

Integrating microbial metagenomics and physicochemical parameters and a new perspective on starter culture for fine cocoa fermentation

Carolina O. de C. Lima ^{a,1}, Aline B.M. Vaz ^{b,1}, Giovanni M. De Castro ^{b,1}, Francisco Lobo ^b, Ricardo Solar ^b, Cristine Rodrigues ^c, Luiz Roberto Martins Pinto ^d, Luciana Vandenberghe ^c, Gilberto Pereira ^c, Andréa Miúra da Costa ^d, Raquel Guimarães Benevides ^a, Vasco Azevedo ^b, Ana Paula Trovatti Uetanabaro ^{d,1}, Carlos Ricardo Soccoll ^c, Aristóteles Góes-Neto ^{a,b,*}

^a Department of Biological Sciences, Universidade Estadual de Feira de Santana (UEFS), Feira de Santana, BA, 44036-900, Brazil

^b Institute of Biological Sciences, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, 31270-901, Brazil

^c Bioprocess Engineering and Biotechnology Department, Universidade Federal do Paraná (UFPR), Curitiba, PR, 81531-980, Brazil

^d Department of Biological Sciences, Universidade Estadual de Santa Cruz (UESC), Ilhéus, BA, 45662-900, Brazil

ARTICLE INFO

Keywords:

Theobroma cacao
Cocoa fermentation
Ecological succession
Microbiome
Starter inoculum
Chocolate

ABSTRACT

Cocoa beans used for chocolate production are fermented seeds of *Theobroma cacao* obtained by a natural fermentation process. The flavors and chemical compounds produced during the fermentation process make this step one of the most important in fine chocolate production. Herein, an integrative analysis of the variation of microbial community structure, using a shotgun metagenomics approach and associated physicochemical features, was performed during fermentation of fine cocoa beans. Samples of Forastero variety (FOR) and a mixture of two hybrids (PS1319 and CCN51) (MIX) from Bahia, Brazil, were analyzed at 7 different times. In the beginning (0 h), the structures of microbial communities were very different between FOR and MIX, reflecting the original plant-associated microbiomes. The highest change in microbial community structures occurred at the first 24 h of fermentation, with a marked increase in temperature and acetic acid concentration, and pH decrease. At 24–48 h both microbial community structures were quite homogenous regarding temperature, acetic acid, succinic acid, pH, soluble proteins and total phenols. During 72–96 h, the community structure resembles an acidic and warmer environment, prevailing few acetic acid bacteria. Taxonomic richness and abundance at 72–144 h exhibited significant correlation with temperature, reducing sugars, succinic, and acetic acids. Finally, we recommend that dominant microbial species of spontaneous fine cocoa fermentations should be considered as inoculum in accordance with the farm/region and GMP to maintain a differential organoleptic feature for production of fine chocolate. In our study, a starter inoculum composed of *Acetobacter pauserianus* and *Hanseniaspora opuntiae* strains is indicated.

1. Introduction

Cocoa pulp mass is a three to six days spontaneous fermentation process, showing a succession of microbial activities of yeasts, lactic acid bacteria (LAB), and acetic acid bacteria (AAB) (De Vuyst and Weckx, 2016). As a result of microbial action, cocoa pulp is degraded into different end-metabolites, which diffuse into the cocoa beans and promotes the death of the embryo, transforming the seeds into beans that are suitable for subsequent drying and roasting for chocolate production (Thompson et al., 2007; Vásquez et al., 2019; Wood and Lass, 2001).

Ecologically, the fermentation process promotes successional environmental conditions that sequentially select the natural microbial community able to develop under these changing conditions. One of the first groups of microorganisms usually found in cocoa fermentation are the yeasts, which promote the removal of pectin and metabolize simple carbohydrates into ethanol, carbon dioxide, and volatile compounds, leading to a temperature increase (Lima et al., 2011). The most common fungal genera are *Saccharomyces*, *Pichia*, and *Hanseniaspora*, as revealed by both culture-dependent and culture-independent studies (Araña-sáñchez et al., 2015; De Vuyst and Weckx, 2016; Koffi et al., 2017;

* Corresponding author. Department of Biological Sciences, Universidade Estadual de Feira de Santana (UEFS), Feira de Santana, BA, 44036-900, Brazil.

E-mail addresses: arigoesneto@icb.ufmg.br, arigoesneto@gmail.com (A. Góes-Neto).

¹ These authors have contributed equally to this work.

Mota-Gutierrez et al., 2018; Pereira et al., 2017; Schwenninger et al., 2016; Papalexandratos et al., 2019). The LAB are usually the second group of microorganisms appearing during the fermentation process, metabolizing glucose and fructose to lactic acid, glycerol, mannitol, carbon dioxide, and several volatile compounds (De Vuyst et al., 2010; De Vuyst and Weckx, 2016; Lima et al., 2011; Moreira et al., 2013; Mota-Gutierrez et al., 2018; Papalexandratos et al., 2011, b; Papalexandratos et al., 2013; Schwenninger et al., 2016). *Lactobacillus* and *Leuconostoc* are ubiquitous LAB genera reported in cocoa beans fermentation (De Vuyst and Weckx, 2016; Mota-Gutierrez et al., 2018; Ouattara et al., 2017; Schwenninger et al., 2016). The last group of microorganisms is the AAB, mainly represented by *Acetobacter* and *Gluconobacter*, which are able to use ethanol to produce acetic acid as the final product of their metabolism (Illegems et al., 2013; De Vuyst and Weckx, 2016; Mota-Gutierrez et al., 2018; Schwenninger et al., 2016). The flavors and chemical substances produced during the fermentation process make this step one of the most important in the chocolate production (Menezes et al., 2016), which reinforces the need for studies that better understand the ideal microbial consortium (Craack et al., 2013; De Vuyst and Weckx, 2016; Lefever et al., 2011; Moreira et al., 2017; Papalexandratos and Nielsen, 2016; Pereira et al., 2012; Visintin et al., 2017).

The cocoa production has a high economic value, with a production of 4587 thousand tonnes world production in 2018 (, 2018 Quarterly Bulletin of Cocoa Statistics). Furthermore, the search for chocolates with differentiated organoleptic characteristics by direct and indirect consumers has been increasing, especially in the region of Ilhéus, Bahia, Brazil. This region is undergoing a transformation among cocoa producers who, nowadays, besides producing the award-winning beans of internationally recognized accreditation of quality, are also producing their own chocolate in their customized agroindustries (Dantas et al., 2020). GMP have been employed to obtain high-quality cocoa beans for the production of fine chocolate.

The current work analyzed 14 samples of cocoa beans (2 groups, 7 distinct periods of time), in a Brazilian Agroindustry (tree to bar concept) that uses good manufacturing practices (GMP), and produces fine chocolate. The community structure and their spatial and temporal variation were analyzed, encompassing taxonomic composition, richness, read abundance and diversity indexes of whole microbiomes at genus and species level, and their correlations and ordination jointly with physicochemical variables, by shotgun metagenomics during 6-day cocoa fermentations.

2. Material and methods

2.1. In situ sample collection and analyses

The fermentation process was carried out with *Theobroma cacao* L. seeds of (i) Forastero variety (FOR) and (ii) a mixture of two hybrids (PS1319 and CCN51) (MIX) at Riachuelo Agroindustry of Mendoá Chocolates (www.mendoachocolates.com.br) (Uruçuca, Bahia, Brazil; Lat –14.7719058; Long –39.0492701). Planting, management, harvesting and fermentation process followed GMP standards used in this agroindustry. Ripe and healthy pods were washed, dried and manually opened with a machete. Cocoa seeds were immediately transferred to wooden boxes (45 × 45 × 45 cm), with an overall capacity of 40 kg, inside the fermentation room, limiting the chances of contamination due to transportation (Supplementary File - Fig. S1).

2.2. Field experimental design

Experimental design consisted of two blocks, one for FOR and one for MIX. In each block, there were three fermentation boxes that were sampled in seven distinct times: 0 (cocoa seeds just after opening the pods), 24, 48, 72, 96, 120 and 144 h. Subsamples of 30 g each were collected in five random points in fermentation mass of each wooden

box and were mixed to form a 150 g sample from which 100 g were stored at -20 °C in sterilized plastic bags for shotgun metagenomics, and 50 g were refrigerated (4 °C) for subsequently physicochemical analyses. Cocoa mass mean temperature was measured in five different points in each wooden box, during all the experiments.

2.3. Metagenomic DNA extraction

The metagenomic DNA of the samples was extracted from lyophilized (LP3, Jouan) pulp fraction of the cocoa mass (pulp + beans) based on Cota-sanchez et al. (2006) with some modifications, such as: (i) initial washing in sorbitol buffer to allow a decrease in the presence of substances that are harmful to a good quality extraction, (ii) use of a larger volume of CTAB solution, and (iii) use of RNase, lysozyme and proteinase K.

2.4. Massively parallel sequencing

A total amount of 1 µg of DNA per sample was used for the preparation of metagenomic libraries according to Illumina's standard protocols, and the sequencing was performed on an Illumina Hiseq X Ten (Novogene Sequencing Laboratory, UC Davis Medical Center). All the raw sequences were deposited in NCBI SRA under Bioproject accession PRJNA552479.

2.5. Bioinformatics analysis

Data preprocessing was as follows: (1) FASTQC (v0.11.4) (Andrews, 2010) for quality analysis; (2) Adapter removal with CUTADAPT (v. 1.18) (Martin, 2011); (3) Alignment using BOWTIE2 (v. 2.3.4.3) (Langmead and Salzberg, 2012) to filter for pairs of reads not mapping on *T. cacao* (genome from NCBI Criollo cocoa genome V2); (4) Filtered reads were merged using FLASH (v1.2.11) (Magoc and Salzberg, 2011).

For the taxonomic assignment, two strategies were used: (a) the first one using all reads against a database using CENTRIFUGUE (v. 1.0.3) (Kim et al., 2016) and KAIJU (v. 1.6.2) (Menzel et al., 2016), and (b) another using only the 16 S and 18 S rDNA reads using MAPSEQ (v. 1.2.3) (Matias Rodrigues et al., 2017). All the complete bacterial, archaeal, and fungal genomes available in NCBI were used for CENTRIFUGUE and KAIJU analyses, whereas the default database comprising curated reference of full-length rRNA genes pre-classified to taxonomic categories based on the NCBI taxonomy were used for MAPSEQ. The complete pipeline of bioinformatics analyses, including a detailed explanation of each step with the command lines can be accessed in Supplementary File – Script S12.

