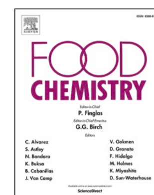
Food Chemistry 365 (2021) 130627

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Fine cocoa beans production: Tracking aroma precursors through a comprehensive analysis of flavor attributes formation

Sebastián Escobar [a,*], Margareth Santander [a], Martha Zuluaga [a], Iván Chacón [a], Jader Rodríguez [a], Fabrice Vaillant [a,b,c]

[a] Corporación Colombiana de Investigación Agropecuaria (Agrosavia), Process & Quality Cocoa Laboratory. Centros de Investigación Palmira, Tibaitatá y La Selva – Km 14 Mosquera-Bogotá, Cundinamarca, P.O. Box 344300, Colombia

[b] Centre de Coopération Internationale en Recherche Agronomique pour le Développement- CIRAD, UMR QualiSud, 1101 avenue Agropolis, CS 24501, 34093. Montpellier Cedex 5, Francia

[c] UMR QualiSud, Univ Montpellier, CIRAD, Montpellier SupAgro, Univ Avignon, Univ La Reunion, Montpellier, France

ARTICLE INFO

Keywords:
Chocolate sensory profile
Chocolate volatile aroma compounds
Cocoa bean metabolomic fingerprint
Postharvest fermentation time
Fine flavor cocoa
Sensory quality

ABSTRACT

The fine flavor cocoa (FFC) market offers cocoa farmers better monetary and nonmonetary benefits than the bulk market. In this work, during cocoa fermentation, flavor formation was studied at different fermentation times based on sensory profiles, volatile compound contents and untargeted metabolomics. It was observed that chocolate quality is influenced by fermentation time. Thus, at 72 h, the sensory profiles showed no outstanding attributes, while at 96 h, the global quality presented a stronger influence of fine attributes, such as fruitiness, florality, spices and nuttiness. Finally, at 120/144 h, these FFC features diminished. Metabolomic fingerprint of cocoa beans (related to peptides, sugars, amino acids, and phenolic compounds) and the volatile fingerprint of chocolate showed a change according to the fermentation time. This allowed the proposal of 96 h as the optimal fermentation time to produce FFC beans. Additionally, 20 volatiles and 48 discriminating metabolites were defined as potential quality biomarkers.

1. Introduction

In general, bulk and fine flavor cocoa (FFC) can be distinguished from each other by their flavor quality. While FFC is characterized by its outstanding aromatic notes such as flowery, fruity, caramelly and nutty, bulk cocoa does not have these same notes, and it has a marked basic cocoa taste (Ríos et al., 2017). The FFC market offers very large economic and social benefits for cocoa farmers compared to the bulk market (Ríos et al., 2017). Therefore, it is imperative for cocoa farmers produce cocoa with high and stable quality to participate in the FFC market.

Cocoa fermentation is essential for chocolate flavor formation. During this stage, the endogenous components of cocoa seeds, mainly proteins and carbohydrates, are transformed into simpler compounds, called flavor precursor metabolites, because during roasting, they react to generate the flavor molecules that are expressed as sensory attributes in chocolate (Santander et al., 2019). In short, depending on the fermentation method, the sensory quality of cocoa beans varies due to the difference in terms of the metabolites generated. Therefore, one of the goals for the cocoa sector is to define fermentation methodologies that allow the production of cocoa beans with higher quality to make chocolates with finer sensory attributes.

On the other hand, farmers empirically ferment cocoa beans without considering product quality, applying their own traditional method and using any kind of fermentation equipment. Additionally, cocoa farmers in Latin America have cocoa plantations averaging 0.5 to 1 ha with an approximate yield of 600 kg/ha/year, which represents noncompetitive production capacity and difficulties when negotiating prices. Additionally, this situation results in cocoa with heterogeneous characteristics between each fermentation batch. This creates a gap that prevents the production of high-quality cocoa being competitive and ensuring regularity in the quality of the cocoa offered (Abbott et al., 2017).

Operation under an associative model may be an alternative to facilitate cocoa farmers to carry out the transformation of seeds to cocoa beans in a central location with the same equipment under unified protocols based on producing the best quality FFC products (Abbott et al., 2017). This may increase competitiveness and ensure higher participation in the FFC market.

Currently, research on cocoa quality has focused mainly on

* Corresponding author.
E-mail address: sescobar@agrosavia.co (S. Escobar).

https://doi.org/10.1016/j.foodchem.2021.130627
Received 22 April 2021; Received in revised form 12 July 2021; Accepted 15 July 2021
Available online 19 July 2021
0308-8146/© 2021 Elsevier Ltd. All rights reserved.

postharvest processes and/or flavor formation, especially from a chemical point of view, considering the precursor metabolites in the cocoa beans and the volatile compounds in chocolate (D'Souza et al., 2018; Megias-Perez et al., 2020; Utrilla-Vázquez et al., 2020). Although those studies contributed considerably to elucidate the flavor formation, they have missed the connection with the real sensorial impact on cocoa quality within the true context of "on-farm" production and failed to generate appropriate postharvest practices.

A study on cocoa flavor formation should be more comprehensive and comprise the dynamics of the "on-farm" transformation from seeds to cocoa beans, considering a holistic approach that links analytical methods with reliable sensorial analysis. This is the objective of this paper, which aimed to analyze the variation in the metabolomic fingerprint of cocoa beans during "on-farm" fermentation, the volatile profile and the sensory attributes of chocolate produced with these beans.

