Full Paper

Microbial composition and metabolic profiles during machine-controlled intra-factory fermentation of cocoa beans harvested in semitropical area of Japan

Hiroya NISHIMURA[1a], Yuh SHIWA[2, 3a], Satoru TOMITA[4a] and Akihito ENDO[1, 5*]

[1]Department of Food, Aroma and Cosmetic Chemistry, Faculty of Bioindustry, Tokyo University of Agriculture, 196 Yasaka, Abashiri-shi, Hokkaido 099-2493, Japan

[2]Department of Molecular Microbiology, Faculty of Life Sciences, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan

[3]NODAI Genome Research Centre, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan

[4]Institute of Food Research, National Agriculture and Food Research Organization (NARO), 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan

[5]Department of Nutritional Science and Food Safety, Faculty of Applied Bioscience, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan

Cocoa bean fermentation is typically performed in a spontaneous manner on farms in tropical countries or areas and involves several microbial groups. Metabolism by microbes markedly affects the quality of cocoa beans fermented and the chocolate produced thereof. The present study characterized the microbiota and their metabolic profiles in temperature- and humidity-controlled intra-factory cocoa fermentation in a semitropical area of Japan. Although environmental factors were uniform, the microbiota of cocoa beans subjected to intra-factory fermentation was not stable between tests, particularly in terms of the cell count levels and species observed. Fermentation was sometimes delayed, and fermenting microbes were present at very low levels after 24 hr of fermentation. Due to the unstable microbiota, the profiles of water-soluble compounds differed between tests, indicating the unstable qualities of the fermented cocoa beans. These results suggest the necessity of starter cultures not only in on-farm fermentation but also in machine-controlled intra-factory cocoa fermentation.

Key words: cocoa fermentation, intra-factory fermentation, Japan, microbiota, metabolomics

INTRODUCTION

The cocoa tree is a tropical plant that is typically cultivated in tropical countries around the equatorial zone, 20 degrees north and south of the equator, including the Ivory Coast, Ghana, Indonesia, Brazil, Ecuador, and Peru [1]. The cocoa fruits (pods) harvested in these tropical countries are split, and the beans and pulp are isolated and used for on-farm fermentation. Cacao pulp is a rich source of water (approx. 80%) and carbohydrates (approx. 10%) and, thus, enhances microbial growth [2]. Fermentation reduces endogenous bitter and astringent compounds. Although wooden boxes and heaps are the main methods used for fermentation, baskets, plastic boxes, platforms, and trays are also employed [1]. Fermented cocoa beans are dried and roasted and then used for cocoa products, including chocolate.

Fermentation typically occurs in a spontaneous manner and continues for between 5 and 10 days. Several microbes, including yeast, lactic acid bacteria (LAB), and acetic acid bacteria (AAB), are involved in fermentation [3]. Although the succession of microbiota during fermentation varies among regions, fermentation is typically initiated by yeasts, followed by LAB and AAB. Spore-forming bacteria, mainly *Bacillus* spp., are sometimes dominant at the end of fermentation [1, 2]. Each microbe has a different impact on fermentation and cocoa quality. Yeasts are responsible for ethanol production and pectinolysis, and LAB are responsible for the production of lactate and acetate as well as citrate metabolism [2, 4]. AAB mainly produce acetate from carbohydrates and also covert ethanol to acetate and acetoin in a latter stage [5]. Several AAB have the potential to perform oxidative fermentation of acetate and lactate to carbon

---

[a]These authors contributed equally to the work.*Corresponding author. Akihito Endo (E-mail: a3endou@nodai.ac.jp)

dioxide and water [6, 7]. Metabolism by these microbial groups has a marked impact on the quality of cocoa fermentation. The growth of spore-forming *Bacillus* spp. is attributed to a reduction in acidity caused by the oxidation of AAB and an elevated fermentation temperature (approx. 45°C) in the last stage of fermentation [8]. They may be regarded as off-flavour producers. While fermentation methods, including containers and durations, vary among areas, the main fermentation microbes are relatively similar. *Saccharomyces cerevisiae, Pichia kudriavzevii, Hanseniaspora guilliermondii,* and *Hanseniaspora uvarum* are the major yeasts, while *Limosilactobacillus fermentum* and *Lactiplantibacillus plantarum* are the predominant LAB. The major AAB are *Acetobacter* spp., particularly *Acetobacter pasteurianus* [1, 4]. These microbes have been examined as starter microbes for microbiologically stable fermentation in several countries [5, 9–11].

Cocoa is mainly cultivated and fermented in tropical areas. Japan belongs to a subfrigid zone, temperate zone, and subtropical zone, depending on the region, and culture of cocoa plants has started in the Ogasawara Islands, which have a semitropical climate. The production size is still small at approx. 10 t of cocoa pods per year. Harvested cocoa pods are transported to the capital region, and fermentation is conducted intra-factory under controlled conditions. The fermentation conditions include temporary storage of the fermenting materials in a refrigerator (approx. 4°C) in the middle of fermentation because the factory operates on weekdays only. The temperature generally increases during the fermentation of cocoa beans and reaches >40°C by the middle of fermentation [11, 12]. These temperature changes in intra-factory fermentation markedly affect microbial profiles and metabolites in cocoa, which have yet to be characterized.

In the present study, the microbiota of intra-factory fermented cocoa in Japan was monitored using a culturing technique and shotgun metagenomic analysis. A metabolome analysis was conducted to examine changes in water-soluble compounds during fermentation. The results obtained herein demonstrated that the microbiota of intra-factory fermented cocoa was unstable, which markedly affected the metabolic profiles in cocoa.

MATERIALS AND METHODS

Cocoa fermentation and sample collection

Ripe cocoa pods were harvested in Hahajima Island in the Ogasawara Islands (26.39 N, 142.09 E), Tokyo, Japan, and transported to a factory in Saitama by boat and truck within 3 days. The pods were manually opened with knives, and the cocoa beans and pulp (approx. 20 kg) were immediately transferred to disposable plastic bags. The plastic bags were then placed into plastic boxes (approx. 50% of the volume of the boxes), which were covered with plastic lids and kept at 37°C under 60% humidity in an incubator. Fermentation took 12 days. The boxes were placed in a refrigerator (approx. 4°C) from the evening of day 4 (approx. 105 hr of fermentation) to the morning of day 7 (approx. 168 hr of fermentation) due to a factory weekend holiday. Fermentation samples were mixed for aeration during fermentation using gloved hands every day, except for holidays. New plastic bags were used for every fermentation. Approximately 100 g fermentation samples were obtained on day 1 (24 hr), day 4 (96 hr), day 8 (192 hr), and day 11 (264 hr) in July, October, and November of 2020 and used for microbial and metabolome

analyses. The average temperatures of Ogasawara Islands in July, October, and November are 28, 26, and 24°C, respectively (data obtained from the Japan Meteorological Agency).

Microbiological analysis by culturing

Fermentation samples were manually homogenized, and the pulp or juice was serially diluted in saline solution. Dilutions were plated onto potato dextrose agar (PDA; Nissui, Japan) supplemented with 100 mg/L chloramphenicol; MRS agar (Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10 g/L D-fructose, 10 mg/L cycloheximide, and 10 mg/L sodium azide (F-MRS agar); AAM agar supplemented with 10 mg/L cycloheximide; and nutrient agar (NA; Nissui) supplemented with 10 mg/L cycloheximide for the isolation of fungi, LAB, AAB, and aerobic bacteria, respectively. The AAM agar was enrichment medium I, as described by Lisdiyanti *et al.* [7], and was composed of 10 g/L D-glucose, 15 g/L polypeptone, 8 g/L yeast extract, 0.5% (v/v) ethanol, and 0.3% (v/v) acetic acid (pH 3.5) and supplemented with 15 g/L agar. These media were incubated at 25°C (PDA) or 37°C (F-MRS agar, AAM agar, and NA) for 2 or 3 days. Eight to ten colonies were selected from each agar based on the colony morphologies and colours and inoculated onto PDA slant agar (for fungi) or into F-MRS broth (for LAB), AAM broth (for AAB), or nutrient broth (for aerobic bacteria), and they were then incubated at the appropriate temperatures overnight. Bacterial isolates were maintained at −80°C in sterile half-strength MRS broth supplemented with 20% (v/v) glycerol, and fungal isolates were kept at 4°C on PDA slant agar.

The identification of bacterial and fungal isolates was conducted based on sequencing of the 16S rRNA gene (for bacteria) and internal transcribed spacer (ITS) region (for fungi) using previously described methods [13]. The 16S rRNA gene and ITS sequences determined in the present study are available in GenBank/EMBL/DDBJ under accession numbers LC716662-LC716683.

Shotgun metagenomic analysis

Total genomic DNA was extracted from the pulp and juice of fermentation samples using a QIAamp DNA Stool Mini Kit (Qiagen, Tokyo, Japan). Fermentation samples were supplemented with 0.6 g of zirconia beads and beaten at 5 m/sec for 300 sec in a bead beater (model Bead Mill 4, Fisher Scientific) prior to extraction. Isolated DNA was sent to Macrogen Japan Corp. (Tokyo, Japan) and used for sequencing. A TruSeq Nano DNA kit (Illumina, San Diego, CA, USA) was used for library preparation, and the library was sequenced by 150-bp paired-end sequencing with an Illumina NovaSeq 6000 system. Kneaddata (v0.10.0, https://github.com/biobakery/kneaddata) was used with the default settings to remove reads mapping to the genome of *Theobroma cacao* (Criollo_cocoa_genome_V2, GCF_000208745.1), to filter low-quality reads, and to remove adapters. The remaining reads were randomly subsampled to 5 M reads using seqtk_sample (v0.16.1, https://github.com/lh3/seqtk) with the default settings. Subsampled FASTQ files were then uploaded to the local Galaxy server [14] for further processing. All reads were assigned taxonomy using Kraken 2 [15] (Galaxy Version 2.1.1+galaxy1) against the NCBI non-redundant nucleotide database (built in May 2021) with the "confidence" parameter changed to "0.05". Species-level taxonomic abundance values were estimated using Bracken [16] (Galaxy Version

2.6.1+galaxy0) based on read counts from Kraken 2. Sequences assigned to insects were removed from the dataset for further analyses. Due to the small number of *Tatumella* genus entries in the NCBI non-redundant nucleotide database used in the present study, reads classified to *Tatumella* spp. in the Kraken 2 analysis were extracted using the extract_kraken_reads.py script (https://github.com/jenniferlu717/KrakenTools). The taxonomic positions of the extracted reads were confirmed with Genome Taxonomy Database (GTDB) release 207 produced with Struo2 [17] using Kraken 2. The sequencing data obtained were also included to prepare a dendrogram using the hclust function with the Ward.D2 algorithm in the R (version 3.6.2).

Sequencing data were deposited in the NCBI Sequence Read Archive under the accession number DRA014576.