2.6. Community ecology analyses

Rarefaction curves were constructed with Inext package (Chao et al., 2013; Hsieh et al., 2016), and taxon diversity was measured using taxon abundance, richness, and the following diversity indexes (Colwell, 2009): Shannon normalized, and Simpson normalized (Gardener, 2014). Homogeneity of abundances (evenness) was measured using the Pielou's index (Legendre and Legendre, 1998), and the diversity was estimated using the Shannon and Simpson indices. We also estimated taxon richness using Chao2 (Chao, 1987). Correlation analyses were based on Spearman's *r*, and Principal Component Analysis (PCA) were carried out using the vegan package (Oksanen et al., 2007). All the analyses were performed using the R package (R113) (R Developmental Core Team., 2019), and the complete script is in the Supplementary File – Script S13.

2.7. Physicochemical analyses of the beans

Prior to the physicochemical analyses of the beans, all the pulp was removed, and they were subsequently crushed in a multiprocessor. Moreover, all the analyses were performed in triplicate. Carbohydrates

(total and reducing sugars) were quantified by using DNS reagent (Miller, 1959), analyzed in a spectrophotometer (UV-SP2000) at 540 nm, and the results were expressed in milligrams of sugar per gram of dry beans (mg.g^{-1}). Soluble Proteins were determined in accordance with Bradford (1976), determined by spectrophotometry at 595 nm, and the results were expressed as milligram of protein per gram of dry beans (mg.g^{-1}). Total phenols were determined using the modified method of Folin-Ciocalteu (Singleton and Rossi, 1965), measuring absorbances at 740 nm, with methanol as a blank control, and the results were expressed as gallic acid equivalents (GAE) in milligrams per 100 g dry mass ($\text{mg GAE/100 g dry mass}$). Alcohols (ethanol and methanol) and organic acids (acetic, citric, lactic, and succinic) were determined by chromatography (HP1200 series; Hewlett-Packard Company), attached to Hi PlexH Agilent column. pH was measured by the potentiometric method (IAL, 1985) while moisture was measured in percentage by the gravimetric method (IAL, 1985), and the value of the dry mass was obtained of the initial mass minus the final mass, after drying.

3. Results and discussion

3.1. Overview

The high-throughput sequencing of metagenomic DNA generated a total of 47,133 gigabases from all the 14 samples, representing 314,221,582 raw reads. The number of reads from each sample ranged from a minimum of 17,712,314 to a maximum of 32,842,601, and GC content varied between 43 and 54% (Table 1). The fraction of reads mapping on the cocoa genome encompassed between a minimum of 0.26% (FOR 144 h) and a maximum of 15.03% (FOR 0 h).

Except for the beginning and the end of the fermentation, taking into account the relative abundance of bacteria and fungi, both cocoa groups exhibited the same behavior of richness index variation over time (Fig. 2A), and a similar pattern was detected in the Shannon diversity index all over the fermentation time (Fig. 2B). During the first 48 h, the variation of the evenness index was divergent between the two cocoa groups (Fig. 2C). Furthermore, a very sharp decrease in evenness indexes from 48 to 72 h, which remain constantly low until the end of the process (Fig. 2C), indicating that the community of microorganisms exhibited an ever-increasing dominance of few taxa, especially in the period between 72 and 144 h.

Hanseniaspora opuntiae and *Acetobacter pasteurianus* were the most abundant yeast and bacterial species in cocoa fermentation in both FOR and MIX fine cocoa fermentation (Fig. 1). From a total of 340 of fungal OTUs, 57 of them were identified at genus level, and only two species showed more than 1% fungal relative abundance. *Hanseniaspora opuntiae* was the most relatively abundant species with 98%, followed by *H. guilliermondii* with 1.3% for both cocoa sample groups (Fig. 1A). From a total of 488 bacterial OTUs number, 102 OTUs were identified at genus level, and a total of 21 bacterial genera showed more than 1% relative

abundance in, at least, one sample (Fig. 1B).

3.2. Fungi: genus and species levels

Saccharomyces, *Hanseniaspora* (anamorph *Kloeckera*), and *Pichia* are the most commonly cited genera in cocoa fermentation (Maura et al., 2016; Arana-sánchez et al., 2015; Schwan and Fleet, 2014). At the present work, *Hanseniaspora* was by far the dominant yeast taxa during all the fermentation including 0 h time. At the initial period of cocoa fermentation, shortly after the pod break (period 0 h), community richness was associated with yeasts: *Hanseniaspora*, *Nakazawaea*, *Pichia*, *Candida*, *Wickerhamomyces*, *Millerozyma*, *Eremothecium* and *Starmerella*, as well as some multicellular saprotrophic fungi, such as *Rhizopus* and *Lichtheimia*, and even some pathotrophs, such as *Fusarium* and *Phyllosticta* (Fig. 1A). Some of these fungi have never been reported at the beginning of cocoa fermentations. Other mycelial fungi were also found, such as *Aspergillus*, *Penicillium*, and *Trichoderma* (Hanada et al., 2010). Among the filamentous fungi, *Rizophorus* was more abundant in FOR samples from 0 to 48 h but did not appear before the end of fermentation.

The performance of yeasts in the first 24 h of fermentation is related to the production of ethanol, carbon dioxide, organic acids, volatile compounds, as well as with the simultaneous increase of temperature (Lima et al., 2011). Additionally, different yeast species are also capable of producing pectinolytic enzymes, reducing pulp viscosity, increasing oxygen availability, and providing supplementary carbohydrate sources for subsequent microorganisms (Gálvez et al., 2007; Schwan et al., 1997). The ability to depolymerize pectin during the initial fermentation stages of cocoa is important for proper aeration of the fermented cocoa beans and also for the subsequent growth of aerobic AAB (Roelofsen, 1953; Schwan et al., 1997; Silva et al., 2005).

The production of volatile compounds in beans, associated with the taste of chocolate, is directly related to the metabolism of yeasts in pulp, such as those of the genus *Pichia* (Koffi et al., 2017; Pereira et al., 2017). The amount and type of volatile compounds are the most important indicators to evaluate the quality of the cocoa beans (Krähmer et al., 2015). In our work, *Pichia* were present in a very low abundance during the fine cocoa fermentation process, except for FOR at 120 h and, especially, at 144 h (Fig. 1A).

The highest abundance of yeasts was observed in the first 24 h (data not shown), and the highest ethanol concentration in beans of FOR and MIX was detected at 48 h (Supplementary File - Fig. S8). This later (48 h) and lower detections of ethanol in beans than usually reported in cocoa pulp can be caused by its fast consumption, especially by AAB, and also due to the time that ethanol lasts to pass through testa and to penetrate into the beans.

Hanseniaspora is associated with intense metabolism of sugars, generating ethanol and exhibiting pectinolytic activity (Gálvez et al., 2007). Additionally, this genus has a peculiar metabolic diversity, with the capacity to assimilate trehalose and D-gluconate (Jindamorakot et al., 2009), conferring to it an evolutionary and ecological advantage. The high relative abundance of *Hanseniaspora* has already been recorded in distinct cocoa fermentations around the world (Arana-sánchez et al., 2015; Crafaack et al., 2013; Hamdouche et al., 2015; Illegheems et al., 2012). *Hanseniaspora opuntiae* was the yeast species with the highest relative abundance in our study (Fig. 2; Supplementary File - Fig. S2 A and B), and it has already been identified in cocoa fermentations (Crafaack et al., 2013; Daniel et al., 2009; Hamdouche et al., 2015; Illegheems et al., 2012; Miguel et al., 2017; Papalexandratou and De Vuyst, 2011; Papalexandratou et al., 2011; Schwenninger et al., 2016; Visintin et al., 2016).

The genus *Saccharomyces*, which is usually found in high relative frequency in studies related to cocoa fermentation (Daniel et al., 2009; Visintin et al., 2016; Papalexandratou et al., 2011; Pereira et al., 2013b; Schwan and Fleet, 2014), was retrieved with low relative abundance (Fig. 1A). Nonetheless, the aforementioned studies are

Table 1
Summary statistics of the metagenomic samples.

| Sample group | Fermentation period (hours) | Total no. of sequences | %GC |
|--------------|-----------------------------|------------------------|-----|
| FOR | 0 | 20,267,169 | 43 |
| | 24 | 21,994,382 | 47 |
| | 48 | 17,712,314 | 48 |
| | 72 | 23,772,429 | 52 |
| | 96 | 18,676,621 | 54 |
| | 120 | 32,842,601 | 52 |
| | 144 | 26,953,770 | 52 |
| MIX | 0 | 22,050,404 | 47 |
| | 24 | 19,431,365 | 52 |
| | 48 | 21,527,855 | 48 |
| | 72 | 22,866,047 | 52 |
| | 96 | 21,311,073 | 50 |
| | 120 | 22,898,538 | 51 |
| | 144 | 21,917,014 | 51 |

