0	0	0	0	15.1118	18.2714	0	0	0	0	0	0	30.11382	0	67.30074	474.401	0	
L1.1	Low T1	Low T1	0	6.118769	0	19.46037	14.75728	4.535765	12.49702	0	7.838837	10.33	0	13.11773	12.75927	15.51184	20.04449	6.690927	0	0	0	1.294574	14.51517	14.93902	0	0	
L1.2	Low T1	Low T1	0	7.144469	0	9.483499	15.932	12.51527	13.93883	0	17.88837	14.56759	0	15.52216	15.19988	18.31665	22.55186	6.690927	0	0	0	1.294574	12.35872	13.93912	0	0	
L1.3	Low T1	Low T1	0	8.88294	0	7.830455	17.21357	9.908189	12.19585	0	18.8031	17.84463	0	19.9141	18.16653	16.69531	1.660706	0	0	0	0	1.405983	6.302399	6.58117	0	0	
L2	Low T2	Low T2	0	6.189379	0	22.85527	0	10.10414	0	10.03178	8.962238	6.744852	0	10.45513	10.27006	13.24604	18.22558	13.98359	0	0	0	0	0	0	0	0	
L2.1	Low T2	Low T2	0	6.189379	0	22.85527	0	10.10414	0	10.03178	8.962238	6.744852	0	10.45513	10.27006	13.24604	18.22558	13.98359	0	0	0	0	0	0	0	0	
L2.2	Low T2	Low T2	0	6.189379	0	22.85527	0	10.10414	0	10.03178	8.962238	6.744852	0	10.45513	10.27006	13.24604	18.22558	13.98359	0	0	0	0	0	0	0	0	
L2.3	Low T2	Low T2	0	6.189379	0	22.85527	0	10.10414	0	10.03178	8.962238	6.744852	0	10.45513	10.27006	13.24604	18.22558	13.98359	0	0	0	0	0	0	0	0	
L3	Low Acid treatment	Low Acid treatment	0	12.82821	0	19.98858	22.48311	27.33563	0	24.84268	19.13853	0	21.50678	22.26135	4.076953	12.64102	15.52624	6.950555	0	0	0	0	3.986074	12.55415	0	0	
L3.1	Low Acid treatment	Low Acid treatment	0	5.492424	0	17.57238	23.5024	23.16615	20.28622	0	19.15701	0	0	21.50678	22.26135	4.076953	12.64102	15.52624	6.950555	0	0	0	0	0	0	0	
L3.2	Low Acid treatment	Low Acid treatment	0	14.83532	0	24.84933	8.255338	6.495973	12.15194	0	14.84287	7.775997	0	14.08004	17.19119	10.2022	19.8805	10.15896	0	0	0	15.42822	9.287101	6.515466	0	0	
L3.3	Low Acid treatment	Low Acid treatment	0	7.128003	0	10.61762	13.55561	7.11881	0	23.66273	0	0	24.45556	19.10312	10.188	10.20075	8.572148	0	0	0	0	24.54891	16.05925	0	0	0	
L3.4	Low Acid treatment	Low Acid treatment	0	16.67275	0	13.93381	8.00005	0	12.50717	0	17.48135	10.29605	0	14.41139	10.29605	0	0	0	0	0	0	1.797629	14.47455	9.931914	0	0	
L3.5	Low Acid treatment	Low Acid treatment	0	16.67275	0	13.93381	8.00005	0	12.50717	0	17.48135	10.29605	0	14.41139	10.29605	0	0	0	0	0	0	1.797629	14.47455	9.931914	0	0	
L4	Low Acid treatment	Low Acid treatment	0	14.83532	0	24.84933	8.255338	6.495973	12.15194	0	14.84287	7.775997	0	14.08004	17.19119	10.2022	19.8805	10.15896	0	0	0	15.42822	9.287101	6.515466	0	0	
L4.1	Low T1	Low T1	0	0	0	0	1.923413	7.2461	0	6.58887	0	0	0.688987	0	0.325725	0	6.957182	2.110511	0	0	0	7.20479	3.211474	0	0	0	
L4.2	Low T1	Low T1	0	17.48407	0	4.10176	6.73672	0.200142	9.869911	8.031683	6.758884	0	0	0.303964	0	0	4.28598	21.75867	0	0	0	2.445834	3.067521	3.406169	0	0	
L4.3	Low T1	Low T1	0	7.5634	5.4626	6.260653	0	0	22.5592	16.70092	11.80775	0	0	0.782189	8.000511	0	0	29.15111	0	0	0	20.05237	7.807113	0	2.130716	0	
L4.4	Low T1	Low T1	0	0	0	0	0	0	0	0	0.469152	0	0	0.967122	0.021932	0	0	1.91074	8.973429	0	0	0	0	0	0	0	
L4.5	Low T1	Low T1	0	21.68466	0	28.61626	10.79543	8.86602	13.20706	0	12.51612	0	0	19.1115	8.88594	16.49493	24.6688	14.2934	0	0	0	10.06441	16.640733	0	20.20253	1.395798	