NMR-based metabolomics

To extract water-soluble metabolites, three thawed cocoa beans were transferred to the same weight of ice-cold water in a 50-mL centrifuge tube and vigorously vortexed at room temperature (~20°C) for 5 min. After removing the beans, that the pulp layers of which had come off, the clear supernatant was collected by centrifuging the remaining suspension at $12,000 \times g$ for 10 min at 4°C. Extraction was performed in triplicate for each sample. An aliquot of the supernatant was used to measure pH values, and the residual was stored at −20°C for later analyses. Analytical sample preparation was performed as previously described [18]. The proton (^{1}H) NMR spectrum for the metabolome analysis was recorded using the Bruker pulse program zgpr as previously reported [18] with some modifications: probe, a CryoProbe CPDUL (Bruker, Billerica, MA, USA); ^{1}H 90° pulse, 13.4 μsec; number of scans, 256. Measurements of two-dimensional NMR spectra followed by metabolite annotations were performed as described previously [19]. Regarding non-targeted NMR metabolomics, a dataset was prepared using Amix software (Bruker) by subdividing ^{1}H NMR spectra into 0.04-ppm-wide integral bins followed by normalization to the signal intensity of the internal standard. Additionally, to investigate temporal changes in the signal intensity of each metabolite, integrals in the regions of interest (ROIs) that contained an isolated signal derived from an individual compound were calculated using the variable-sized buckets option of Amix.

Statistical analysis

Regarding non-targeted metabolomics, the variables of 230 bins, excluding those with a solvent signal, were subjected to a principal component analysis (PCA) using the SIMCA software (ver. 14; Umetrics, UmeÅ, Sweden) and applying Pareto scaling. A correlation analysis and hierarchical clustering analysis (HCA) were performed by Spearman's rank correlation and Ward's method, respectively, using R (ver. 4.2.1). Data on the HCA heatmap and temporal signal intensities were visualized using heatmap3 with the ward.D2 algorithm and ggplot2 R-packages, respectively [20, 21].

RESULTS

Changes in pH during cocoa fermentation

Changes in pH were monitored during fermentation. At the beginning of fermentation (24 hr), pH was stable among the 3 batches and ranged between 3.7 and 3.8 (Fig. 1). It gradually

increased during fermentation and reached 4.7 in July and 5.2 in October and November at 264 hr of fermentation.

Microbiota as determined by the culturing technique

Viable cell counts during cocoa fermentation varied among the tests. In the test in July, yeast and LAB were already $>10^8$ and $>10^7$ colony forming unit (CFU)/g of cocoa mass, respectively, after 24 hr of fermentation (Fig. 2A). Yeast cell numbers remained relatively stable until 192 hr but decreased to $<10^7$ CFU/g at 264 hr, whereas those of LAB were stable and maintained at $>10^8$ CFU/g until 264 hr. AAB were not isolated at 24 hr but markedly increased to $>10^8$ CFU/g by 96 hr and were subsequently maintained at that level until 264 hr. Aerobe cell numbers were approx. 10^6 at 24 hr, but no aerobes were recovered after 96 hr.

During fermentation in October, the viable cell count for yeast was $<10^6$ CFU/g at 24 hr, and it gradually increased until 192 hr (Fig. 2B). LAB were not isolated at 24 hr but markedly increased to $>10^8$ CFU/g by 96 hr and maintained a viable level until 264 hr. Aerobes were initially dominant (approx. 10^8 CFU/g at 24 hr) during this fermentation period but were not isolated after 96 hr. AAB were not isolated during this fermentation period.

During fermentation in November, aerobes initially dominated ($>10^7$ CFU/g at 24 hr) and disappeared after 96 hr (Fig. 2C). Moderate populations of yeast and LAB (approx. 10^5 CFU/g) were observed at 24 hr but increased after 96 hr of fermentation. AAB were not isolated until 96 hr but gradually increased in the latter stages of fermentation (after 192 hr).

A total of 302 colonies were collected from different culture media and identified by ITS sequencing for yeast and 16S rRNA gene sequencing for bacteria. Regarding yeast, *P. kudriavzevii* was observed from all fermentation samples that had yeast cell numbers $>10^7$ CFU/g (Fig. 2D). *Candida* spp., *Hanseniaspora* spp., and other species were occasionally isolated during fermentation. Moulds were not recovered in the present study. Concerning LAB, *L. fermentum* was the sole LAB isolated from fermentation in July, whereas *L. plantarum* was the major LAB during fermentation in October and November. *A. pasteurianus* and *Acetobacter tropicalis* were the major AAB in July, while *A. tropicalis* was the sole AAB species in November. Aerobes were only isolated at 24 hr (Fig. 2A–2C). The species isolated were from the phylum Pseudomonadota (formerly Proteobacteria) and included *Pantoea dispersa*, *Tatumella morbirosei*, *Tatumella saanichensis*, *Serratia* sp., and *Enterobacter* sp. (data not shown).
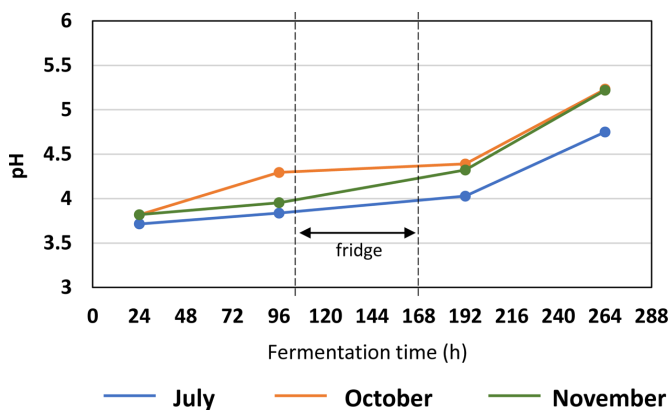

Fig. 1. Changes in pH during cocoa fermentation.

Spore-forming *Bacillus* spp. were not isolated in the present study.

Microbiota as determined by the shotgun metagenomic approach

The sequencing of metagenomic DNA from the 12 fermentation samples generated 31,861,887 ± 3,538,135 reads (average ± standard deviation, SD) and a total read bp of 4.8 ± 0.5 Gbp (average ± SD). Sequencing data were annotated with Kraken 2, and sequence data that originated from plants (mainly from *Tatumella cacao*) and insects were removed from the dataset and used to examine the taxonomic composition.

The shotgun metagenomic approach revealed that the microbiota for cocoa fermentation markedly changed between 24 and 96 hr. In July, *P. kudriavzevii* and *L. fermentum* were dominant, followed by fungi (*Aspergillus* spp. and *Lasiodiplodia*
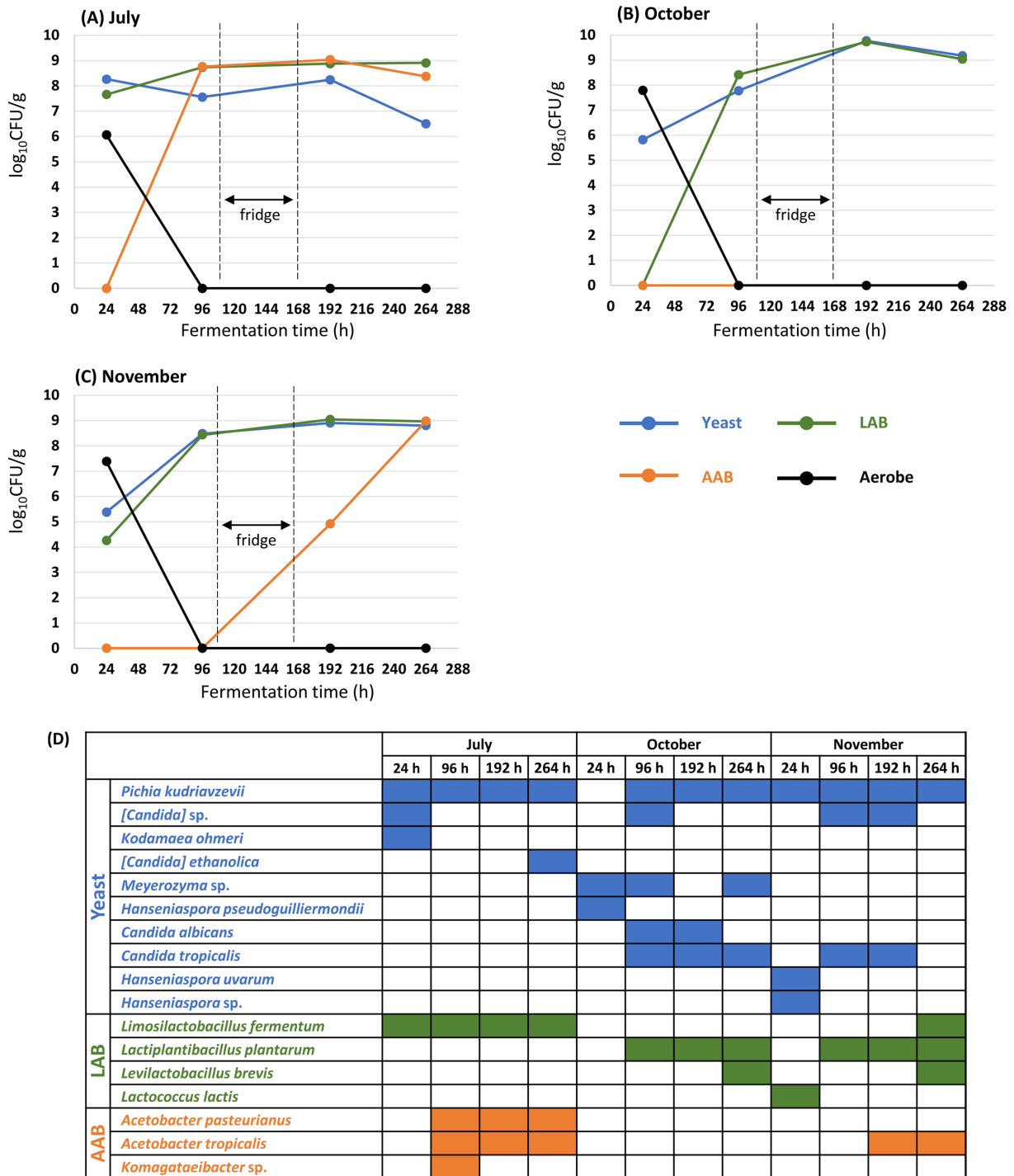

Fig. 2. Viable cell counts of yeast, LAB, AAB, and aerobic bacteria during cocoa fermentation in July (A), October (B), and November (C). The species isolated at different time points in the three fermentation periods are indicated by colours (D).

LAB: lactic acid bacteria; AAB: acetic acid bacteria.

theobromae) and *Limosilactobacillus oris,* after 24 hr of fermentation (Fig. 3). However, the relative abundance of *A. pasteurianus* in the microbiota was 88 to 92% after 96 hr of fermentation in July. *L. fermentum* was detected to have a relative abundance of 1–2% after 96 hr of fermentation.

Tatumella spp. and *Pantoea* spp. in the phylum Pseudomonadota were the major microbes after 24 hr of fermentation in October (Fig. 3), and no species belonging to yeast, LAB, or AAB were detected with a relative abundance >1%. However, *P. kudriavzevii* and *L. plantarum* were dominant after 96 hr of fermentation in October, and the sum of their relative abundances was between 88 and 93%. Similar changes in the microbiota were observed during fermentation in November. *Tatumella* spp. and *Pantoea* spp. were dominant after 24 hr of fermentation, and *P. kudriavzevii* and *L. plantarum* were dominant in the microbiota after 96 hr (Fig. 3). One marked difference observed between October and November was an increased population of *Gluconobacter oxydans* after 264 hr of fermentation in November.