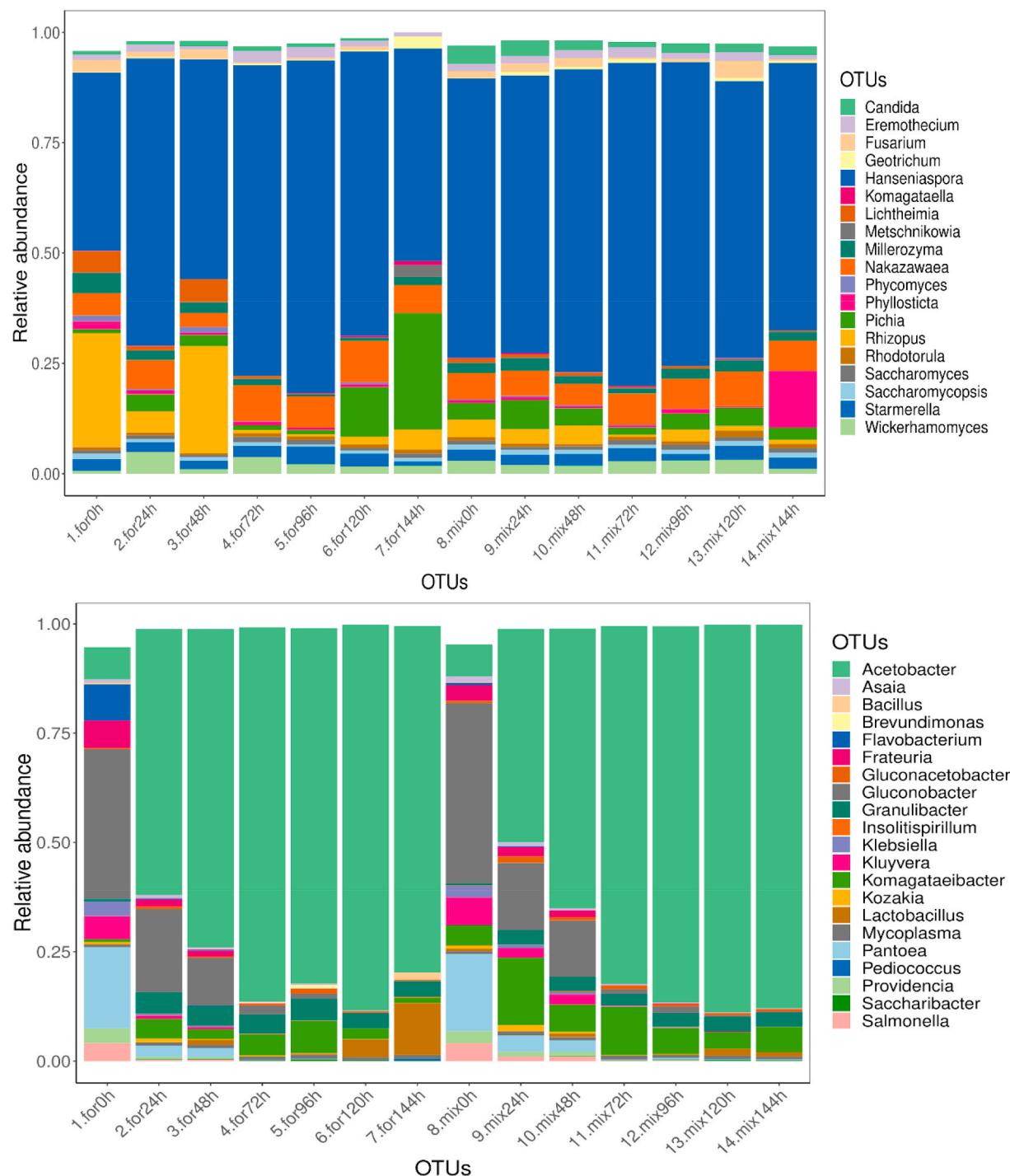


Fig. 1. Relative abundance of OTUs at fungal (A) and bacteria (B) genus level, from FOR and MIX cocoa fermentation, which showed more than 1% relative abundance among Fungi and Bacteria, respectively, in each fermentation stage (0–144 h).

culture-dependent, and the genus *Saccharomyces* is composed of easily cultivated species, which might have influenced the results. Investigations conducted by culture-independent methods pointed out that *Saccharomyces* can occur with high (Papalexandratou et al., 2019) or low (Papalexandratou and De Vuyst, 2011) relative abundances.

In a study in which both culture-dependent and culture-independent methods were used for the same samples, there was a difference between the relative abundance of *Saccharomyces* and *Hanseniaspora*, respectively (Arana-sá'ñchez et al., 2015; Pereira et al., 2013b). In another study, performed by culture-independent methods (PCR-DGGE), *Saccharomyces* and *Hanseniaspora* were identified as the most abundant,

but the authors reported the importance of further studies to overcome the limitation of PCR-DGGE technique with selected 18 S rDNA primers (Pereira et al., 2013a).

More recently, *Hanseniaspora uvarum/opuntiae*, *Saccharomyces cerevisiae*, and *Pichia kudriavzevii* were cited as the most abundant yeast species determined by high-throughput sequencing of ITS2 amplicons of different single-variety fermentations of fine cacao from Nicaragua (Papalexandratou et al., 2019).

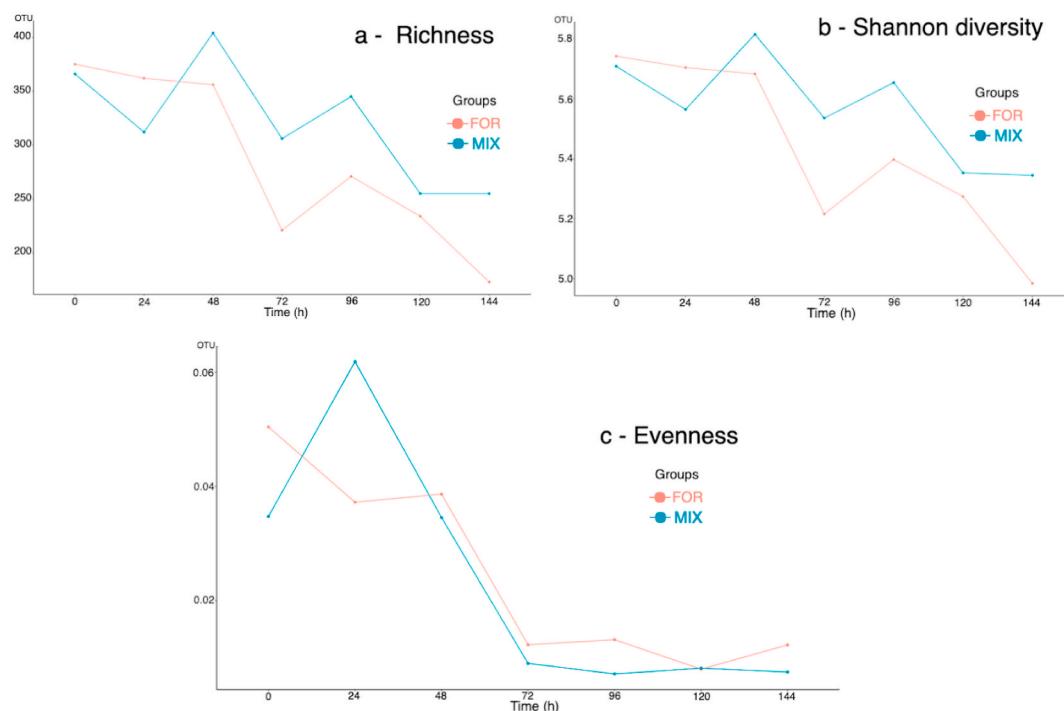


Fig. 2. Diversity analysis of fungal and bacterial OTUs at genus level from FOR and MIX fine cocoa fermentation sample groups in each fermentation stage. (blue line: FOR) (red line: MIX). (A) Richness. (B) Shannon Diversity. (C) Evenness.

3.3. Bacteria: genus and species levels

Acetobacter was the dominant bacterial genus and had the most remarkable increase in relative abundance throughout the fermentation process. *Gluconobacter*, *Pantoea*, *Frateuria*, *Kluyvera* and many others, including some potential pathogenic genera of bacteria (such as *Klebsiella*, *Providencia*, and *Salmonella*, for instance) showed the opposite pattern since they exhibited a decrease of their relative abundances during the fermentation process (Fig. 1).

Two species of *Acetobacter* showed more than 1% of bacterial relative abundance. *Acetobacter pasteurianus* was by far the dominant bacterial species (77% and 81% of relative abundance), followed by *A. ghanensis* (12% and 10%), *A. senegalensis* (3.3% and 1.6%), *A. tropicalis* (1.9% and 1.1%), and *A. ascendens* (1.9% and 2%), for FOR and MIX, respectively (Supplementary Fig. S2).

In our study the high dominance of the family *Acetobacteraceae*, especially *Acetobacter* and *Gluconobacter* (Fig. 1A), is directly related to the ethanol consumption (Supplementary File - Fig. S8), which is produced at the beginning by the fermentation of the pulp, reaching its highest concentration in the pulp between the period of 12 h and 48 h and migrating to beans afterwards (Lefeber et al., 2011; Pereira et al., 2013a). The highest ethanol concentration in beans was observed at 48 h (Supplementary File - Fig. S8).

Species of *Gluconobacter* have already been identified as those that most oxidize ethanol, producing acetic acid, during cocoa fermentation (Pereira et al., 2012). At 0 h, *Gluconobacter* was the most abundant genus of *Acetobacteraceae* and until 48 h showed a striking presence in both fermentations.

The production and, consequently, the consumption of ethanol reflected the production of acetic acid by AAB of both FOR and MIX sample groups in beans (Supplementary File - Figs. S6–A). The period of the highest absorption of acetic acid by the beans was 72 h after the beginning of the fermentation, reaching values similar to those reported in other studies (Ho et al., 2015; Oliveira et al., 2011; Pereira et al., 2013a). The presence of acetic acid in the beans contributes to the death of the embryo and decreases the pH values, which favors the release of

flavor precursors (Craack et al., 2013). Nonetheless, acidity is desirable for flavor development but not for final cocoa products. Thus, it is necessary that these acids are consumed or evaporated. The decrease in acetic acid content occurred from 72 h in the FOR sample group while a constant concentration was detected in the MIX sample group over the fermentation process (Supplementary File - Figs. S6–A). This can be associated with a higher relative abundance of *Acetobacteraceae* in the MIX compared to FOR group (Supplementary File - Fig. S2). Nevertheless, this cannot be considered a problem because, in the drying period (data not shown), the acetic acid concentration decreased in both varieties.

The dominance of *Acetobacter* in cocoa fermentation has already been documented in different regions and by different methods (Camu et al., 2008; Craack et al., 2013; Gabaza et al., 2019; Garcia-armisen et al., 2010; Hamdouche et al., 2015; Moreira et al., 2017; Papalexandratou et al., 2011c; Pereira et al., 2012, 2013abib_Pereira_et_al_2012bib_Pereira_et_al_2013a; Ramos et al., 2014; Schwenninger et al., 2016; Visintin et al., 2016). These data suggest that *Acetobacter* is a core bacterial genus in cocoa bean fermentation worldwide.

In our work, *Acetobacter pasteurianus* was the most abundant among five species identified with relative abundance above 1% (Supplementary File - Fig. S2). Moreover, *A. pasteurianus* has already been identified in other studies on cocoa fermentation (Ardhana and Fleet, 2003; Bortolini et al., 2016; Camu et al., 2007, 2008; Craack et al., 2013; Garcia-armisen et al., 2010, 2011; Hamdouche et al., 2015; Ho et al., 2014; Illegems et al., 2012; Menezes et al., 2016; Schwenninger et al., 2016).

AAB are known as relatively slow-growing bacteria due to their oxidation of sugars, alcohols, and polyols into their corresponding organic acids, aldehydes, or ketones from which they obtain energy (De Ley et al., 1984; Taban and Saichana, 2017). In spontaneous cocoa fermentation, they are one of the dominant bacterial groups, which means that this environment offers both abiotic and biotic conditions for their successful growth, as well as the adaptation of some AAB species to diverse media of cocoa fermentation mass.