L4.6	Low Acid treatment	Low Acid treatment	0	11.50963	0	14.19076	20.95743	5.34231	9.048843	5.97731	7.043132	11.05834	0	9.625478	10.03771	8.99443	2.24668	10.2721	0	0	0	2.566032	5.184435	3.234976	0	0	
L4.7	Low Acid treatment	Low Acid treatment	0	0.609635	0	10.27495	10.54394	51.62173	18.90573	1.648708	0.799674	0	0	1.072342	0	0	1.504318	1.800074	0	0	0	0.43173	0	0	0	0	
L4.8	Low Moist	Low Moist	0	0	0	0	0	0	0	0	0	0	0	3.033033	0	0	0	3.521819	75.3748	0	0	66.55827	0	165.291	863.8967	0	
L4.9	Low Moist	Low Moist	0	0	0	0	0	0	0	0	0	0	0	3.033033	0	0	0	0	0	0	0	1.1868	0	0	0	0	
L5	Normal T1	Normal T1	0	6.918027	0	9.460446	0	0	0	0	0	0	0	4.795527	8.14561	0	4.498348	4.652118	0	0	0	0	0	0	0	0	
L5.1	Normal T1	Normal T1	0	11.22752	5.49151	9.521817	0	0	0	0	0	0	0	2.763144	10.93631	0	0	3.866269	0	0	0	0	0	0	0	0	
L5.2	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.3	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.4	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.5	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.6	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.7	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.8	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L5.9	Normal T1	Normal T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L6	Normal T2	Normal T2	0	0.215936	0	0	0	0	0	0	0	0	0	0.801991	0.757957	0	0	6.453355	0	0	0	74.92114	0	5.77202	56.11004	0	
L6.1	Normal T2	Normal T2	0	9.940208	0	13.21055	0	0	0	0	0	0	0	10.04599	14.81841	12.81382	15.20083	16.33388	0	0	0	12.12071	0	7.85792	10.6713	7.959465	
L6.2	Normal T2	Normal T2	0	10.40022	0	7.714492	7.631867	0	0	0	0	0	0	20.95848	14.17996	15.79719	37.39304	32.27696	0	0	0	11.68499	0	8.76796	11.44472	7.279716	
L6.3	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L6.4	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L6.5	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L6.6	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L6.7	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L6.8	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L6.9	Normal T2	Normal T2	0	0	0	0	0	0	0	0	0	0	0	14.13525	0	0	0	0	0	0	0	0	0	0	0	0	
L7	Normal T3	Normal T3	0	11.22752	5.49151	9.521817	0	0	0	0	0	0	0	2.763144	10.93631	0	0	3.866269	0	0	0	0	0	0	0	0	
L7.1	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.2	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.3	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.4	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.5	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.6	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.7	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.8	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L7.9	Normal T3	Normal T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L8	Normal T4	Normal T4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L8.1	Normal T4	Normal T4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L8.2	Normal T4	Normal T4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
L8.3	Normal T4	Normal T4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0</										