The metagenomic analysis showed that *Tatumella citrea* was predominant after 24 hr of fermentation in October and that *T. citrea* and *Tatumella ptyseos* were predominant after 24 hr in November. However, these species were not isolated by the culturing technique. *T. morbirosei* and *T. saanichensis* were isolated after 24 hr of fermentation in October and November, respectively. Therefore, the sequences assigned to *T. citrea* and *T. ptyseos* were further identified by GTDB (Release 207). This analysis revealed that the *Tatumella* species were mainly composed of *T. morbirosei* (approx. 89%) and *T. citrea* (approx. 10%) after 24 hr of fermentation in October and mainly composed of *T. citrea* (approx. 54%), *T. saanichensis* (approx. 31%), and *T. morbirosei* (approx. 11%) after 24 hr of fermentation in November.

Shannon's diversity index was commonly highest after 24 hr in the three fermentation periods, and diversity markedly decreased after 96 hr of fermentation (Fig. 4A). This decrease was attributed to the dominance of 1 or 2 specific species after 96 hr of fermentation. HCA based on the composition of the microbiota produced three major clusters (Fig. 4B). The first cluster contained 24-hr-old samples from all three of the tested months, while the second cluster included 96-, 192-, and 264-hr-old samples from October and November. The third cluster contained 96-, 192-, and 264-hr-old samples from July.

Nuclear Magnetic Resonance (NMR)-based metabolomics

Proton and two-dimensional NMR analyses of the water-soluble extract of fermented cocoa beans detected many signals annotated as various chemical species, including aldose, alditol, aldonic acid, uronic acid, amino acid, alcohol, organic acid, polyphenol, and other miscellaneous metabolites (Supplementary Fig. 1). Based on ^{1}H NMR spectra, a non-targeted PCA was initially performed to comprehensively compare differences in the water-soluble metabolite profile among samples (Fig. 5A). The first and second principal components (PC1 and PC2) explained 61.9 and 21.9% of the total variance, respectively. In the score plot, samples showed a clear separation into three classes. The samples collected after a short period of fermentation (24 hr) organized into a single group regardless of the preparation month, whereas those from longer periods were divided into
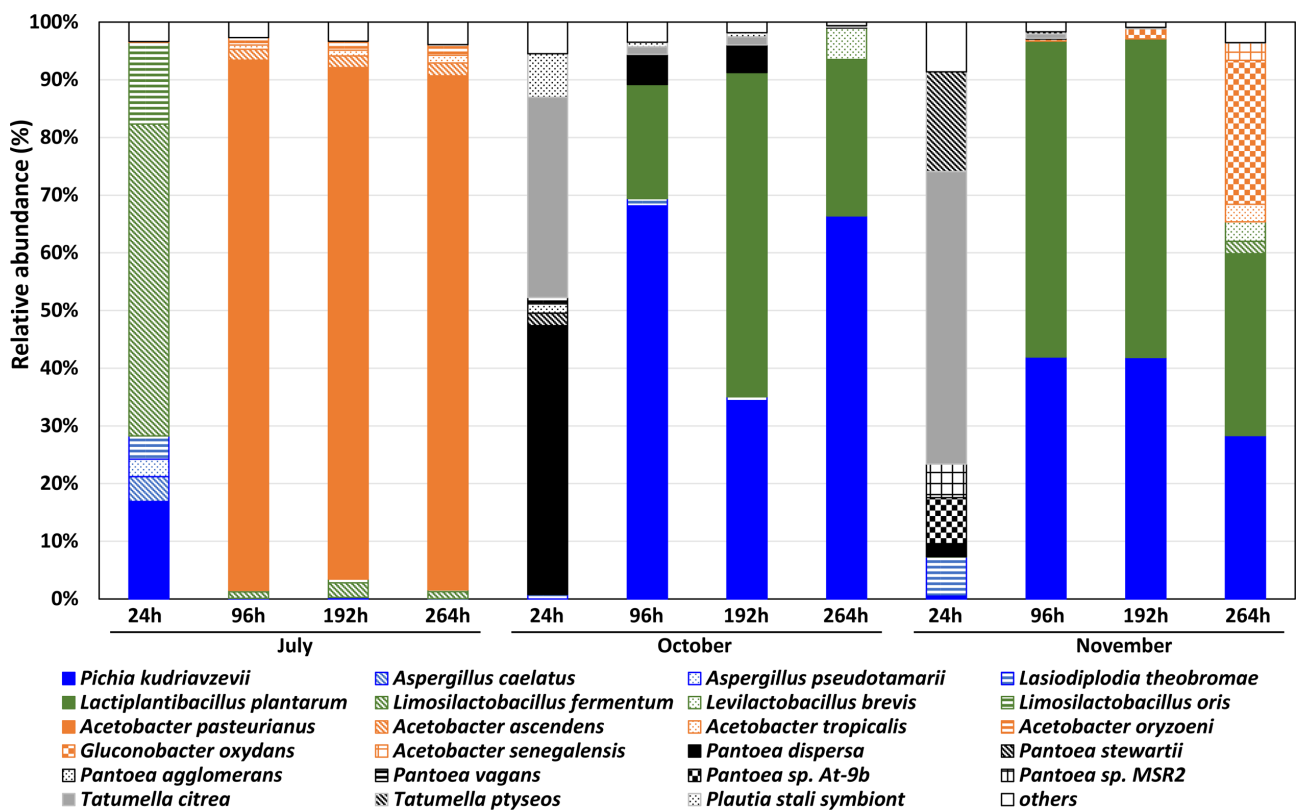

Fig. 3. Relative abundance of each microbe based on a shotgun metagenomic analysis. Different colours represent different microbial groups: blue, fungi; green, lactic acid bacteria (LAB); orange, acetic acid bacteria (AAB); and grey-black, other bacteria.
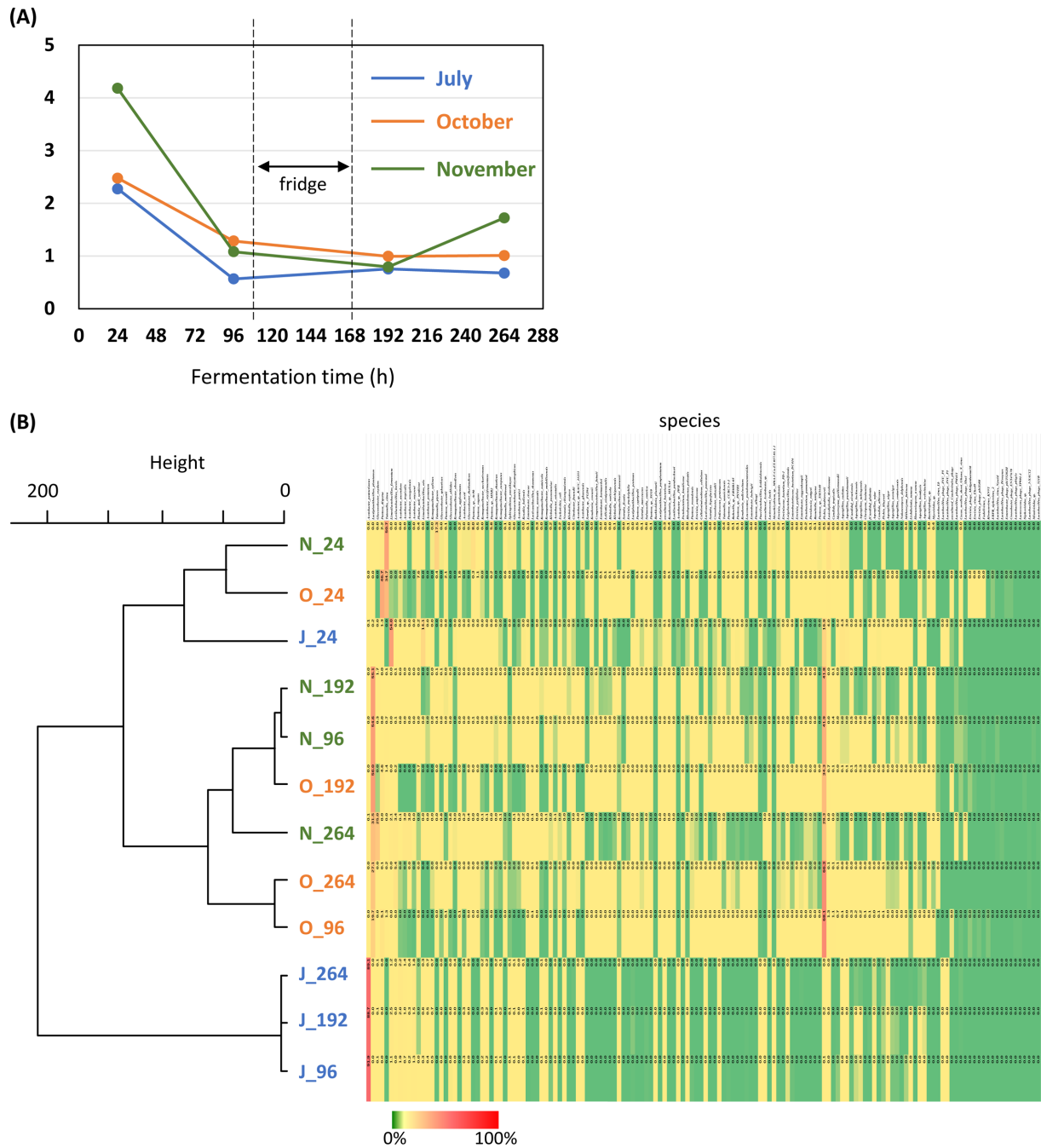
Fig. 4. Shannon index during cocoa fermentation (A) and hierarchical clustering analysis (HCA) based on the composition of the microbiota and heatmap indicating microbial composition (B). In (B), J, O, and N indicate fermentation in July, October, and November, respectively, and 24, 96, 192, and 264 indicate the fermentation times (hr).

two classes comprised of the samples from July and those from October and November. The loading plot showed that these separations were primarily explained by NMR signals derived from ethanol, aldoses (glucose and fructose), alditols (mannitol and glycerol), organic acids (acetic, lactic, and citric acids), and 2,3-butanediol (inseparable from 1,2-propanediol). Samples were additionally compared using a targeted approach applying the signal intensities of 43 ROIs comprising 39 compounds and four unannotated signals. The heatmap of HCA data is shown in Fig. 5B. Samples were separated into three clusters, and the

separations were consistent with those observed in PCA data. These results highlighted the differences in the water-soluble compound profiles by 96 hr between the samples from July and those from October and November.