Based on both culture-dependent and culture-independent studies, cocoa fermentation is especially associated with two bacterial groups,

LAB and AAB, which apparently constitute two distinct seral communities in this ecological succession (Bortolini et al., 2016; Moreira et al., 2013; Hamdouche et al., 2015; Papalexandratos et al., 2011, 2013; Visintin et al., 2011). Our results, based on fine cocoa fermentation, also recovered a high number of AAB OTUs, but a shiny relative abundance of *Lactobacillus* was detected only in the final periods (120 and 144 h) in FOR, and even lower in MIX group (Fig. 1B). Since the studied samples were of cocoa beans, especially worked under GMP conditions, differences in cultivation and processing probably have profoundly affected the typical microbial profile in cocoa fermentation.

Additionally, our study showed that the absence of a high relative abundance of LAB did not avoid the presence of lactic acid in fine cocoa beans (Supplementary File - Figs. S6–B). Other studies, which also evaluated lactic acid concentration, pointed out that, in the pulp, its concentration decreases along the fermentation while, in beans, it remains constant until the end of the process (Lefever et al., 2011; Pereira et al., 2013a). Succinic acid may also impair the quality of beans and chocolate making them more acidic and has already been identified as a product of LAB metabolism (Quattara et al., 2016; Schwan and Wheals, 2004). In our study, the succinic acid concentration showed no strong changes during the entire experiment (0–144 h) (Supplementary File - Figs. S6–B). Nonetheless, it was one of the physicochemical variables that contribute to the separation of the groups (A, B, C, D) in the ordination analysis (Fig. 3). Altogether, our results revealed that LAB were not necessary to fine cocoa spontaneous fermentation and were similar to some studies that detected the same situation for inoculated *in vitro* and controlled cocoa fermentation (Ho et al., 2015; John et al., 2019). Inocula of yeasts, LAB, and AAB were tested jointly and separately for *in vitro* cocoa fermentation, and both bacterial groups (LAB and AAB) were reported as not essential for cocoa bean fermentation (Ho et al., 2018).

Enterobacteriales, as already showed by other studies (Camu et al., 2008; Garcia-armisen et al., 2010; Hamdouche et al., 2015; Illeghemps et al., 2012; Lefever et al., 2011; Papalexandratos et al., 2011, 2019; Papalexandratos and De Vuyst, 2011; Pereira et al., 2013a), exhibited a

high relative frequency at 0 h (Fig. 1B). *Enterobacteriaceae* are associated with methylglyoxal detoxification and mixed-acid fermentation pathways, which involve the production of lactate, acetate, succinate and ethanol, and also pectinolysis and citrate assimilation. Among them, three genera have already been identified as being able to perform mixed-acid fermentation pathway in cocoa fermentation: *Pantoea*, *Serratia*, and *Erwinia* (Illeghemps et al., 2015). Only *Pantoea* was detected in our research in a high bacterial relative abundance at 0 h, and it decreased during fermentation time (Fig. 1B), as most of the bacteria species detected in our study.

3.4. Physicochemical analysis

The temperature variation displayed a similar behavior for FOR and MIX cocoa fermentations. The highest increase in temperature occurred in the first 24 h (10–15 °C). Then, a decelerated increase in temperature was observed during 24–96 h. The temperature at the beginning of the fermentation was 26 °C for both sample groups and reached maximum values of 49 °C in MIX and 46 °C in FOR at 96 h, respectively (Supplementary File - Fig. S3).

Initial pH (6.2) and also its variation exhibited similar patterns for both FOR and MIX fermentations. There was a strong decrease in the first 48 h (4.4–4.5) and then a low variation was detected until 96 h. At 144 h pH values reached 5.1 and 4.6, respectively (Supplementary Fig. S4).

Total and reducing sugars showed a similar variation pattern along the fermentation process in both sample groups. In MIX, there was a high decrease in the first 24 h (total sugar: 72.96 mg g⁻¹; reducing sugars: 18.6 mg g⁻¹), the lowest concentrations during fermentation. The highest concentrations for total and reducing sugars inside the beans were at the final stages (92.56 mg g⁻¹, 120 h, and 42.37 mg g⁻¹, 144 h) for FOR and MIX, respectively), as reported by Moreira et al. (2013). In FOR samples, there was a constant decrease of total sugars up to 96 h (82.83 mg g⁻¹), and an increase in the concentration of the reducing

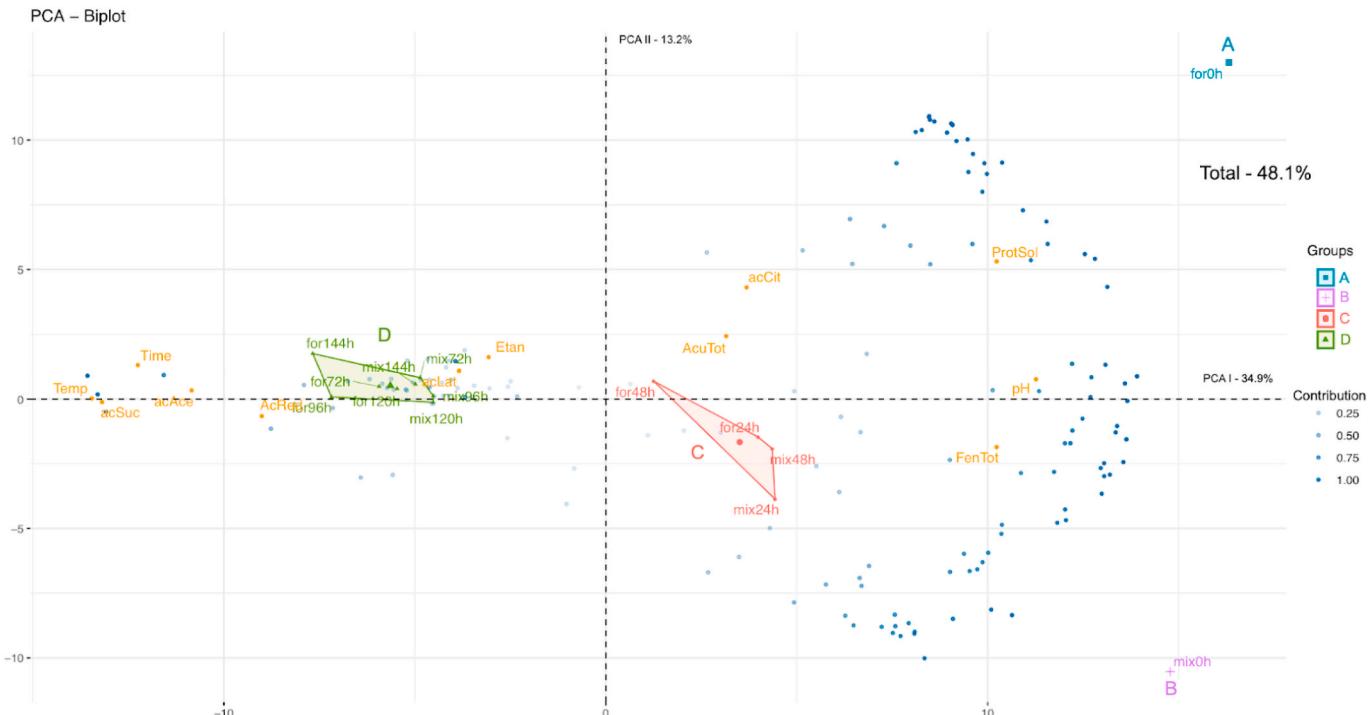


Fig. 3. Principal Component Analysis (PCA) of bacterial and fungal OTUs at genera level and physical-chemical variables of FOR and MIX fine cocoa fermentation in different times (0, 24, 48, 72, 120 and 144 h). Etan = ethanol; FenTot = total phenol; acAc = acetic acid; acCit = Citric acid; acLat = lactic acid; acSuc = succinic acid; AcTot = total sugars; AcRed = reducing sugars; ProtSol = soluble proteins; Temp = Temperature. A (blue), B (purple), C (red) and D (green) represent the groups formed by ordination analysis.

sugars in the first 72 h (37.77 mg g^{-1}) (Supplementary Fig. S5).

The beginning of the fermentation process (0 h) was marked by a higher concentration of citric acid in the beans (FOR: 8.39 mg g^{-1} ; MIX: 6.30 mg g^{-1}), with a slight decrease over fermentation time (Supplementary Figs. S6–A). The patterns of variation of acetic acid concentration in cocoa beans of both sample groups were remarkably similar, except for the end of the fermentation. There was a continuous and very sharp increase, reaching a maximum value at 72 h (FOR: 27.36 mg g^{-1} ; MIX: 23.21 mg g^{-1}), followed by a continuous decrease until the end of fermentation in FOR (8.37 mg g^{-1} at 144 h) and fluctuations in MIX, with a final concentration rather higher (20.49 mg g^{-1} at 144 h) than that of FOR sample group (Supplementary Figs. S6–A).

A slight increase in lactic acid concentration in the beans was observed for FOR sample group (3.9 mg g^{-1}) at 48 h and even lower for MIX (1.6 mg g^{-1}) at 96 h (Supplementary Figs. S6–B). The pattern of variation of succinic acid concentration was very low and similar for FOR and MIX during the fermentation process (Supplementary

Figs. S6–B).

The pattern of variation of total phenols concentration in the beans similarly decreased in both cocoa sample groups along the fermentation. At the first 48 h occurred the main phenol concentration decrease, followed by a slight increase at 48–72 h, and a subsequent decrease until the end of fermentation (Supplementary Fig. 7). Variation of soluble protein concentration in the beans also displayed very similar behavior in both sample groups along the fermentation process: a sharp decrease during the first 24 h (minimum values along all the process), followed by a small fluctuation until the end of fermentation (Supplementary Fig. S7).