to be continued

6 AUTHOR'S FINAL CONSIDERATIONS AND FUTURE PERSPECTIVES

Presenting the thesis in the format of drafted manuscripts has the advantage of making the thesis more compact, so it reports objectively the main positive results obtained over the doctoral research. However, only a small fraction of the research usually has the potential to be published, hence the majority of the developed work is not reported. Furthermore, in this thesis format, each chapter corresponds to a different manuscript and connections between them may not be properly addressed in the text. To overcome some of these disadvantages, this section will briefly address part of our initial objectives, difficulties, general conclusions, as well as the final considerations.

When assessing changes in the contaminant community and the metabolic profile under different fermentative yields, we believe to have contributed to expanding the knowledge about the effects of contaminating microorganisms in ethanol distilleries' productivity. The main challenge of this work was to conduct scientific research in the industrial system. Such a noncontrolled and diverse environment directly impacts the sample evaluation. It is worth mentioning that sampling was conducted in collaboration with a distillery operating in a continuous fermentation line, which brought greater complexity and difficulties to this work.

Therefore, sampling methodologies were adapted to ensure that industry safety measures were respected and the sample's quality maintained. Before the actual sampling, several tests were carried out to define sampling methodologies, reduce the impact of the samples' transportation, and to test distinct protocols for analytical steps. After defining and testing the chosen methodology, 16 sampling procedures were performed across the fermentative line between August and November 2017. Until mid-September, the distillery operated with two fermentative lines simultaneously but for operational reasons, this practice was abandoned. We believe that this did not bring any prejudice to the study. Samples were simultaneously collected for metabarcoding analysis, extracellular metabolomics, and intracellular metabolic profile analysis.

Our first attempt was to perform a temporal sampling to evaluate blocks of consecutive days, but distillery's operational conditions, unexpected changes of the distillery process, and climatic conditions prevented the success of this monitoring. It is important to highlight that we tried to standardize the evaluated samples and reduce as much as possible the influence of external parameters. Some samples were disregarded, for example, due to industrial technical issues reported after sampling or changes in must production.

Our first objective was to assess the contaminant microbiota under distinct ethanol yields, but we also tried to correlate other parameters such as flocculation and the usage of antibiotics. Unfortunately, distillery measurements of those parameters were not rigorous as the control of fermentative efficiency and we ended up not being able to proceed with the complementary evaluations.

Microbiota assessments were conducted at the University of Copenhagen through metabarcoding approaches. Due to legal difficulties to ship the collected material, DNA extractions were performed in Brazil. Considering logistic and costs involved in this analyzes, it was necessary to previously define samples to be evaluated. Based on the distillery data we chose some samples of the must, tank 1, and tank 3 to be analyzed.

At first, we would evaluate the bacterial and yeast community through metabarcoding approaches. However, we did not succeed with the yeast community analysis. In addition to ITS amplification difficulties, the use of a single primer pair was not an assertive choice to evaluate the contaminant yeast community since 99% of the reported reads were assigned to the *Saccharomyces* genus. We are aware that, with metabarcoding techniques, it is not possible to distinguish species, but we were expecting to detect other yeast contaminants. Thus, we only

successfully evaluated the bacterial community, and our results were presented in the chapter/manuscript: ASSESSMENT OF BACTERIAL COMMUNITIES PROFILE IN A FIRST-GENERATION ETHANOL DISTILLERY ON NORMAL AND LOW FERMENTATIVE YIELD.

Our second objective was to explore the distillery's metabolic profile under distinct fermentation yield. Metabolomics analyses were performed in our lab in Brazil, which allowed testing several analytical parameters and metabolites extraction protocols specifically for this work. It is worth mentioning that all tests were carried out with a mass spectrometry specialist, to ensure analytical quality. Another advantage of performing the sample analysis in our lab was the ability to assess all collected samples. To evaluate the fermentative line extracellular and intracellular profile, samples from must, all fermentative tanks, and acid treatment were analysed through untargeted GC-MS. Please notice that in both metabolomics chapters we evaluated the same sample set, but not in the metabarcoding chapter.