Temporal changes in the signal intensity of each compound are shown in Fig. 6. The levels of glucose, fructose, sucrose, citric acid, malic acid, and glutamine were high at 24 hr and quickly decreased by 96 hr (Fig. 6A), suggesting that they were derived from raw cocoa beans and were commonly consumed during the initial stage of fermentation. However, the compounds
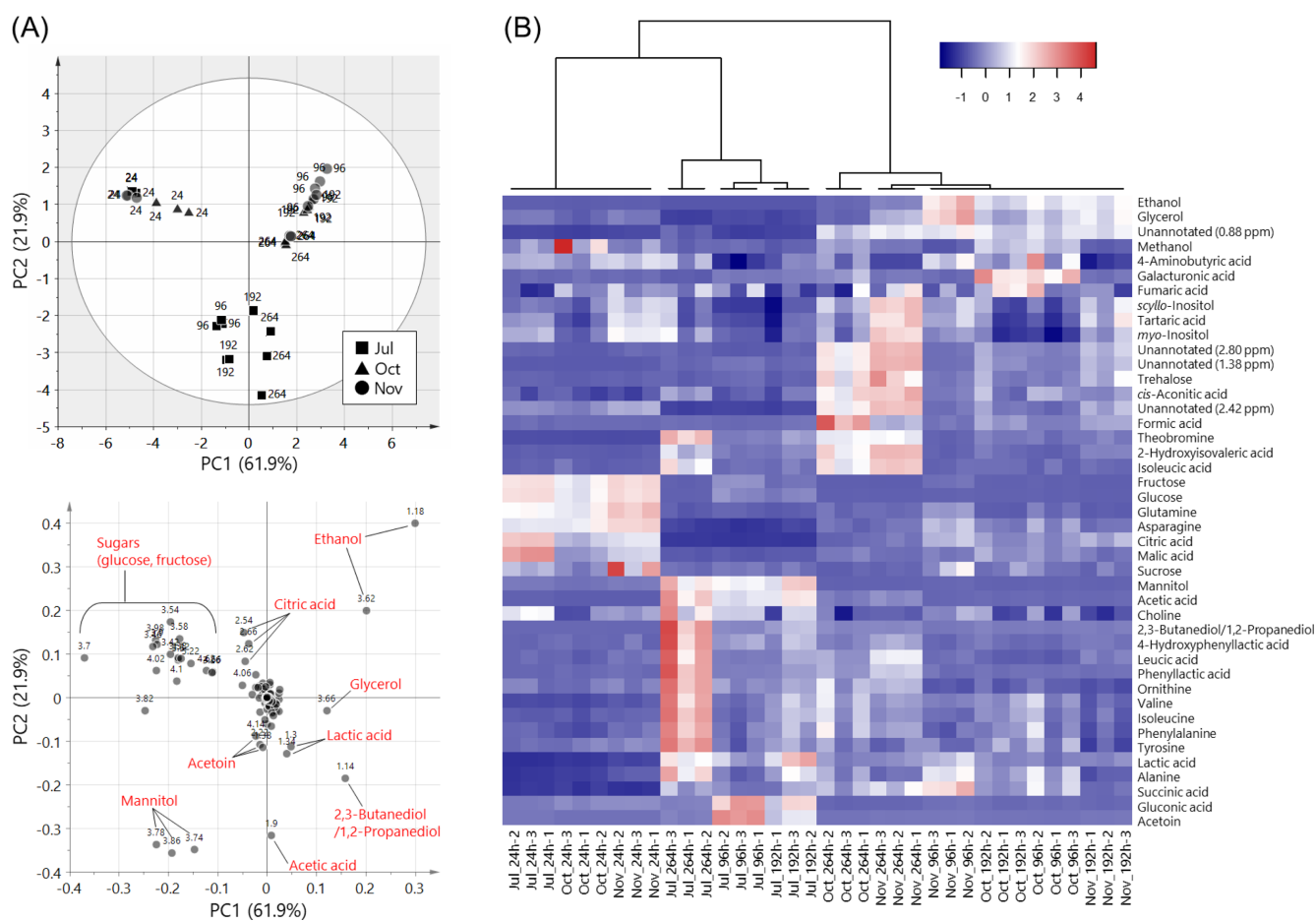
Fig. 5. Metabolomic comparison of water-soluble compound profiles among fermented cocoa samples. Non-targeted principal component analysis (PCA) scores and loading plots (A) and a heatmap of a regions of interest (ROI)-based hierarchical clustering analysis (B). Numerical labels in score and loading plots represent the fermentation time (hr) and chemical shifts (ppm) in NMR spectral bins. The major metabolites responsible for the feature space are indicated on the loading plot in red.

that increased along with this consumption differed between the samples from July and those from October and November. The levels of gluconic acid, acetoin, acetic acid, mannitol, and lactic acid increased during fermentation in July (Fig. 6B), whereas those of ethanol and an unannotated signal at 0.88 ppm markedly increased during fermentation in October and November (Fig. 6C). The compounds that markedly increased after 264 hr of fermentation were also involved in the difference between cocoa preparation months: 2,3-butanediol (1,2-propanediol), some amino acids (phenylalanine, tyrosine, isoleucine, and valine), and their possible transaminated products (phenyllactic acid, 4-hydroxyphenyllactic acid, and leucic acid) were responsible for the samples in July (Fig. 6B), whereas trehalose, *cis*-aconitic acid, and an unannotated signal at 1.38 ppm were responsible for those in October and November (Fig. 6C). Although temporal changes in most metabolites were very similar between October and November, the samples in October exhibited specific increases in galacturonic acid, methanol, and formic acid signals (Fig. 6D). Theobromine, which is a dominant polyphenol in cocoa beans [22], rapidly increased from 192 to 264 hr, suggesting that the exudation of the inner chemical was enhanced after 192 hr. This would be due to cell disruption inside cocoa beans caused by microbial metabolites and fermentation heat at the end of

fermentation, as described previously [1]. Collectively, these results indicated that two large compositional changes, occurring between 24 and 96 hr and between 192 and 264 hr of fermentation, resulted in the differences observed among the cocoa preparation months.

Correlation analysis

A relationship was expected between the microbiota and water-soluble compound profiles observed above. A comparison of NMR signal intensities in ROIs between samples in July and those in October and November revealed significant differences in many compounds, which are listed in Table 1. The relative ratios of mannitol, acetoin, gluconic acid, glucose, and acetic acid were >10-fold greater in July samples, whereas those of ethanol, citric acid, trehalose, galacturonic acid (only significant at 264 hr), and unannotated signals (0.88, 1.38, and 2.42 ppm) were <0.1-fold lower in July samples. The correlation analysis showed that some of these compounds were strongly correlated ($R>0.600$) with relative microbial dominance at the species level. The relative dominance of *A. pasteurianus* was strongly correlated with the compounds detected in July, including mannitol, acetoin, gluconic acid, and ethanol, with correlation coefficients of $R=0.803$, 0.708, 0.634, and −0.651, respectively. The relative dominance of *L.*
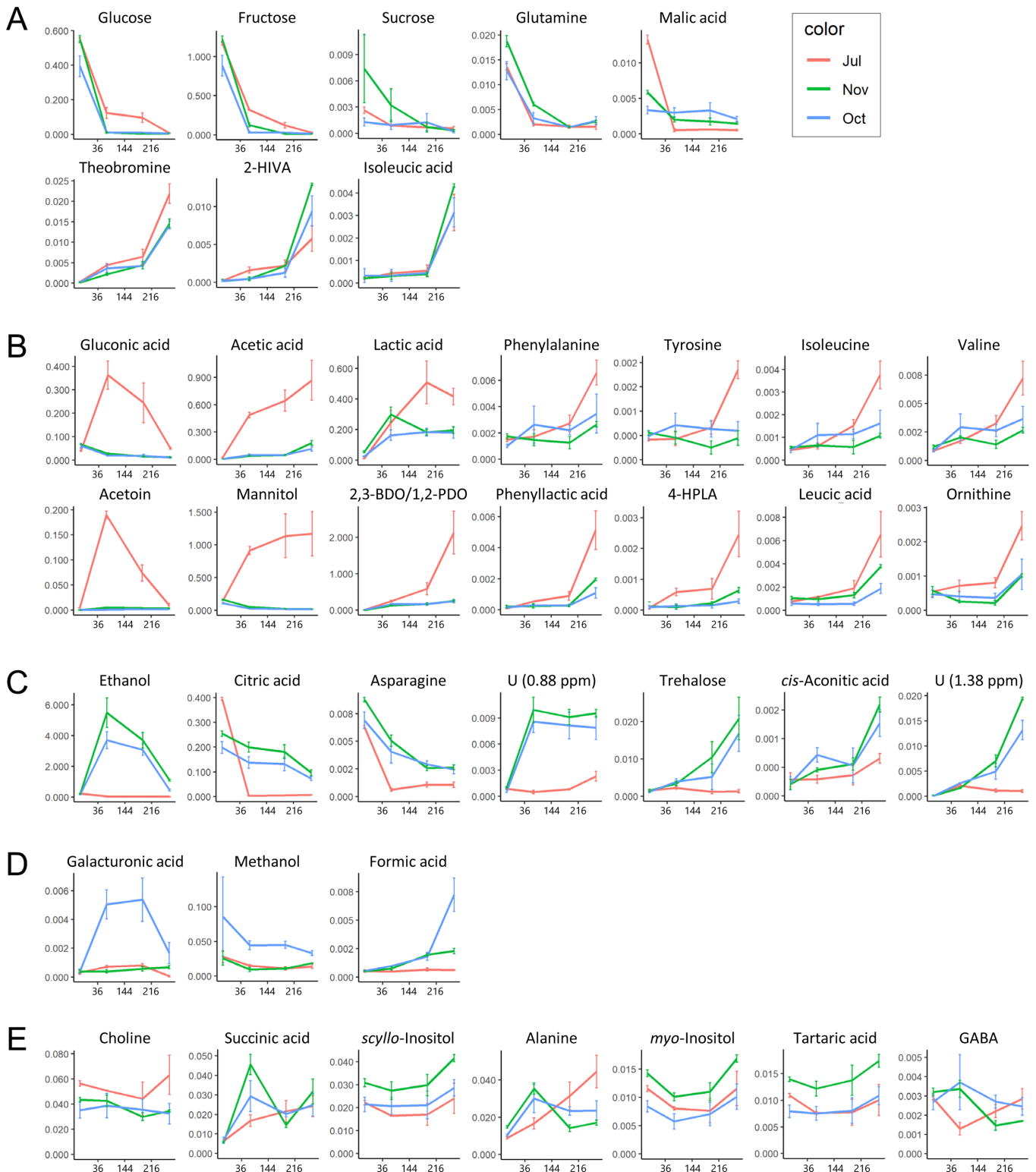
Fig. 6. Temporal changes in signal intensities of water-soluble metabolites during cocoa fermentation. Metabolites exhibiting common decreases or increases (A), those responsible for characterizing samples in July (B), those contributing to both samples in October and November (C), those relatively specific to samples in October (D), and those with unclear changes during fermentation (E). The x and y axes indicate the fermentation times (hour) and ROI-based signal intensities relative to the internal standard, respectively. Data are shown as the mean ± SD (n=3). 2-HIVA: 2-hydroxyisovaleric acid; 2,3-BDO: 2,3-butanediol; 1,2-PDO: 1,2-propanediol; 4-HPLA: 4-hydroxyphenyllactic acid; U: unannotated; GABA: 4-aminobutyric acid.