The behavior of the variation of ethanol concentration in the beans was also very similar along most of the fermentation (0–96 h) for both sample groups. FOR and MIX exhibited two peaks of ethanol concentration (first at 48 h for both and second one at the end of the process, 120 h and 144 h, respectively). The strongest increase occurred at the beginning of the process culminating in the highest values detected at

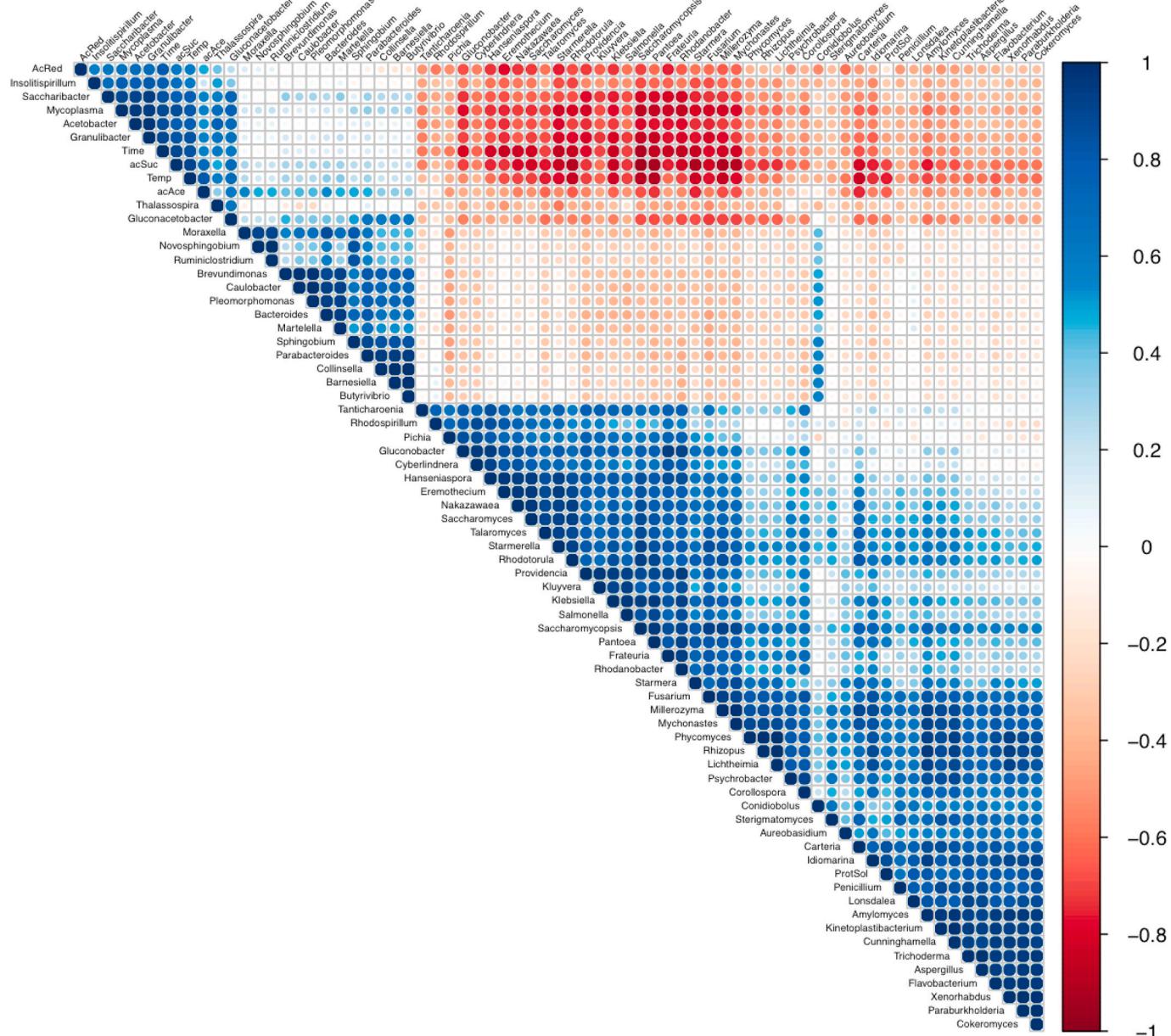


Fig. 4. Correlation analysis of bacterial and fungal OTUs at genus level and physicochemical variables in each stage from FOR and MIX fine cocoa fermentation sample groups. AcRed = reducing sugars; acSuc = succinic acid; Temp = temperature; acAc = acetic acid; ProtSol = soluble proteins.

48 h (FOR: 3.58 mg g⁻¹, MIX: 1.84 mg g⁻¹) (Supplementary Fig. S8).

3.5. Integrative analyses: ordination, cluster, and correlation analyses

Despite the pods being collected from cocoa trees cultivated in the same farm and under the same practices of cultivation and harvesting, the microbial communities were rather different between the two sample groups (FOR and MIX) at 0 h (Fig. 3).

Ordination analysis (PCA) of microbial communities and physicochemical features of the beans in FOR and MIX sample groups clearly separated the 14 samples (2 cocoa groups x 7 times) in four distinct groups: A, B, C, and D (Fig. 3). Groups A and B, which comprised FOR 0 h and MIX 0 h, respectively, are considerably separated from each other and also from the groups C and D. A total of 87 distinct genera of microorganisms were identified in FOR 0 h (A) and 91 in MIX 0 h, 12 and 18 genera were, respectively, exclusive of each one (data not shown). All the samples of 24–48 h and 72–114 h of both FOR and MIX formed another two well-defined and separated groups: C and D, respectively (Fig. 3).

The physicochemical features of the beans that most influenced the separation of these well-defined groups were temperature, succinic acid, pH, acetic acid, total phenols, and soluble proteins (Fig. 3). A total of 54 genera of microorganisms including the strongly correlated taxa *Hanseniaspora*, *Starmerella*, *Nakazawaea*, *Saccharomyces* (Fig. 4) were associated with group C (24–48 h). Group D (72–144 h) was mostly associated with the correlated genera *Acetobacter*, *Saccharibacter*, *Granulibacter*, *Thalassospira*, *Mycoplasma* and *Paracoccus*. Taking into account the genera and relative abundance of the microorganisms of the groups C and D, a strong negative correlation between them was detected (Fig. 4).

In the first 24 h of fermentation for both FOR and MIX, there was an alteration in the bacterial diversity, and the genus *Acetobacter* became the most abundant, and remained as such until the end of the fermentations (144 h) (Fig. 1). *Hanseniaspora* was the most relatively abundant yeast genus during all the fermentation in pulp (0–144 h) in both FOR and MIX groups (Fig. 1).

As already well-described in the specialized literature on cocoa fermentation, for FOR and MIX sample groups, an increase in the temperature of the cocoa mass was detected during the collection of samples of fine cocoa bean fermentation (Supplementary Fig. S3). In the laboratory, a decrease in pH (Supplementary Fig. S4) was detected, as well as a sharp increase in the concentration of acetic acid (Supplementary Figs. S6–A) in beans. The concentrations of citric and lactic acids (Supplementary Figs. S6–A) apparently did not undergo major changes during the fermentation process in the beans of both sample groups. However, succinic acid, in spite of a low concentration in the beans of FOR and MIX sample groups, had a strong influence at ordination (PCA) analysis (Fig. 3). The concentration of ethanol in beans increased considerably in the first 48 h and then declined over time (Supplementary Fig. S8).

The increase in temperature, decrease in pH, and the presence of acids and alcohols that diffuse into the beans played a role in the disintegration of the cell compartments of the cotyledons, and the release of enzymes derived from the beans themselves with subsequent degradation of soluble proteins and phenols (Supplementary Fig. S7) and total sugars, as well as an increase in reducing sugars (Supplementary Fig. S5). Furthermore, as previously described, the increase of the temperature (Supplementary Fig. S3), and decrease of pH (Supplementary Fig. S4), establishes warm and acidic environment necessary to initiate the protein proteolysis in cocoa beans and formation of flavor precursors (Aculey et al., 2010; Biehl et al., 1969; Crafack et al., 2013; Schwan and Wheals, 2004). According to the ordination (PCA) analysis, temperature and pH, succinic and acetic acids, total phenol and soluble proteins were the physicochemical variables that most contributed to the separation of A, B, C and D groups in our study (Fig. 3). Correlation analysis (Fig. 4) showed that, in general, reducing sugars, acetic and

succinic acids, and temperature were the physicochemical variables that more negatively correlated to microbial occurrence and abundance.

Ordination analysis (PCA) clear separation of FOR 0 h (group A) and MIX 0 h (group B) from the other groups (groups C and D) at initial fermentation process (0 h) (Fig. 3) was mainly based on taxon diversity (composition, richness, and abundance) and reflected the plant-associated microbiota, as well as the practices during harvesting and handling of pods. A total of 54 distinct bacterial and fungal genera strongly contributed to the marked separation of the groups A (FOR 0 h) and B (MIX 0 h), and many of them have already been reported in the fermentation of cocoa, performing functions related to ecological succession (Supplementary File – Table S11).

Regarding correlation analysis (Fig. 4), an interesting negative correlation occurs between two AAB genera, *Acetobacter* and *Gluconobacter*. At the beginning of the fermentation process (0 h), *Gluconobacter* had a higher relative abundance than *Acetobacter*, but at 24 h a marked change was detected, and *Acetobacter* relative abundance was higher than *Gluconobacter* (Fig. 1A). It might have occurred because cocoa pulp fermentable sugars (10–15% of glucose, fructose, and sucrose) (Pettipher, 1986) are more preferred as a carbon source by *Gluconobacter* than by *Acetobacter*. *Gluconobacter* species can obtain energy more efficiently by the metabolism of the sugars via pentose phosphate pathway (De Ley et al., 1984). Moreover, *Gluconobacter* strains have the ability to perform oxidative fermentation of various sugars, sugar alcohols, and sugar acids, leading to the formation of several valuable products (Sai-chana et al., 2015). Correlation analysis exhibited a negative correlation between *Gluconobacter* and the mainly physicochemical variables and *Acetobacter*, principally and probably because *Gluconobacter* is a better user of sugars than *Acetobacter*. *Gluconobacter* is also able to produce acids from D-glucose and D-fructose (De Ley et al., 1984) actively changing and facilitating the environment to the next major bacterial group, *Acetobacter*. In accordance with correlation analysis (Fig. 4), the mainly physicochemical variables that were significantly correlated to *Acetobacter* were temperature and succinic acid concentration (both increased along fermentation time) and reducing sugar (decreased with the fermentation time), which reinforces that the principal substrate for *Acetobacter* was ethanol produced by *Hanseniaspora*.

Hanseniaspora was negatively correlated to reducing sugar, which decreased along the fermentation time (Fig. 4) because they also consumed the pulp sugars and competed with *Gluconobacter* and other microorganisms for such niche. Amongst the yeasts, *Pichia*, *Eremothecium*, *Nakazawaea*, *Saccharomyces*, *Starmerella*, *Saccharomyopsis* were some of those that also were positively correlated with *Hanseniaspora*. Besides, they could be producing pectinolytic and other enzymes for proper aeration of the fermented cocoa beans, and for the subsequent growth of aerobic AAB (Gálvez et al., 2007; Roelofsen, 1953; Schwan et al., 1997; Schwan et al., 1997; Silva et al., 2005). The aerobic phase is characterized by oxidative and condensation reactions such as oxidation of protein-polyphenol complexes and carbonyl-amino condensation that reduce astringency (Afoakwa et al., 2008).