In the chapter "UNTARGETED METABOLOMICS OF FIRST GENERATION ETHANOL: COMPARING THE METABOLIC PROFILE OF NORMAL AND LOW ETHANOL YIELDS" it was performed an exploratory analysis of distillery extracellular metabolites on distinct fermentative yields. We also checked the potential of metabolomics in identifying promising productivity indicators for future studies. Initially, we also aimed to evaluate the microbial dynamics, but without prior knowledge of the sample's microorganisms, it was not possible to speculate about metabolic interactions. In this way, in this chapter only superficial extrapolations were made considering the microorganisms commonly reported in distilleries. Nevertheless, for the exploration of these data from a biotechnological perspective, we do not consider it extremely important to identify the origin of each metabolite in an initial analysis.

Originally, we also intended to correlate the intracellular metabolic profile with changes in yeast metabolism. However, due to the sample's characteristics, it was not possible to assign any metabolite to a specific microorganism. We looked for methodologies to separate some microorganisms before the metabolite extraction, but since metabolites change rapidly in response to environmental conditions, no cell separation method could be used without compromising the final metabolic profile. Thus, in the chapter "UNTARGETED INTRACELLULAR METABOLIC PROFILE OF FERMENTATIVE TANKS FROM A BRAZILIAN ETHANOL DISTILLERY UNDER DISTINCT ETHANOL YIELDS" we explored the intracellular community metabolome and tried to identify potential metabolic changes, in commonly reported yeast and contaminating microorganisms, related to stressing conditions of the distillery. We are aware that some expected metabolites have not been detected and no direct correlation between extracellular and intracellular metabolites was observed. This can be partially explained by the chosen extraction methodology and future works should attempt to overcome this.

Interestingly, when extrapolating the data from all the chapters, it was possible to observe some potential relations between the reported bacterial community dynamics and the metabolic profiles. We can highlight the possible relation between oleic acid and *Lactobacillus* in the must samples, for example. This metabolite is known to be a major constituent of yeast cells and a growth factor for lactic acid bacteria and meanwhile is an inhibitor of other bacterial genera, e.g. *Bacillus*. Its greater abundance in the must samples of the low fermentative productivity group is probably related to a greater abundance of *Lactobacillus*, which was reported as an important contaminant of Low productivity must samples. It is also important to highlight that *Geobacillus* were previously classified as *Bacillus* and may have their growth inhibited by this metabolite. Another interesting fact was the higher abundance of mannitol in the intracellular metabolome of fermentative tanks on samples belonging to the activity

group. This metabolite is related to the presence of heterofermentative bacteria and in the microbiota evaluation of Low productivity samples, we detected a different OTU pattern for *Lactobacillus* contamination in fermentative tanks of Normal and Low productivity. These data together suggest that there are an increase or predominance of heterofermentative *Lactobacillus* in tanks under Low productivity conditions.

Despite all the difficulties to conduct a study in an industrial environment, we succeed to evaluate relevant differences in the bacterial community and metabolic profile of low and normal fermentative yields. We have demonstrated that the bacterial contaminants community is different in distinct ethanol yields, especially in the must. These differences are mainly related to the genera *Lactobacillus* and *Geobacillus* and future studies should clarify this interaction and the role of must contaminants in the fermentation success. It is common knowledge that bacterial monitoring in the ethanol distilleries is quantitative, but our data suggest that qualitative analyses should be considered as well. Through the metabolic profile analyses, differences regarding productivity were also detected. Our results suggest that some metabolites have the potential to be explored as productivity monitoring tools, which could aid in decision making throughout the ethanol fermentation.

In the future, we hope to contribute with distilleries decision making in controlling the fermentation process to achieve higher efficiencies. For now, we strongly recommend that distilleries should be more careful with the fermentation feedstock, improve asepsis of all process stages, and invest in more efficient ways to monitor more detrimental bacterial contaminants, especially in the must.