Table 1. Comparative analysis of metabolite intensities among fermented cocoa preparations

Annotated compounds	hr	Signal intensity relative to DSS-d_6[1]						Comparative analysis[2]		
		July		October		November		July vs. October & November		
		Mean	SD	Mean	SD	Mean	SD	Ratio[3]	p-value	
2,3-Butanediol/1,2-Propanediol	24	8.6E-3	1.3E-3	6.1E-3	9.2E-4	3.4E-3	3.9E-4	1.79	0.012	*
	96	2.4E-1	2.2E-2	1.6E-1	1.9E-2	1.2E-1	1.5E-2	1.71	0.002	**
	192	5.9E-1	1.7E-1	1.7E-1	3.7E-2	1.7E-1	2.5E-2	3.48	0.049	*
	264	2.1E+0	5.9E-1	2.5E-1	4.7E-2	2.5E-1	1.8E-2	8.57	0.031	*
2-Hydroxyisovaleric acid	24	2.4E-4	1.4E-4	2.2E-4	9.4E-5	1.9E-4	2.1E-4	1.19	0.715	
	96	1.6E-3	4.0E-4	4.6E-4	2.4E-4	4.9E-4	2.0E-4	3.42	0.026	*
	192	2.2E-3	7.7E-4	1.3E-3	5.9E-4	2.2E-3	8.5E-5	1.27	0.418	
	264	5.8E-3	1.7E-3	9.4E-3	2.0E-3	1.3E-2	1.3E-4	0.52	0.009	**
4-Aminobutyric acid	24	2.9E-3	8.3E-5	2.7E-3	4.2E-4	3.2E-3	1.9E-4	0.97	0.665	
	96	1.3E-3	3.3E-4	3.7E-3	1.4E-3	3.4E-3	4.3E-4	0.37	0.002	**
	192	2.2E-3	3.3E-4	2.7E-3	3.3E-4	1.5E-3	2.5E-4	1.06	0.733	
	264	2.9E-3	5.2E-4	2.5E-3	4.6E-4	1.7E-3	1.3E-5	1.37	0.104	
4-Hydroxyphenyllactate	24	7.6E-5	4.6E-5	1.0E-4	5.0E-5	1.3E-4	1.4E-4	0.65	0.412	
	96	5.9E-4	1.2E-4	1.3E-4	6.6E-5	9.8E-5	6.8E-5	5.11	0.014	*
	192	6.9E-4	3.3E-4	1.6E-4	9.7E-5	2.2E-4	5.2E-5	3.61	0.119	
	264	2.5E-3	7.4E-4	2.9E-4	7.4E-5	6.5E-4	8.3E-5	5.27	0.038	*
Acetic acid	24	2.2E-2	1.4E-3	6.5E-3	6.4E-4	5.4E-3	2.0E-4	3.71	0.001	***
	96	4.9E-1	3.1E-2	4.9E-2	9.8E-3	4.0E-2	5.7E-3	10.97	0.001	**
	192	6.4E-1	1.2E-1	5.1E-2	7.0E-3	4.9E-2	5.9E-3	12.97	0.012	*
	264	8.7E-1	2.2E-1	1.2E-1	2.3E-2	1.8E-1	2.5E-2	5.80	0.028	*
Acetoin	24	4.3E-3	2.7E-4	5.0E-4	2.2E-5	3.7E-4	9.5E-5	10.06	0.001	***
	96	1.9E-1	7.2E-3	1.2E-3	3.4E-4	5.1E-3	7.4E-4	59.82	0.000	***
	192	7.4E-2	1.6E-2	2.2E-3	3.8E-4	4.3E-3	4.9E-4	22.54	0.016	*
	264	1.0E-2	2.8E-3	1.7E-3	3.3E-5	3.8E-3	4.5E-4	3.73	0.032	*
Alanine	24	8.9E-3	5.4E-4	1.0E-2	5.1E-4	1.5E-2	7.0E-4	0.71	0.014	*
	96	1.7E-2	3.0E-3	3.0E-2	7.6E-3	3.5E-2	3.0E-3	0.51	0.001	**
	192	3.1E-2	7.4E-3	2.3E-2	2.3E-3	1.4E-2	1.8E-3	1.68	0.075	
	264	4.4E-2	8.8E-3	2.4E-2	5.3E-3	1.7E-2	1.4E-3	2.18	0.027	*
Asparagine	24	6.3E-3	1.1E-4	6.9E-3	7.6E-4	8.8E-3	1.9E-4	0.81	0.022	*
	96	6.3E-4	1.3E-4	4.1E-3	1.0E-3	5.1E-3	5.0E-4	0.14	0.000	***
	192	1.1E-3	2.3E-4	2.9E-3	3.1E-4	2.6E-3	2.4E-4	0.39	0.000	***
	264	1.1E-3	2.4E-4	2.4E-3	4.1E-4	2.6E-3	2.2E-4	0.42	0.000	***
Choline	24	5.6E-2	2.2E-3	3.5E-2	6.4E-3	4.4E-2	1.8E-3	1.43	0.001	***
	96	5.1E-2	2.8E-4	3.9E-2	9.7E-3	4.3E-2	4.5E-3	1.25	0.017	*
	192	4.4E-2	1.3E-2	3.6E-2	8.5E-3	3.0E-2	3.2E-3	1.35	0.269	
	264	6.3E-2	1.6E-2	3.2E-2	8.2E-3	3.5E-2	8.4E-4	1.89	0.073	
cis-Aconitic acid	24	2.8E-4	1.2E-4	2.2E-4	1.1E-4	2.0E-4	8.9E-5	1.33	0.431	
	96	2.9E-4	6.9E-5	7.2E-4	1.3E-4	4.6E-4	3.8E-5	0.50	0.007	**
	192	3.6E-4	1.7E-4	5.3E-4	3.1E-4	5.6E-4	1.1E-4	0.67	0.224	
	264	6.5E-4	8.5E-5	1.3E-3	2.3E-4	1.6E-3	1.3E-4	0.46	0.000	***
Citric acid	24	3.9E-1	6.0E-3	2.0E-1	2.4E-2	2.5E-1	1.1E-2	1.74	0.000	***
	96	3.2E-3	8.8E-4	1.4E-1	2.3E-2	2.0E-1	2.0E-2	0.02	0.000	***
	192	4.9E-3	1.8E-3	1.3E-1	2.7E-2	1.8E-1	2.9E-2	0.03	0.000	***
	264	7.2E-3	7.9E-4	7.3E-2	8.5E-3	9.6E-2	1.2E-2	0.09	0.000	***
Ethanol	24	2.2E-1	5.2E-3	1.7E-1	2.5E-2	2.2E-1	1.1E-2	1.13	0.121	
	96	4.0E-2	7.3E-4	3.7E+0	5.4E-1	5.5E+0	9.6E-1	0.01	0.000	***
	192	3.3E-2	6.0E-3	3.1E+0	3.7E-1	3.7E+0	5.0E-1	0.01	0.000	***
	264	2.9E-2	8.6E-3	4.5E-1	5.9E-2	1.1E+0	6.3E-2	0.04	0.004	**
Formic acid	24	5.3E-4	1.1E-4	5.6E-4	3.0E-5	5.2E-4	2.7E-5	0.99	0.930	
	96	4.9E-4	2.7E-5	9.8E-4	2.1E-5	7.6E-4	1.4E-4	0.57	0.001	**
	192	6.7E-4	1.4E-4	1.8E-3	3.5E-4	2.0E-3	9.2E-5	0.36	0.000	***
	264	6.2E-4	3.2E-5	7.2E-3	1.5E-3	2.3E-3	2.4E-4	0.13	0.017	*
Fructose	24	1.2E+0	2.3E-2	8.8E-1	1.3E-1	1.2E+0	4.0E-2	1.11	0.221	
	96	3.2E-1	5.9E-3	3.3E-2	1.1E-2	1.3E-1	2.3E-2	4.02	0.000	***
	192	1.2E-1	3.6E-2	2.9E-2	2.8E-3	1.7E-2	1.7E-3	5.31	0.038	*
	264	2.9E-2	7.5E-3	1.7E-2	2.5E-3	1.2E-2	4.4E-4	2.00	0.069	

[1] n=3.
[2] *p<0.05, **p<0.01, ***p<0.001 by Welch's t-test.
[3] Ratio=average intensity of samples in July/average intensity of samples October and November.