Regarding the cluster analysis based on the heatmap (occurrence and abundance) for *Bacteria* (Fig. 5a), the samples were separated in two main clusters (I and II). The first cluster (I) comprised two subclusters (IA and IB). Subcluster IA was formed by FOR and MIX samples at 0 h (groups A and B of Fig. 3), with a very low relative abundance of *Acetobacter*. Subcluster IB (group C of Fig. 3) displayed a much higher relative abundance of *Acetobacter* when compared to subcluster IA (Fig. 5a and Supplementary File - Figure S2 B). The second cluster (II) was composed of the samples with a very high relative abundance of *Acetobacter* (Fig. 5a), which is equivalent to group D of Fig. 4.

Based on the heat map analysis (occurrence and abundance) for *Fungi*, one main cluster (IC) was retrieved (Fig. 5b), and it comprised all the samples of both FOR and MIX from 0 to 120 h. This large group was characterized by the high relative abundance of *Hanseniaspora* (Fig. 5b and Supplementary File - Fig. S2). Inside this main cluster, there was a subcluster (IC1) uniting the samples of FOR at 0 h and 48 h, which were

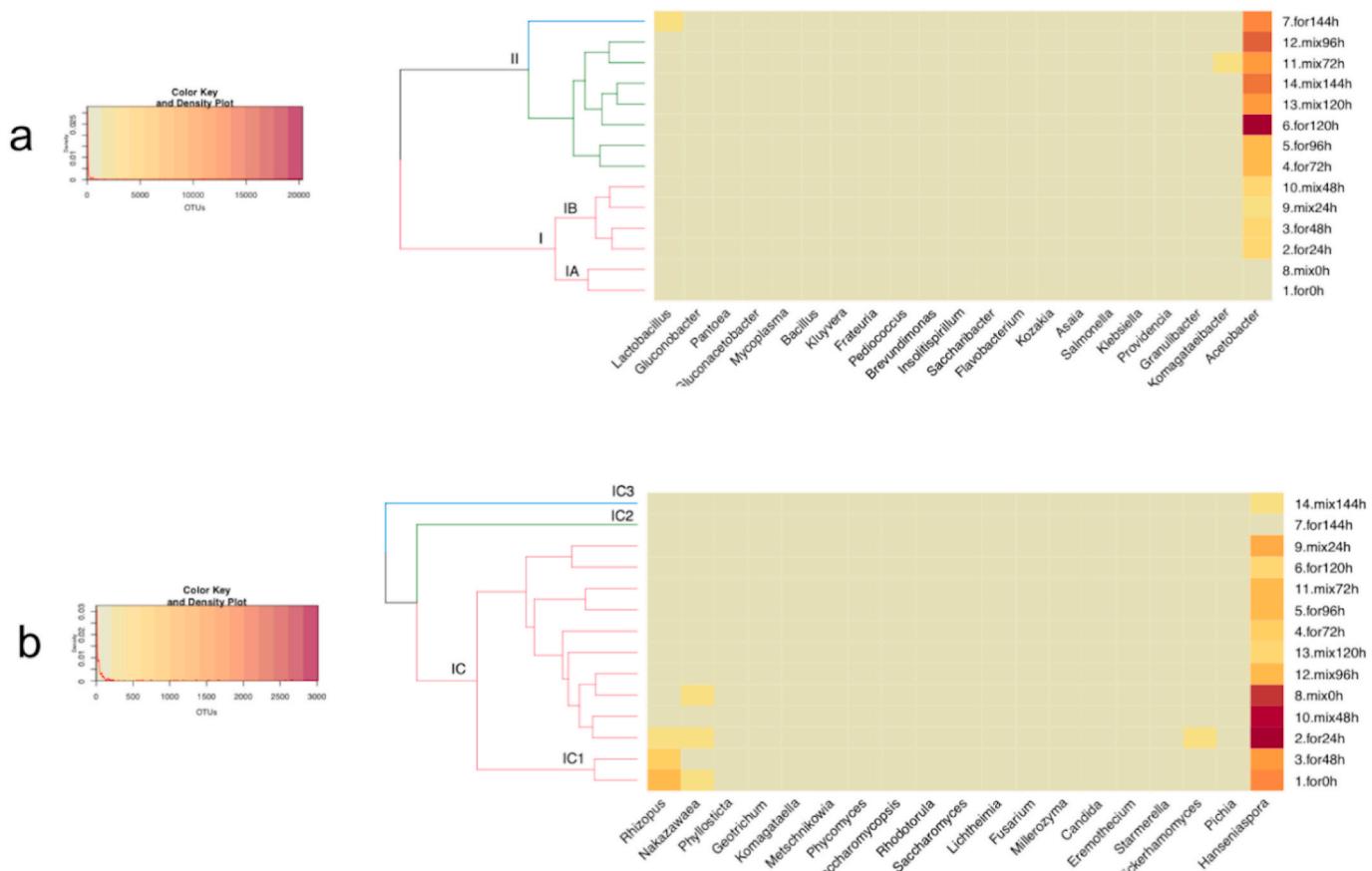


Fig. 5. Heatmap analysis of microbial OTUs from FOR and MIX fine cocoa fermentation sample groups. **(a)** The most abundant bacterial genera; **(b)** the most abundant fungal genera. IA, IB: bacterial subclusters. IC: fungal cluster.

characterized by a higher relative abundance of *Rhizopus* than all the other samples.

Moreover, there were two isolated subclusters, IC2 and IC3 (Fig. 5), one composed by the sample of FOR at 144 h, and the other consisting of the sample of MIX at 144 h, both characterized by the low relative abundance of *Hanseniaspora*, with the former comprising the lowest abundance of this fungal genus (Fig. 5b).

The initial populations of the different microbial groups vary from place to place and growth kinetics depending on the composition of the fresh pulp, as well as the practices applied during fermentation (Nielsen et al., 2013; Papalexandratos and Nielsen, 2016). Besides, small methodological changes can cause impact on cocoa microbial fermentation. Mota-Gutierrez et al. (2018), studying Forastero variety cocoa fermentations by 16 S and ITS amplicon metagenomics, pointed out that the relative abundances of several bacterial and yeast species along the fermentations were influenced if they were carried out in boxes or in heaps.

During all the fermentation processes, some microbial groups are recurrently cited (yeasts, LAB, AAB, and Enterobacteriaceae), but the late and low relative abundance of LAB only in FOR may mean that this bacterial group is not important to fine cocoa fermentation. Nonetheless, even at genus level, a considerable variation in the most frequent microorganisms was detected and, therefore, biotic and abiotic factors of cocoa cultivation and fermentation processes must be taken into account. More recently, correlations matrices of oligotypes abundance of the overall fermentations of two Colombian farms showed a clear clustering by farm, suggesting that farm protocols generate a unique fingerprint in the dynamics and interactions of the microbial communities (Pacheco-Montealegre et al., 2020).

Our findings showed that microbial fermentation profile of FOR and MIX fine cocoa beans, produced in the same farm and under the same GMP, showed taxonomic composition, richness, relative abundance, and evenness of microbial communities in cocoa fermentation quite similar, but taxonomic composition was different when compared to other studies.

As the market for fine chocolate is increasing and seeking for new organoleptic features, we suggest a new perspective on inoculum for cocoa fermentation aiming at fine chocolate production. Following this suggestion, microbial cocoa fermentation inoculants should be composed of the main profiles of the microbial composition also considering the region/farm where fermentation is performed. We suggest that a starter inoculum composed of *Hanseniaspora opuntiae* and *Acetobacter pasteurianus* strains (the fungal and bacterial dominant species in both cocoa sample groups) should be tested in fermentation process for FOR and MIX cocoa beans.

The next steps in our research program will comprise the evaluation of the suggested starter inoculum, (composed of *Hanseniaspora opuntiae* and *Acetobacter pasteurianus* strains), as well as the *in silico* metabolic reconstruction of these microbiomes.

Declaration of competing interest

None.

Acknowledgments

The authors would like to thank all the colleagues that contributed directly or indirectly to this work: the Graduate Program in

Biotechnology of the State University of Feira de Santana (UEFS), the Graduate Programs in Microbiology and in Bioinformatics of Federal University of Minas Gerais (UFMG), and, especially Raimundo Mororó, at Riachuelo agroindustry of Mendoá Chocolates. We are also grateful to the Foundation of support for Research in the State of Minas Gerais (FAPEMIG) and Foundation of support for Research in the State of Bahia (FAPESB). This work was funded by the Coordination of Superior Level Staff Improvement (CAPES, Grant PROCAD 88881.068458/2014–01). The funder had no role in study design, data collection, analysis, and decision to publish or prepare the manuscript. AG-N receives a research grant for productivity from the National Council for Scientific and Technological Development (CNPq), Brazil (no. 310764/2016–5) and APTU receives a senior postdoc grant from CNPq (no. 104327/2019–7).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jenvman.2020.111198>.