Table 1. Continued

| | | Signal intensity relative to DSS-d_6[1] | | | | | | Comparative analysis[2] | | |
| | | July | | October | | November | | July vs. October & November | | |
Annotated compounds	hr	Mean	SD	Mean	SD	Mean	SD	Ratio[3]	p-value	
Fumaric acid	24	1.7E-4	5.1E-5	3.2E-4	6.8E-5	3.4E-4	7.7E-5	0.53	0.011	*
	96	3.2E-4	1.2E-4	3.9E-4	2.1E-4	2.6E-4	8.4E-5	0.97	0.932	
	192	1.9E-4	6.2E-5	4.7E-4	9.3E-5	2.6E-4	8.1E-5	0.51	0.032	*
	264	3.3E-4	9.1E-5	2.8E-4	1.8E-4	4.0E-4	1.5E-4	0.97	0.913	
Galacturonic acid	24	3.2E-4	1.2E-4	4.1E-4	1.5E-4	4.0E-4	8.3E-5	0.78	0.353	
	96	7.2E-4	7.6E-5	5.0E-3	1.0E-3	3.9E-4	1.2E-4	0.27	0.123	
	192	8.1E-4	1.2E-4	5.4E-3	1.5E-3	5.7E-4	1.6E-4	0.27	0.117	
	264	7.0E-5	5.6E-5	1.7E-3	7.3E-4	6.9E-4	1.2E-4	0.06	0.012	*
Glucose	24	5.5E-1	1.0E-2	3.9E-1	5.9E-2	5.5E-1	2.0E-2	1.16	0.111	
	96	1.2E-1	3.2E-2	9.2E-3	3.8E-3	1.1E-2	3.3E-3	12.44	0.024	*
	192	9.6E-2	2.8E-2	9.0E-3	4.5E-4	4.1E-3	7.4E-4	14.66	0.030	*
	264	8.1E-3	7.9E-4	5.6E-3	1.0E-3	5.2E-3	2.7E-4	1.50	0.009	**
Glutamine	24	1.4E-2	6.1E-4	1.3E-2	1.8E-3	1.9E-2	1.1E-3	0.86	0.195	
	96	2.0E-3	3.4E-4	3.2E-3	1.2E-3	6.1E-3	3.7E-4	0.43	0.012	*
	192	1.5E-3	3.4E-4	1.5E-3	1.2E-4	1.5E-3	1.4E-4	1.04	0.774	
	264	1.5E-3	5.2E-4	2.7E-3	8.4E-4	2.6E-3	2.5E-4	0.56	0.032	*
Gluconic acid	24	4.0E-2	2.5E-3	6.2E-2	9.9E-3	6.7E-2	6.2E-4	0.62	0.000	***
	96	3.6E-1	6.0E-2	2.0E-2	3.8E-3	2.8E-2	3.1E-3	15.05	0.010	*
	192	2.4E-1	8.5E-2	1.8E-2	5.6E-3	1.5E-2	2.2E-3	14.45	0.044	*
	264	5.0E-2	3.7E-3	1.1E-2	2.4E-3	1.2E-2	7.1E-4	4.23	0.001	**
Glycerol	24	3.6E-2	1.2E-3	2.5E-2	4.0E-3	3.6E-2	3.8E-3	1.19	0.111	
	96	1.3E-2	4.5E-4	7.9E-2	1.2E-2	1.4E-1	2.0E-2	0.12	0.001	**
	192	1.2E-2	3.3E-3	6.1E-2	1.2E-2	9.5E-2	1.4E-2	0.15	0.001	***
	264	6.4E-3	9.1E-4	2.6E-2	2.3E-3	6.6E-2	5.3E-3	0.14	0.007	**
Isoleucine	24	4.4E-4	3.2E-5	4.6E-4	1.4E-4	5.7E-4	4.7E-5	0.86	0.202	
	96	6.3E-4	1.4E-4	1.1E-3	5.2E-4	6.4E-4	7.1E-5	0.72	0.232	
	192	1.5E-3	2.7E-4	1.1E-3	2.1E-4	5.7E-4	3.0E-4	1.76	0.028	*
	264	3.8E-3	6.1E-4	1.6E-3	5.8E-4	1.1E-3	1.0E-4	2.80	0.007	**
Isoleucic acid	24	2.1E-4	5.5E-5	3.3E-4	3.1E-4	2.1E-4	3.2E-5	0.79	0.557	
	96	4.4E-4	8.9E-5	3.4E-4	2.7E-4	3.1E-4	1.5E-4	1.36	0.252	
	192	5.5E-4	2.4E-4	4.5E-4	1.1E-5	3.9E-4	1.1E-4	1.33	0.428	
	264	3.1E-3	8.1E-4	3.1E-3	6.5E-4	4.3E-3	7.9E-5	0.84	0.348	
Lactic acid	24	9.8E-3	1.7E-3	2.5E-2	3.3E-3	5.3E-2	5.6E-3	0.25	0.006	**
	96	2.5E-1	7.9E-2	1.6E-1	3.4E-2	3.0E-1	4.9E-2	1.06	0.806	
	192	5.1E-1	1.4E-1	1.8E-1	2.5E-2	1.8E-1	1.8E-2	2.78	0.055	
	264	4.2E-1	5.4E-2	1.8E-1	3.8E-2	2.0E-1	2.2E-2	2.21	0.010	**
Leucic acid	24	7.8E-4	1.4E-4	6.0E-4	1.7E-4	1.1E-3	1.4E-4	0.94	0.728	
	96	1.2E-3	5.9E-5	5.3E-4	1.3E-4	9.7E-4	1.7E-4	1.56	0.012	*
	192	1.9E-3	6.5E-4	5.8E-4	1.6E-4	1.3E-3	2.2E-4	2.02	0.104	
	264	6.5E-3	2.0E-3	1.9E-3	4.3E-4	3.8E-3	1.4E-4	2.31	0.066	
Malic acid	24	1.3E-2	6.2E-4	3.4E-3	5.3E-4	5.9E-3	2.5E-4	2.88	0.000	***
	96	5.4E-4	2.1E-4	3.0E-3	6.8E-4	2.0E-3	3.4E-4	0.21	0.001	***
	192	6.5E-4	3.7E-5	3.3E-3	1.1E-3	1.8E-3	4.9E-4	0.25	0.010	**
	264	5.5E-4	1.1E-4	2.1E-3	3.7E-4	1.5E-3	9.5E-5	0.31	0.001	***
Mannitol	24	1.6E-1	5.3E-3	1.1E-1	1.1E-2	1.6E-1	8.2E-3	1.15	0.165	
	96	9.2E-1	6.5E-2	3.0E-2	4.9E-3	5.1E-2	9.0E-3	22.55	0.002	**
	192	1.1E+0	3.4E-1	2.1E-2	2.3E-3	2.5E-2	3.8E-3	49.77	0.029	*
	264	1.2E+0	3.4E-1	1.6E-2	2.3E-3	2.1E-2	1.6E-3	63.53	0.028	*
Methanol	24	2.8E-2	8.7E-3	8.6E-2	5.7E-2	2.6E-2	9.8E-3	0.51	0.236	
	96	1.5E-2	1.4E-3	4.5E-2	6.4E-3	9.6E-3	2.9E-3	0.56	0.198	
	192	1.1E-2	2.4E-3	4.5E-2	5.1E-3	1.1E-2	1.4E-3	0.40	0.079	
	264	1.4E-2	3.1E-3	3.3E-2	3.5E-3	1.9E-2	5.0E-4	0.54	0.017	*
myo-Inositol	24	1.2E-2	5.5E-4	8.4E-3	9.9E-4	1.4E-2	6.1E-4	1.02	0.872	
	96	8.0E-3	3.2E-4	5.7E-3	1.4E-3	1.0E-2	7.1E-4	1.01	0.916	
	192	7.7E-3	2.1E-3	7.1E-3	2.0E-3	1.1E-2	1.6E-3	0.85	0.441	
	264	1.2E-2	3.0E-3	1.0E-2	2.2E-3	1.7E-2	6.3E-4	0.86	0.452	

[1] n=3.
[2] *p<0.05, **p<0.01, ***p<0.001 by Welch's t-test.
[3] Ratio=average intensity of samples in July/average intensity of samples October and November.

Table 1. Continued

		Signal intensity relative to DSS-d_6[1]						Comparative analysis[2]		
		July		October		November		July vs. October & November		
Annotated compounds	hr	Mean	SD	Mean	SD	Mean	SD	Ratio[3]	p-value	
Ornithine	24	5.5E-4	1.4E-4	4.8E-4	1.1E-4	5.8E-4	1.2E-4	1.04	0.822	
	96	7.2E-4	1.6E-4	4.1E-4	1.4E-4	2.6E-4	4.1E-5	2.15	0.034	*
	192	8.1E-4	1.5E-4	3.6E-4	1.4E-4	2.2E-4	6.7E-5	2.81	0.009	**
	264	2.5E-3	4.2E-4	1.1E-3	4.5E-4	1.0E-3	4.8E-5	2.38	0.013	*
Phenylalanine	24	1.5E-3	1.5E-4	1.0E-3	1.3E-4	1.7E-3	2.0E-4	1.08	0.575	
	96	1.7E-3	5.2E-4	2.6E-3	1.4E-3	1.5E-3	4.5E-4	0.84	0.576	
	192	2.7E-3	6.3E-4	2.2E-3	3.6E-4	1.3E-3	4.6E-4	1.58	0.088	
	264	6.6E-3	9.4E-4	3.5E-3	1.5E-3	2.7E-3	2.1E-4	2.15	0.005	**
Phenyllactic acid	24	8.7E-5	1.6E-5	1.7E-4	6.4E-5	2.1E-4	8.4E-5	0.46	0.015	*
	96	5.3E-4	1.4E-5	3.0E-4	3.9E-5	2.6E-4	1.4E-4	1.93	0.001	***
	192	9.0E-4	2.8E-4	2.8E-4	8.1E-5	2.7E-4	5.9E-5	3.26	0.055	
	264	5.1E-3	1.2E-3	1.1E-3	3.4E-4	2.0E-3	8.5E-5	3.36	0.029	*
scyllo-Inositol	24	2.2E-2	1.1E-3	2.1E-2	3.1E-3	3.1E-2	1.8E-3	0.85	0.151	
	96	1.7E-2	2.2E-4	2.1E-2	3.7E-3	2.7E-2	3.9E-3	0.69	0.014	*
	192	1.7E-2	4.7E-3	2.1E-2	6.0E-3	3.0E-2	4.6E-3	0.67	0.074	
	264	2.4E-2	6.6E-3	2.9E-2	3.5E-3	4.2E-2	1.5E-3	0.68	0.076	
Sucrose	24	2.6E-3	3.9E-4	1.3E-3	4.7E-4	7.4E-3	3.9E-3	0.61	0.361	
	96	9.1E-4	1.1E-4	9.7E-4	5.3E-4	3.2E-3	1.9E-3	0.44	0.159	
	192	7.0E-4	3.6E-4	1.3E-3	9.8E-4	7.3E-4	5.8E-4	0.70	0.449	
	264	5.8E-4	1.9E-4	2.3E-4	1.6E-4	3.5E-4	1.0E-4	2.02	0.096	
Succinic acid	24	6.4E-3	5.6E-4	7.1E-3	1.4E-3	5.7E-3	4.5E-4	1.00	0.989	
	96	1.7E-2	3.0E-3	2.9E-2	7.9E-3	4.6E-2	5.2E-3	0.45	0.004	**
	192	2.2E-2	5.5E-3	2.0E-2	1.4E-3	1.5E-2	1.2E-3	1.24	0.317	
	264	2.4E-2	5.4E-3	2.5E-2	5.7E-3	3.2E-2	6.3E-3	0.85	0.356	
Tartaric acid	24	1.1E-2	3.5E-4	7.9E-3	1.2E-3	1.4E-2	4.2E-4	1.00	0.993	
	96	7.6E-3	2.7E-4	7.5E-3	1.3E-3	1.2E-2	1.4E-3	0.77	0.108	
	192	7.7E-3	2.5E-3	8.1E-3	2.4E-3	1.4E-2	2.9E-3	0.71	0.185	
	264	1.0E-2	2.9E-3	1.1E-2	1.7E-3	1.7E-2	1.2E-3	0.71	0.131	
Theobromine	24	3.6E-4	4.7E-5	2.3E-4	1.1E-4	1.3E-4	4.3E-5	2.01	0.007	**
	96	4.4E-3	3.6E-4	3.6E-3	9.8E-4	2.1E-3	2.8E-4	1.53	0.015	*
	192	6.4E-3	1.8E-3	4.2E-3	7.5E-4	4.4E-3	9.0E-4	1.50	0.174	
	264	2.2E-2	2.4E-3	1.4E-2	9.4E-4	1.5E-2	1.1E-3	1.52	0.025	*
Trehalose	24	1.6E-3	1.1E-4	1.2E-3	2.2E-4	1.6E-3	3.3E-4	1.19	0.126	
	96	2.3E-3	3.8E-4	4.0E-3	7.2E-4	3.4E-3	5.8E-4	0.61	0.005	**
	192	1.2E-3	5.4E-4	5.3E-3	3.4E-3	1.0E-2	4.1E-3	0.15	0.014	*
	264	1.4E-3	4.9E-4	1.7E-2	4.8E-3	2.1E-2	5.7E-3	0.07	0.000	***
Tyrosine	24	4.2E-4	2.8E-5	5.0E-4	8.4E-5	5.6E-4	5.0E-5	0.79	0.013	*
	96	4.3E-4	1.4E-4	7.1E-4	2.4E-4	4.6E-4	1.3E-4	0.74	0.246	
	192	6.7E-4	1.4E-4	6.3E-4	1.7E-4	2.6E-4	1.4E-4	1.50	0.128	
	264	1.9E-3	1.8E-4	6.0E-4	1.9E-4	4.5E-4	1.5E-4	3.53	0.001	***
Valine	24	8.5E-4	4.3E-5	8.6E-4	1.1E-4	1.2E-3	1.5E-4	0.83	0.109	
	96	1.7E-3	2.2E-4	2.9E-3	1.2E-3	2.0E-3	1.1E-4	0.69	0.091	
	192	3.3E-3	7.1E-4	2.6E-3	5.3E-4	1.4E-3	3.7E-4	1.61	0.069	
	264	7.2E-3	1.5E-3	3.7E-3	1.1E-3	2.7E-3	2.5E-4	2.29	0.026	*
Unannotated 2.80 ppm	24	1.9E-3	7.1E-5	1.1E-3	3.6E-5	2.0E-3	3.7E-4	1.21	0.193	
	96	1.0E-3	9.7E-5	3.7E-3	3.8E-4	2.6E-3	2.7E-4	0.33	0.001	***
	192	1.5E-3	4.9E-4	5.1E-3	1.6E-3	8.0E-3	1.4E-3	0.22	0.001	**
	264	1.8E-3	4.2E-4	1.3E-2	1.7E-3	1.8E-2	3.6E-4	0.12	0.000	***
Unannotated 2.42 ppm	24	4.5E-3	1.3E-4	4.0E-3	6.4E-4	6.1E-3	2.9E-4	0.90	0.368	
	96	8.0E-4	6.2E-5	2.8E-3	4.5E-4	3.0E-3	4.1E-4	0.27	0.000	***
	192	9.5E-4	4.1E-4	4.1E-3	1.4E-3	6.2E-3	1.1E-3	0.18	0.001	***
	264	6.4E-4	1.6E-4	1.1E-2	1.3E-3	1.4E-2	2.3E-4	0.05	0.000	***

[1] n=3.
[2] *p<0.05, **p<0.01, ***p<0.001 by Welch's t-test.
[3] Ratio=average intensity of samples in July/average intensity of samples October and November.

Table 1. Continued

| | | Signal intensity relative to DSS-d_6[1] | | | | | | Comparative analysis[2] | |
| | | July | | October | | November | | July vs. October & November | |
Annotated compounds	hr	Mean	SD	Mean	SD	Mean	SD	Ratio[3]	p-value
Unannotated 1.38 ppm	24	8.9E-5	4.5E-5	5.8E-5	1.3E-5	9.9E-5	1.2E-4	1.13	0.815
	96	2.1E-3	1.6E-4	2.6E-3	1.0E-4	1.7E-3	1.6E-4	1.00	0.971
	192	1.2E-3	3.5E-4	5.0E-3	1.6E-3	7.0E-3	1.2E-3	0.20	0.001 ***
	264	1.1E-3	2.7E-4	1.3E-2	1.9E-3	1.9E-2	2.5E-4	0.06	0.000 ***
Unannotated 0.88 ppm	24	8.8E-4	5.4E-5	6.1E-4	1.8E-4	1.0E-3	7.6E-5	1.07	0.634
	96	5.0E-4	2.0E-4	8.6E-3	1.3E-3	9.9E-3	1.5E-3	0.05	0.000 ***
	192	8.2E-4	9.3E-5	8.1E-3	1.5E-3	9.1E-3	8.8E-4	0.09	0.000 ***
	264	2.3E-3	5.5E-4	7.9E-3	1.3E-3	9.5E-3	4.4E-4	0.27	0.000 ***

[1] n=3.
[2] *p<0.05, **p<0.01, ***p<0.001 by Welch's t-test.
[3] Ratio=average intensity of samples in July/average intensity of samples October and November.

plantarum and *P. kudriavzevii*, dominant species in samples from October and November, was correlated with ethanol, trehalose, unannotated (1.38 ppm), mannitol, and gluconic acid, with correlation coefficients of R=0.834, 0.734, 0.650, −0.747, and −0.782 for *L. plantarum* and R=0.844, 0.749, 0.754, −0.728, and −0.532 for *P. kudriavzevii*, respectively. The relative dominance of *L. plantarum* was also strongly correlated with the intensity of an unannotated signal (0.88 ppm; R=0.829). *L. fermentum* exhibited a strong correlation with acetoin (R=0.777). Although their correlations were not strong (R <|0.600|), it is important to note that *A. pasteurianus*, *L. fermentum*, and *G. oxydans* showed positive correlations with the acetic acid level, with correlation coefficients of R=0.443, 0.545, and 0.471, respectively. These correlations between the microbiota and metabolite profiles in the present study suggested a relationship between microbial metabolism in cocoa bean fermentation and compositional differences among the different preparation months.

DISCUSSION

In the terms of cell count levels, intra-factory cocoa fermentation was unstable in the three different months examined herein. The largest difference observed among the three tests was in the levels of AAB. Their levels peaked at 96 hr, at approx. 10^8 to 10^9 CFU/g, and were maintained until 264 hr in July; AAB were not isolated in October, and their levels reached approx. 10^9 CFU/g after 264 hr in November. The metagenomic analysis fully agreed with these results. AAB were dominant, with a relative abundance of >90% after 96 hr of fermentation in July. They were not detected in October, and they were only detected after 264 hr in November, showing a relative abundance of >20% (the sum of *Acetobacter* spp. and *G. oxydans*; Fig. 3). AAB are one of the major microbial groups during cocoa fermentation [23], and not detecting/isolating them is unusual. They generally start to grow after yeasts and LAB, which occurs after approx. 48 hr of fermentation [3]. AAB are aerobes that require oxygen for growth, and the aeration resulting from the turning of cocoa heaps enhances the growth of AAB [24]. In the present study, the plastic boxes used for fermentation were closed with lids, which may have affected the supply of oxygen for fermentation. However, the boxes were opened every day, and the samples were mixed for aeration. Therefore, this would not be a major factor contributing to the unstable population of AAB in the present study. Storage in a refrigerator for more than 60 hr during fermentation may have resulted in a bottleneck for the growth of AAB. The temperature within a heap sometimes increases to between 40 and 50°C in the stage of AAB fermentation [25, 26]. AAB are responsible for acetate production mainly from the metabolism of carbohydrates and the oxidation of ethanol produced by yeasts [3]. Acetate levels markedly increased during fermentation in July but not during October or November (Fig. 6), although the levels of AAB also increased at the end of fermentation in November. Acetate itself affects the taste of processed cocoa products (chocolate). It also forms acetate esters via a reaction with alcohols, which are flavour-active compounds in processed products [27]. Acetate also indirectly contributes to the formation of several flavour precursors, by decomposing cocoa bean cotyledons in combination with an increased temperature and other microbial metabolites [3]. Therefore, unstable levels of AAB and acetate are concerns in cocoa fermentation.

The levels of yeasts and LAB were unstable at the beginning of fermentation (24 hr) but generally stabilized after 48 hr of fermentation in the three tests. After 24 hr, their levels had already reached 10^7 to 10^9 CFU/g in July, but they were lower than 10^6 CFU/g in October and November. The metagenomic analysis showed consistent results, and the two microbial groups had very minor populations (relative abundance <1%) at 24 hr in October and November. *Enterobacteriaceae* aerobes, mainly comprising *Tatumella* spp. and *Pantoea* spp., were dominant in the microbiota at 24 hr in October and November, as confirmed by culturing and metagenomic approaches, but disappeared after 96 hr. These aerobes are plant symbionts that are also sometimes found at the beginning of cocoa fermentation [28–30] and are generally regarded as minor contributors to cocoa fermentation [4]. Therefore, we concluded that fermentation in October and November was delayed for unidentified reasons. Since the fermentation conditions were stable in the temperature- and humidity-controlled machine used in the present study, other factors, including autochthonous microbes on pod surfaces, affected the fermenting microbiota. In spontaneous fermentation, environmental factors (such as the weather and temperatures at the fermentation sites as well as microbial contamination from insects and workers' hands during fermentation), the fermentation method, geographic location, and degree of mixing are promising

factors for the development of different microbiota [2, 31], and they were uniform in the present study. Camu *et al.* compared the cocoa microbiota between on-farm heap fermentation and intra-factory heap fermentation, with the heaps covered with banana leaves in both cases, to examine the impact of environmental contamination and found no marked differences in the diversity of LAB and AAB between the two types of fermentation [24]. Therefore, they considered cocoa pod surfaces, not the general environment, to be the main inoculum for spontaneous cocoa bean heap fermentation. The microbiota of cocoa pod surfaces may be affected by environmental factors, particularly different temperatures and precipitation amounts, in the semitropical area of Japan examined; however, this was not characterized in the present study. Plant phenology has been shown to affect microbial community profiles on plant surfaces [32, 33].

Although the fermentation conditions markedly differ between on-farm fermentation and machine-controlled intra-factory fermentation, the major microbial species observed in the present study were those reported for on-farm fermentation in tropical countries. The microbiota was relatively simple during intra-factory fermentation, and only one or two species were observed as the major fermenting microbes in each group of yeast, LAB, and AAB. The Shannon indices markedly decreased between 24 and 96 hr of fermentation (Fig. 4A). *P. kudriavzevii*, which was the major yeast observed in the three tests, is the 2nd most prevalent yeast in spontaneous on-farm fermentation in some countries, including Australia, Brazil, Cuba, Ecuador, Ivory Coast, and Malaysia [1, 4]. It is a pectinolytic organism that breaks down cocoa pulp and is tolerant of the high temperatures and ethanol observed in cocoa fermentation [34]. *S. cerevisiae* is the most frequently observed yeast species in spontaneous cocoa fermentation in most countries [1, 4], but it was not detected in the present study. Ethanol production is one of the key roles of *S. cerevisiae* in fermentation, and ethanol accumulation was positively correlated with *P. kudriavzevii* in the present study.

L. fermentum and *L. plantarum* are two representative LAB species in spontaneous cocoa fermentation in several countries [1], and they were also observed during intra-factory fermentation. They did not coexist during fermentation. Although they both belong to the LAB group, they share different metabolic properties [35]. *L. plantarum* is a homofermentative organism that mainly produces lactate from glucose metabolism, while *L. fermentum* is a heterofermentative that produces lactate, ethanol/acetate, and CO_2 from glucose. Heterofermentative LAB sometimes possess properties for the conversion of fructose to mannitol by using fructose as an external electron acceptor [36], and this property is conserved in *L. fermentum* [35]. Mannitol accumulated in *L. fermentum*-enriched fermentation in July but not in *L. plantarum*-enriched fermentation in October or November; however, a positive correlation with mannitol was observed for *A. pasteurianus* but not for *L. fermentum*. This may be due to the excessive levels of relative abundance of *A. pasteurianus* at 96, 192, and 264 hr in July, as demonstrated by metagenomics. The culturing technique showed that *L. fermentum* and *A. pasteurianus* had similar CFU levels in July (Fig. 2A). Therefore, the excessive abundance of *A. pasteurianus*, which was >90%, in fermentation during July may have been due to differing efficiencies of DNA extraction between the microbes. The efficiency of DNA isolation is generally higher for Gram-negative bacteria than for Gram-positive bacteria because of the different integrities

of their cell walls [37–39], and this is sometimes a bottleneck for a sequencing-based analysis. *A. pasteurianus* is known as a mannitol consumer in cocoa fermentation [31]. The abundance of this microbe is also correlated with the accumulation of acetoin and gluconic acid, which are metabolites produced from glucose and lactate, respectively, by the organism [40].

HCA of the metagenomic data revealed that the microbiota markedly changed between 24 and 96 hr of fermentation to an *A. pasteurianus*-enriched microbiota in July and *P. kudriavzevii-L. plantarum*-enriched microbiota in November and December (Figs. 3 and 4). The microbiota did not show marked changes between 96 and 192 hr, which may have been due to the storage of fermenting samples in a refrigerator. The differentiated microbiota had a marked impact on the compounds that accumulated or were consumed by fermenting microbes. The three fermentation periods showed similar metabolic profiles at 24 hr but were separated into two groups, July fermentation and October and November fermentation, after 96 hr (Fig. 5). This separation was due to different levels of increases or decreases in more than 10 compounds, including gluconic acid, acetoin, acetic acid, mannitol, lactic acid, ethanol, 2,3-butanediol (1,2-propanediol), trehalose, and amino acids. Some of these compounds in fermented cocoa have been identified as factors that affect the flavour and taste of chocolate. For example, several amino acids are associated with bitterness and a cocoa flavour, and mannitol is one of the compounds that correlated with bitterness, astringency, and the cocoa flavour in chocolate [41, 42]. The marked differences in metabolite profiles among the fermentation periods indicated the unstable quality of the fermented cocoa beans. These results suggest the necessity of starter microbes for intra-factory fermentation. The application of starter microbes has been examined in detail in on-farm cocoa fermentation, which is generally due to the difficulties associated with controlling the environmental factors that affect the microbiota [9, 43, 44]. Further studies are needed on the impact of starter microbes on machine-controlled intra-factory fermentation.