References

- Aculey, P.C., Snitkjaer, P., Owusu, M., Bassompierre, M., Takrama, J., Norgaard, Lars N., Petersen, M.A., Nielsen, D.S., 2010. Ghanaian cocoa bean fermentation characterized by spectroscopic and chromatographic methods and chemometrics. *Journal of Food Science* 75, 300–307.
- Afoakwa, E.O., Paterson, A., Fowler, M., Ryan, A., 2008. Flavor formation and character in cocoa and chocolate: a critical review. *Crit. Rev. Food Sci. Nutr.* 48, 840–857.
- Andrews, S., 2010. FastQC A quality control tool for high throughput sequence data". Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. (Accessed 15 January 2017).
- Arana-sánchez, A., Segura-garcía, L.E., Kirchmayr, M., Orozco-ávila, I., Lugo-cervantes, E., Gschaeider-mathis, A., 2015. Identification of predominant yeasts associated with artisan Mexican cocoa fermentations using culture-dependent and culture-independent approaches. *World J. Microbiol. Biotechnol.* 31, 359–369.
- Ardhana, M.M., Fleet, G.H., 2003. The microbial ecology of cocoa bean fermentations in Indonesia. *Int. J. Food Microbiol.* 8, 87–99.
- Biehl, E.R., Nieh, E., Hsu, K.C., 1969. Substituent effects on the reactivity of arynes. Product distributions as an index of relative reactivities of arynes in methylamine and dimethylamine solvents. *J. Organ. Chem.* 34 (11), 3595–3599.
- Bortolini, C., Patroni, V., Puglisi, E., Morelli, L., 2016. Detailed analyses of the bacterial populations in processed cocoa beans of different geographic origin, subject to varied fermentation conditions. *Int. J. Food Microbiol.* 236, 98–106.
- Bradford, M.M., 1976. A rapid sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254.
- Camu, N., Winter, T., Verbrugge, K., Cleenwerck, I., Vandamme, P., Takrama, J.S., Vancanneyt, M., De Vuyst, L., 2007. Dynamics and biodiversity of populations of lactic acid bacteria and acetic acid bacteria involved in spontaneous heap fermentation of cocoa beans in Ghana. *Appl. Environ. Microbiol.* 73, 1809–1824.
- Camu, N., Gonzalez, A., De Winter, T., Van Schoor, A., De Bruyne, K., Vandamme, P., Takrama, J.S., De Vuyst, L., 2008. Influence of turning and environmental contamination on the dynamics of lactic acid bacteria and acetic acid bacteria populations involved in spontaneous cocoa bean heap fermentation in Ghana. *Appl. Environ. Microbiol.* 74, 86–98.
- Chao, A., 1987. Estimating the population size for capture and recapture data with unequal matchability. *Biometrics* 43, 783–791.
- Chao, A., et al., 2013. Rarefaction and extrapolation with hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monograph* 84, 45–67.
- Colwell, R.K., 2009. In: Levin, S.A., et al. (Eds.), *Biodiversity: Concepts, Patterns, and Measurement* in the Princeton Guide to Ecology. *Princeton University Press*, pp. 257–263.
- Cota-sanchez, J.H., Remarchuk, K., Ubayesena, K., 2006. Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Mol. Biol. Rep.* 24, 161–167.
- Crafack, M., Mikkelsen, M.B., Saerens, S., Knudsen, M., Blennow, A., Lowor, S., Takrama, J., Swiegers, J.H., Petersen, G.B., Heimdal, H., Nielsen, D.S., 2013. Influencing cocoa flavor using *Pichia kluveri* and *Kluveromyces marxianus* in a defined mixed starter culture for cocoa fermentation. *Int. J. Food Microbiol.* 167, 103–116.
- Daniel, H., Vrancken, G., Takrama, J.F., Camu, N., Vos, P., De Vuyst, L., 2009. Yeast diversity of Ghanaian cocoa bean heap fermentations. *FEMS Yeast Res.* 9, 774–783.
- Dantas, P.C.C., Pires, M.M., Uetanabaro, A.P.T., Gomes, A.S., Novais, A.C.P., 2020. O mercado de chocolate no sul da Bahia. DRD - Desenvolvimento Regional em debate 10, 54–73. <https://doi.org/10.24302/drdd.v10i0.2373>.
- De Ley, J., Gills, M., Swings, J. (1984). Acetobacteriaceae. In: Krieg, N. R., Holt, J. G. (Eds.), *Bergey's Manual of Systematic Bacteriology*, 9a. Ed., Williams & Wilkins, London. 267 – 278.
- De Vuyst, L., Weckx, S., 2016. The cocoa bean fermentation process: from ecosystem analysis to starter culture development. *J. Appl. Microbiol.* 121, 5–17.
- De Vuyst, L., Lefeber, T., Papalexandratos, Z., Camu, N., 2010. The functional role of lactic acid bacteria in cocoa bean fermentation. *Biotechnology of Lactic Acid Bacteria* 301–325. <https://doi.org/10.1002/9780813820866.ch17>.
- Gabaza, M., Joossens, M., Cnockaert, M., Muchuweti, M., Raes, K., Vandamme, P., 2019. Lactococci dominate the bacterial communities of fermented maize sorghum and millet slurries in Zimbabwe. *Int. J. Food Microbiol.* 289, 77–87.
- Gálvez, S.L., Loiseau, G., Paredes, J.L., Barel, M., Guiraud, J., 2007. Study on the microflora and biochemistry of cocoa fermentation in the Dominican Republic. *Int. J. Food Microbiol.* 114, 124–130.
- Garcia-armisen, T., Papalexandratos, Z., Hendryckx, H., Camu, N., Vrancken, G., De Vuyst, L., Cornelis, P., 2010. Diversity of the total bacterial community associated with Ghanaian and Brazilian cocoa bean fermentation samples as revealed by a 16 S rRNA gene clone library. *Appl. Microbiol. Biotechnol.* 87, 2281–2292.
- Garcia-Armisen, T., Garcia-Armisen, K., Passerat, J., Triest, D., Servais, P., Cornelis, P., 2011. Antimicrobial resistance of heterotrophic bacteria in sewage-contaminated rivers. *Water Res.* 45, 788–796.
- Gardener, M., 2014. Diversity comparing in Community ecology. In: Gardener, M. (Ed.), *Pelagic Publishing* 196–270.
- Hamdouche, Y., Guehi, T., Durand, N., Kedjebo, K.B.D., Montet, D., Meile, J.C., 2015. Dynamics of microbial ecology during cocoa fermentation and drying: towards the identification of molecular markers. *Food Contr.* 48, 117–122.
- Hanada, R.E., Pomella, A.W.V., Costa, H.S., Bezerra, J.L., Loguerio, L.L., Pereira, J.O., 2010. Endophytic fungal diversity in *Theobroma cacao* (cacao) and *T. grandiflorum* (cupuacu) trees and their potential for growth promotion and biocontrol of black-pod disease. *Fungal biology* 114, 901–910.
- Ho, V.T.T., Zhao, J., Fleet, G., 2014. Yeasts are essential for cocoa bean fermentation. *Int. J. Food Microbiol.* 174, 72–87.
- Ho, V.T.T., Zhao, J., Fleet, G., 2015. The effect of lactic acid bacteria on cocoa bean fermentation. *Int. J. Food Microbiol.* 205, 54–67.
- Ho, V.T.T., Bach, L.G., Vo, D.V.N., 2018. Preparation and characterization of advanced PtRu/TiO₂ 7MoO₃ 702 catalysts for direct methanol fuel cells. *Appl. Mech. Mater.* 876, 57–63.
- Hsieh, T.C., Ma, K.H., Chao, A., McInerny, G., 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* 7, 1451–1456.
- IAL – Instituto Adolfo Lutz, 1985. Normas analíticas do Instituto Adolfo Lutz. *Métodos químicos e físicos para análise de alimentos*, 3a. São Paulo, 1.
- Illegems, K., De Vuyst, L., Papalexandratos, Z., Weckx, S., 2012. Phylogenetic analysis of a spontaneous cocoa bean fermentation metagenome reveals new insights into its bacterial and fungal community diversity. *PLoS One* 7, 5.
- Illegems, K., De Vuyst, L., Weckx, S., 2013. Complete genome sequence and comparative analysis of *Acetobacter pasteurianus* 386B, a strain well-adapted to the cocoa bean fermentation ecosystem. *BMC Genom.* 14, 526.
- Illegems, K., Weckx, S., De Vuyst, L., 2015. Applying meta-pathway analyses through metagenomics to identify the functional properties of the major bacterial communities of a single spontaneous cocoa bean fermentation process sample. *Food Microbiol.* 50, 54–63.
- Jindamorakot, S., Ninomiya, S., Limtong, S., Yongmanitchai, W., Tuntirungkij, M., Potcharoen, W., Tanaka, K., Kawasaki, H., Nakase, T., 2009. Three new species of bipolar budding yeasts of the genus *Hanseniaspora* and its anamorph *Kloeckera* isolated in Thailand. *FEMS Yeast Res.* 9, 1327–1337.
- John, U., Lu, Y., Wohlrab, S., Groth, M., Janouškovec, J., Kohli, G.S., Mark, F.C., Bickmeyer, U., Farhat, S., Felder, M., Frickenhaus, S., Guillou, L., Keeling, P.J., Moustafa, A., Porcel, B.M., Valentini, K., Glöckner, G., 2019. An aerobic eukaryotic parasite with functional mitochondria that likely lacks a mitochondrial genome. *Science Advances*. 5 <https://doi.org/10.1126/sciadv.aav1110>. PMID: 31032404.
- Kim, D., Song, L., Breitwieser, F.P., Salzberg, S.L., 2016. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* 26, 1721–1729.
- Koffi, O., Samagaci, L., Goualé, B., Niame, S., 2017. Diversity of yeasts involved in cocoa fermentation of six major cocoa-producing regions in ivory coast. *Eur. Sci. J.* 13, 496–516.
- Krämer, A., Engel, A., Kadow, D., Ali, N., Umaharan, P., Kroh, L.W., Schulz, H., 2015. Fast and neat — determination of biochemical quality parameters in cocoa using near infrared spectroscopy. *Food Chem.* 181, 152–159.
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Lefebre, T., Gobert, W., Vrancken, G., Camu, N., De Vuyst, L., 2011. Dynamics and species diversity of communities of lactic acid bacteria and acetic acid bacteria during spontaneous cocoa bean fermentation in vessels. *Food Microbiol.* 28, 457–464.
- Legendre, P., Legendre, L., 1998. Developments in environmental modelling. *Numerical ecology*. 24.
- Lima, L.J.R., Almeida, M.H., Nout, M.J.R., Zwietering, M.H., 2011. *Theobroma cacao L.*, "the food of the gods": quality determinants of commercial cocoa beans, with particular reference to the impact of fermentation. *Crit. Rev. Food Sci. Nutr.* 51, 731–761.
- Magoč, T., Salzberg, S.L., 2011. FLASH: fast length Adjustment of short reads to improve genome assemblies. *Bioinformatics* 27 (21), 2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>.
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. <https://doi.org/10.14806/ej.17.1.200>.
- Matias Rodrigues, J.F., Schmidt, T.S.B., Tackmann, J., Von Mering, C., 2017. MAPseq: highly efficient k-mer search with confidence estimates, for rRNA sequence analysis. *Bioinformatics* 33, 3808–3810. <https://doi.org/10.1093/bioinformatics/btx517>.

- Maura, D., Hazan, R., Kitao, T., Ballok, A.E., Rahme, L.G., 2016. Evidence for direct control of virulence and defense gene circuits by the *Pseudomonas aeruginosa* quorum sensing regulator, MvfR. *Sci. Rep.* 6, 34083.
- Menezes, A.G.T., Batista, N.N., Ramos, C.L., Silva, A.R.A., Efraim, P., Pinheiro, A.C.M., Schwan, R.F., 2016. Investigation of chocolate produced from four different Brazilian varieties of cocoa (*Theobroma cacao* L.) inoculated with *Saccharomyces cerevisiae*. *Food Research International* 81, 83–90.
- Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7, 11257.
- Miguel, M.G.C.P., Reis, L.V.C., Efraim, P., Santos, C., Lima, N., Schwan, R.F., 2017. Cocoa fermentation: microbial identification by MALDI-TOF MS, and sensory evaluation of produced chocolate. *Food Sci. Technol.* 77, 362–369.
- Miller, G.L., 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31, 426–428.
- Moreira, I.M.V., Miguel, M.G.C.P., Duarte, W.F., Dias, D.R., Schwan, R.F., 2013. Microbial succession and the dynamics of metabolites and sugars during the fermentation of three different cocoa (*Theobroma cacao* L.) hybrids. *Food Res. Int.* 54, 9–17.
- Moreira, I.M.V., Vilela, L.F., Miguel, M.G.C.P., Santos, C., Lima, N., Schwan, R.F., 2017. Impact of a microbial cocktail used as a starter culture on cocoa fermentation and chocolate flavor. *Molecules* 22, 766.
- Mota-Gutierrez, J., Botta, C., Ferrocino, I., Giordano, M., Bertolino, M., Dolci, P., Cannoni, M., Cocolin, L., 2018. Dynamics and biodiversity of bacterial and yeast communities during fermentation of cocoa beans. *Appl. Environ. Microbiol.* 84, 10.1128/AEM.01164-18.
- Nielsen, D.S., Crafaack, M., Jespersen, L., Jakobsen, M., 2013. The microbiology of cocoa fermentation. In: Chocolate in Health and Nutrition. *Humana Press*, Totowa, NJ, pp. 39–60.
- Oksanen, J.F., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szeocs, E., Wagner, H., 2007. The vegan package. *Community ecology package* 10, 631–637.
- Oliveira, H.S.S., Mamede, M.E.O., Góes-neto, A., Koblitz, M.G.B., 2011. Improving chocolate flavor in poor-quality cocoa almonds by enzymatic treatment. *J. Food Sci.* 76, 755–759.
- Ouattara, H.D., Ouattara, H.G., Adom, J.N., Goualié, B.G., Koua, G.A., Doué, G.G., Niamke, S.L., 2016. Screening of lactic acid bacteria capable to breakdown citric acid during Ivorian cocoa fermentation and response of bacterial strains to fermentative conditions. *Br. Biotechnol. J.* 10, 1–10.
- Ouattara, H.D., Ouattara, H.G., Droux, M., Reverchon, S., Nasser, W., Niamke, S.L., 2017. Lactic acid bacteria involved in cocoa bean fermentation from Ivory Coast: species diversity and citrate lyase production. *Int. J. Food Microbiol.* 256, 11–19.
- Pacheco-Montealegre, Mauricio Edilberto, Dávila-Mora, Lízeth Lorena, Botero-Rute, Lina Marcela, Reyes, Alejandro, Caro-Quintero, Alejandro, 2020. Fine resolution analysis of microbial communities provides insights into the variability of cocoa bean fermentation. *Front. Microbiol.* 11, 650.
- Papalexandratou, Z., De Vuyst, L., 2011. Assessment of the yeast species composition of cocoa bean fermentations in different cocoa-producing regions using denaturing gradient gel electrophoresis. *FEMS Yeast Res.* 11, 564–574.
- Papalexandratou, Z., Nielsen, D.S., 2016. It's Gettin'Hot in here: breeding robust yeast starter cultures for cocoa fermentation. *Trends Microbiol.* 24, 168–170.
- Papalexandratou, Z., Vrancken, G., Bruyne, K., Vandamme, P., De Vuyst, L., 2011b. Spontaneous organic cocoa bean box fermentations in Brazil are characterized by a restricted species diversity of lactic acid bacteria and acetic acid bacteria. *Food Microbiol.* 28, 1326–1338.
- Papalexandratou, Z., Camu, N., Falovy, G., De Vuyst, L., 2011c. Comparison of the bacterial species diversity of spontaneous cocoa bean fermentations carried out at selected farms in Ivory Coast and Brazil. *Food Microbiol.* 28, 964–973.
- Papalexandratou, Z., Lefeber, T., Bahrim, B., Lee, O.s., Daniel, H.m., De Vuyst, L., 2013. Hanseniaspora opuntiae, *Saccharomyces cerevisiae*, *Lactobacillus fermentum*, and *Acetobacter pasteurianus* predominate during well-performed Malaysian cocoa bean box fermentations, underlining the importance of these microbial species for a successful cocoa bean fermentation process. *Food Microbiol.* 35, 73–85. <https://doi.org/10.1016/j.fm.2013.02.015>.
- Papalexandratou, Z., Falony, G., Romanens, E., Jimenez, J.c., Amores, F., Daniel, H., De Vuyst, L., 2011a. Species diversity, community dynamics, and metabolite kinetics of the microbiota associated with traditional Ecuadorian spontaneous cocoa bean fermentations. *Appl. Environ. Microbiol.* 77 (21), 7698–7714. <https://doi.org/10.1128/AEM.05523-11>. Epub 2011 Sep 16. PMID: 21926224.
- Papalexandratou, Z., Kaasik, K., Kauffmann, L.V., Skorstengaard, A., Bouillon, G., Espensen, J.L., Castro-Mejía, J.L., et al., 2019. Linking cocoa varietals and microbial diversity of Nicaraguan fine cocoa bean fermentations and their impact on final cocoa quality appreciation. *Int. J. Food Microbiol.* 304, 106–118.
- Pereira, G.V.M., Miguel, M.G.C.P., Ramos, C.L., Schwan, R.F., 2012. Microbiological and physicochemical characterization of small-scale cocoa fermentations and screening of yeast and bacterial strains to Develop a defined starter culture. *Appl. Environ. Microbiol.* 78, 5395–5405.
- Pereira, G.V.M., Magalhães, K.T., Almeida, E.G., Coelho, I.S., Schwan, R.F., 2013a. Spontaneous cocoa bean fermentation carried out in a novel-design stainless steel tank: influence on the dynamics of microbial populations and physical-chemical properties. *Int. J. Food Microbiol.* 161, 121–133.
- Pereira, G.V.M., Magalhães-guedes, K.T., Schwan, R.F., 2013b. rDNA-based DGGE analysis and electron microscopic observation of cocoa beans to monitor microbial diversity and distribution during the fermentation process. *Food Res. Int.* 53, 482–486.
- Pereira, G.V.M., Alvarez, J.P., Neto, D.P.C., Soccol, V.T., Tanobe, V.O.A., Rogez, H., Goes-neto, A., Soccol, C.R., 2017. Great intraspecies diversity of *Pichia kudriavzevii* in cocoa fermentation highlights the importance of yeast strain selection for flavor modulation of cocoa beans. *LWT - Food Sci. Technol. (Lebensmittel-Wissenschaft -Technol.)* 84, 290–297.
- Pettipher, G.L., 1986. Analysis of cocoa pulp and the formulation of a standardized artificial cocoa pulp medium. *J. Sci. Food Agric.* 37, 297–309.
- Quarterly Bulletin of Cocoa Statistics, Vol. XLIV - No. 1-Cocoa year 2017/2018. Disponível em: <<https://www.icco.org/about-us/icco-news/394-august-2018-quarterly-bulletin-of-coco-a-statistics.html>>. Acesso em: 11 november 2018.
- R Development Core Team, 2019. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. www.r-project.org.
- Ramos, C.L., Dias, D.R., Miguel, M.G.C.P., Schwan, R.F., 2014. Impact of different cocoa hybrids (*Theobroma cacao* L.) and *S. cerevisiae* UFLA CA11 inoculation on microbial communities and volatile compounds of cocoa fermentation. *Food Res. Int.* 64, 908–918.
- Roelofsen, P.A., 1953. *Biochim. biophys. Acta* 10, 410–413.
- Saichana, N., Matsushita, K., Adachi, O., Frébort, I., Frebortov, J., 2015. Acetic acid bacteria: a group of bacteria with versatile biotechnological applications. *Biotechnol. Adv.* 33, 1260–1271.
- Schwan, R.F., Fleet, G.H., 2014. Cocoa and Coffee Fermentations. CRC Press, Boca Raton.
- Schwan, R.F., Wheals, A.E., 2004. The microbiology of cocoa fermentation and its role in chocolate quality. *Crit. Rev. Food Sci. Nutr.* 44, 205–221.
- Schwan, R.F., Cooper, R.M., Wheals, A.E., 1997. Endopolygalacturonase secretion by *Kluveromyces marxianus* and other cocoa pulp-degrading yeasts. *Enzym. Microb. Technol.* 21, 234–244.
- Schwenninger, S.M., Leischfeld, S.F., Gantenbein-demarchi, C., 2016. High-throughput identification of the microbial biodiversity of cocoa bean fermentation by MALDI-TOF MS. *Lett. Appl. Microbiol.* 63, 347–355.
- Silva, E.G., Borges, M.F., Medina, C., Piccoli, R.H., Schwan, R.F., 2005. Pectinolytic enzymes secreted by yeasts from tropical fruits. *FEMS Yeast Res.* 5, 859–865.
- Singleton, V.L., Rossi, J., 1965. A Colorimetry of total phenolics with phosphomolybdic phosphotungstic acid reagents. *Am. J. Enol. Vitic.* 16, 144–158.
- Taban, B.M., Saichana, N., 2017. Physiology and biochemistry of acetic acid bacteria. In: Sengun, I.Y. (Ed.), Acetic Acid Bacteria Fundamentals and Food Applications. CRC Press, Boca Raton, FL, pp. 71–91.
- Thompson, S.S., Miller, K.B., Lopez, A.S., 2007. Cocoa and coffee. In: Doyle, M.P., Beuchat, L.R. (Eds.), *Food Microbiology: Fundamentals and Frontiers*, third ed. American Society for Microbiology, Washington DC, USA, pp. 837–849.
- Vásquez, Z.S., de Carvalho Neto, D.P., Pereira, G.V., Vandenberghe, L.P., de Oliveira, P. Z., Tiburcio, P.B., Rogez Góes-Neto, A., Soccol, C.R., 2019. Biotechnological approaches for cocoa waste management: a review. *Waste Manag.* 90, 72–83.
- Visintin, S., Alessandria, V., Valente, A., Dolci, P., Cocolin, L., 2016. Molecular identification and physiological characterization of yeasts, lactic acid bacteria and acetic acid bacteria isolated from heap and box cocoa bean fermentations in West Africa. *Int. J. Food Microbiol.* 216, 69–78.
- Visintin, S., Batista, N., Schwan, F., Cocolin, L., 2017. Impact of *Saccharomyces cerevisiae* and *Torulaspora delbrueckii* starter cultures on cocoa beans fermentation. *Int. J. Food Microbiol.* 257, 31–40.
- Wood, G.A.R., Lass, R.A., 2001. *Cocoa*. Blackwell Science, Oxford, United Kingdom, 620 pp.