CONFLICT OF INTEREST

The authors report no conflicts of interest.

REFERENCES

1. Figueroa-Hernández C, Mota-Gutierrez J, Ferrocino I, Hernández-Estrada ZJ, González-Ríos O, Cocolin L, Suárez-Quiroz ML. 2019. The challenges and perspectives of the selection of starter cultures for fermented cocoa beans. Int J Food Microbiol 301: 41–50. [Medline] [CrossRef]

2. Schwan RF, Wheals AE. 2004. The microbiology of cocoa fermentation and its role in chocolate quality. Crit Rev Food Sci Nutr 44: 205–221. [Medline] [CrossRef]

3. De Vuyst L, Leroy F. 2020. Functional role of yeasts, lactic acid bacteria and acetic acid bacteria in cocoa fermentation processes. FEMS Microbiol Rev 44: 432–453. [Medline] [CrossRef]

4. Díaz-Muñoz C, De Vuyst L. 2022. Functional yeast starter cultures for cocoa fermentation. J Appl Microbiol 133: 39–66. [Medline] [CrossRef]

5. Moens F, Lefeber T, De Vuyst L. 2014. Oxidation of metabolites highlights the microbial interactions and role of *Acetobacter pasteurianus* during cocoa bean fermentation. Appl Environ Microbiol 80: 1848–1857. [Medline] [CrossRef]

6. Yamada Y, Yukphan P. 2008. Genera and species in acetic acid bacteria. Int J Food Microbiol 125: 15–24. [Medline] [CrossRef]

7. Lisdiyanti P, Kawasaki H, Seki T, Yamada Y, Uchimura T, Komagata K. 2001. Identification of *Acetobacter* strains isolated from Indonesian sources, and proposals of *Acetobacter syzygii* sp. nov., *Acetobacter cibinongensis* sp. nov., and *Acetobacter orientalis* sp. nov. J Gen Appl Microbiol 47: 119–131. [Medline] [CrossRef]

8. Schwan RF, Vanetti MCD, Silva Do, Lopez A, De Moraes CA. 1986. Characterization and distribution of aerobic, spore-forming bacteria from cacao fermentations in Bahia. J Food Sci 51: 1583–1584.

9. Lefeber T, Papalexandratou Z, Gobert W, Camu N, De Vuyst L. 2012. On-farm implementation of a starter culture for improved cocoa bean fermentation and its influence on the flavour of chocolates produced thereof. Food Microbiol 30: 379–392. [Medline] [CrossRef]

10. 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: 84. [Medline] [CrossRef]

11. de C Lima CO, Vaz ABM, De Castro GM, Lobo F, Solar R, Rodrigues C, Martins Pinto LR, Vandenberghe L, Pereira G, Miúra da Costa A, Benevides RG, Azevedo V, Trovatti Uetanabaro AP, Soccol CR, Góes-Neto A. 2021. Integrating microbial metagenomics and physicochemical parameters and a new perspective on starter culture for fine cocoa fermentation. Food Microbiol 93: 103608. [Medline] [CrossRef]

12. Verce M, Schoonejans J, Hernandez Aguirre C, Molina-Bravo R, De Vuyst L, Weckx S. 2021. A combined metagenomics and metatranscriptomics approach to unravel Costa Rican cocoa box fermentation processes reveals yet unreported microbial species and functionalities. Front Microbiol 12: 641185. [Medline] [CrossRef]

13. Endo A, Koizumi R, Nakazawa Y, Shiwa Y, Maeno S, Kido Y, Irisawa T, Muramatsu Y, Tada K, Yamazaki M, Myoda T. 2021. Characterization of the microbiota and chemical properties of pork loins during dry aging. MicrobiologyOpen 10: e1157. [Medline] [CrossRef]

14. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46 W1: W537–W544. [Medline] [CrossRef]

15. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20: 257. [Medline] [CrossRef]

16. Lu J, Breitwieser FP, Thielen P, Salzberg SL. 2017. Bracken: estimating species abundance in metagenomics data. PeerJ Comput Sci 3: e104. [CrossRef]

17. Youngblut ND, Ley RE. 2021. Struo2: efficient metagenome profiling database construction for ever-expanding microbial genome datasets. PeerJ 9: e12198. [Medline] [CrossRef]

18. Tomita S, Watanabe J, Kuribayashi T, Tanaka S, Kawahara T. 2021. Metabolomic evaluation of different starter culture effects on water-soluble and volatile compound profiles in *nozawana* pickle fermentation. Food Chem (Oxf) 2: 100019. [Medline]

19. Tomita S, Nemoto T, Matsuo Y, Shoji T, Tanaka F, Nakagawa H, Ono H, Kikuchi J, Ohnishi-Kameyama M, Sekiyama Y. 2015. A NMR-based, non-targeted multistep metabolic profiling revealed L-rhamnitol as a metabolite that characterised apples from different geographic origins. Food Chem 174: 163–172. [Medline] [CrossRef]

20. Zhao S, Guo Y, Sheng Q, Shyr Y. 2014. Advanced heat map and clustering analysis using heatmap3. BioMed Res Int 2014: 986048. [Medline] [CrossRef]

21. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Berlin: Springer.

22. Caligiani A, Acquotti D, Cirlini M, Palla G. 2010. 1H NMR study of fermented cocoa (Theobroma cacao L.) beans. J Agric Food Chem 58: 12105–12111. [Medline] [CrossRef]

23. Soumahoro S, Ouattara HG, Droux M, Nasser W, Niamke SL, Reverchon S. 2020. Acetic acid bacteria (AAB) involved in cocoa fermentation from Ivory Coast: species diversity and performance in acetic acid production. J Food Sci Technol 57: 1904–1916. [Medline] [CrossRef]

24. Camu N, González A, De Winter T, Van Schoor A, De Bruyne K, Vandamme P, Takrama JS, Addo SK, De Vuyst L. 2008. Influence of turning and environmental contamination on the dynamics of populations of lactic acid and acetic acid bacteria involved in spontaneous cocoa bean heap fermentation in Ghana. Appl Environ Microbiol 74: 86–98. [Medline] [CrossRef]

25. Camu N, De Winter T, Verbrugghe K, Cleenwerck I, Vandamme P, Takrama JS, 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. [Medline] [CrossRef]

26. Papalexandratou Z, Kaasik K, Kauffmann LV, Skorstengaard A, Bouillon G, Espensen JL, Hansen LH, Jakobsen RR, Blennow A, Krych L, Castro-Mejía JL, Nielsen DS. 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. [Medline] [CrossRef]

27. Meersman E, Steensels J, Struyf N, Paulus T, Saels V, Mathawan M, Allegaert L, Vrancken G, Verstrepen KJ. 2015. Tuning chocolate flavor through development of thermotolerant saccharomyces cerevisiae starter cultures with increased acetate ester production. Appl Environ Microbiol 82: 732–746. [Medline] [CrossRef]

28. Yang WT, Yi YJ, Xia B. 2023. Unveiling the duality of *Pantoea dispersa*: a mini review. Sci Total Environ 873: 162320. [Medline] [CrossRef]

29. Papalexandratou Z, Vrancken G, De Bruyne K, Vandamme P, De Vuyst L. 2011. 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. [Medline] [CrossRef]

30. Torres AR, Araújo WL, Cursino L, Hungria M, Plotegher F, Mostasso FL, Azevedo JL. 2008. Diversity of endophytic enterobacteria associated with different host plants. J Microbiol 46: 373–379. [Medline] [CrossRef]

31. De Vuyst L, Weckx S. 2016. The cocoa bean fermentation process: from ecosystem analysis to starter culture development. J Appl Microbiol 121: 5–17. [Medline] [CrossRef]

32. Ginnan NA, De Anda NI, Campos Freitas Vieira F, Rolshausen PE, Roper MC. 2022. Microbial turnover and dispersal events occur in synchrony with plant phenology in the perennial evergreen tree crop *Citrus sinensis*. MBio 13: e0034322. [Medline] [CrossRef]

33. Vermote L, Verce M, Mozzi F, De Vuyst L, Weckx S. 2022. Microbiomes associated with the surfaces of Northern Argentinian fruits show a wide species diversity. Front Microbiol 13: 872281. [Medline] [CrossRef]

34. Samagaci L, Ouattara H, Niamké S, Lemaire M. 2016. *Pichia kudrazevii* and *Candida nitrativorans* are the most well-adapted and relevant yeast species fermenting cocoa in Agneby-Tiassa, a local Ivorian cocoa producing region. Food Res Int 89: 773–780. [Medline] [CrossRef]

35. Adler P, Bolten CJ, Dohnt K, Hansen CE, Wittmann C. 2013. Core fluxome and metafluxome of lactic acid bacteria under simulated cocoa pulp fermentation conditions. Appl Environ Microbiol 79: 5670–5681. [Medline] [CrossRef]

36. Zaunmüller T, Eichert M, Richter H, Unden G. 2006. Variations in the energy metabolism of biotechnologically relevant heterofermentative lactic acid bacteria during growth on sugars and organic acids. Appl Microbiol Biotechnol 72: 421–429. [Medline] [CrossRef]

37. Auer GK, Weibel DB. 2017. Bacterial cell mechanics. Biochemistry 56: 3710–3724. [Medline] [CrossRef]

38. Nadkarni MA, Martin FE, Hunter N, Jacques NA. 2009. Methods for optimizing DNA extraction before quantifying oral bacterial numbers by real-time PCR. FEMS Microbiol Lett 296: 45–51. [Medline] [CrossRef]

39. de Bruin OM, Birnboim HC. 2016. A method for assessing efficiency of bacterial cell disruption and DNA release. BMC Microbiol 16: 197. [Medline] [CrossRef]

40. Sainz F, Jesús Torija M, Matsutani M, Kataoka N, Yakushi T, Matsushita K, Mas A. 2016. Determination of dehydrogenase activities involved in D-glucose oxidation in gluconobacter and acetobacter strains. Front Microbiol 7: 1358. [Medline] [CrossRef]

41. Kitani Y, Putri SP, Fukusaki E. 2022. Investigation of the effect of processing on the component changes of single-origin chocolate during the bean-to-bar process. J Biosci Bioeng 134: 138–143. [Medline] [CrossRef]

42. Afoakwa EO, 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. [Medline] [CrossRef]

43. Schwan RF. 1998. Cocoa fermentations conducted with a defined microbial cocktail inoculum. Appl Environ Microbiol 64: 1477–1483. [Medline] [CrossRef]

44. Crafack M, Keul H, Eskildsen CE, Petersen MA, Saerens S, Blennow A, Skovmand-Larsen M, Swiegers JH, Petersen GB, Heimdal H, Nielsen DS. 2014. Impact of starter cultures and fermentation techniques on the volatile aroma and sensory profile of chocolate. Food Res Int 63: 306–316. [CrossRef]