2. Materials and methods

2.1. Large-scale postharvest cocoa transformation process

On-farm postharvest cocoa bean transformation was carried out in the three main cocoa-producing regions in the southern Colombian Pacific located in the district of Buenaventura Harbour: Cisneros (Cis), Bajo Calima (Bcal) and Sabaletas (Sab). The same equipment and methodology were used in each region by implementing a central postharvest cocoa transformation unit. Stages such as the harvesting of the pods from the farms, transportation of the cocoa pods to the central postharvest transformation site, fermentation and drying strictly followed the same guidelines in all three regions (Fig. S1. A-H). For this, a total of 32 small cocoa farmers (with cocoa farms between 0.5 and 1 ha) in each region participated in the process, contributing to the mixture of cocoa materials available on their farms, all of them belonging to universal and Colombian hybrid commercial cultivars. Cocoa pods were harvested during the peak harvesting season in November 2019. The cocoa pods were transported to the central transformation site the day after being harvested.

2.1.1. Fermentation of cocoa seeds

Fermentation was developed as reported by Santander, Vaillant, Sinuco, Rodríguez, & Escobar (2021) with some modifications. The cocoa pods were opened, and the cocoa seeds were manually extracted. All of the extracted cocoa seeds were mixed. Then, 300 kg from the total mass of cocoa seeds obtained was distributed in 6 batches of 50 kg each. Each batch of cocoa seeds was placed in independent compartments of the fermenter system shown in Fig. S1. G-H. Thus, per region, the six fermentation batches were considered 6 replicates that represent a mix of cocoa seeds from different farms. The first aeration of the cocoa mass was performed manually after 48 h with a shovel, moving the mass of seeds from one compartment to another empty compartment. Then, the cocoa mass was mixed every 12 h until 120 h of processing was reached in the case of Bajo Calima and Sabaletas; however, this time was extended to 144 h for Cisneros due to a particular reason in this region, and it was not possible to take the sample at 120 h.

2.1.2. Cocoa sampling and drying

All cocoa seeds were sampled from the center of the mass of each of the 6 compartments of the fermenter system at 72, 96 and 120/144 h. All samples were stored at −18 °C. For the metabolomic analysis, 50 g of fermented cocoa was sampled and dried by forced convection with hot air at 50 °C for 24 h (Rodriguez-Campos et al., 2012). Additionally, a sample of 1,500 g of cocoa seeds was collected and subsequently dried by forced convection with hot air at 50 °C for 72 h to produce chocolate and perform the sensory and volatile analyses.

2.2. Chocolate sensory analysis

Chocolates from the cocoa fermented for 72, 96 and 120/144 h from the three regions with 67% (w/w) cocoa solids were produced. These samples were obtained under a standard protocol in the chocolate-making laboratory. The fermented and dried cocoa beans were roasted at 125° C for 60 min. The roasted cocoa nibs were subsequently deshelled. 67% cocoa liquor was mixed with 33% sugar before the refining stage. A three-roll refiner was used. The refined chocolate mass was conched for 6 h at 80 °C with the addition of 5% cocoa butter and 0.01% soy lecithin (as emulsifier). The tempering stage consisted of initial heating of the chocolate above 40 °C, then cooling to 29° C and finally heating to 32 °C before being molded.

A descriptive sensory analysis was performed using the consensus profile method (International Organization for Standardization, 2016) in the sensory analysis laboratory. The chocolate sensory analysis was developed according to Santander et al. (2021). The chocolate samples were blindly tasted by eight trained judges with experience evaluating chocolate (more than five years). Each sample was evaluated twice (one replication). The panel consisted of three females and five males, aged between 25 and 50 years old. They were selected based on their abilities to identify and describe differences in chocolate samples and for having good reproducibility. Continuous panel performance and monitoring methods are applied to ensure all judges maintained a high level of competence.

There were two types of evaluation sessions. First, the judges agreed on a vocabulary with sensory attributes to evaluate the samples. 12 descriptors were selected to describe the flavor of chocolate samples. The descriptors were classified as basic flavors: sweetness, acidity, bitterness, astringency; fine flavors: cocoa taste, cocoa aroma, fruitiness, nuttiness, floral, spices, and off-flavors: burned and animal. Finally, global quality was also considered. In the second session, they were seated in individual booths, and the chocolates (5 g) were served at room temperature (at 20 ± 1 °C) on identical plastic plates pre-labeled with a random three-digit code. Samples were presented successively, one by one, in a monadic sequence presentation order. A completely randomized design was used for the sample presentation. Judges scored chocolate samples using a 10-point scale from 0 (not perceivable) to 10 (very intense) regarding the sensory attributes defined initially. For neutralization between samples, unsalted crackers and non-carbonated water (at 37 °C) were available.

For analysis of the results from the sensorial panel, Cronbach's alpha coefficient (CA) was measured to analyze the internal consistency and agreement among the panelists (or reproducibility). Table S1 shows that there was a high concordance between all judges. Subsequently, a formative model of partial least squares structural equation modeling (PLS-SEM) was performed in order to describe the contribution of the attributes to the sensory profile. The sensory attributes, as measured variables, were grouped to generate three latent variables as "flavor-type descriptors" (Fig. S2. A). The latent variable "Fine" referred to the sensory attributes of fruitiness, florality, nuttiness, cocoa aroma, cocoa taste, and spices. The latent variable "basic" referred to the chocolate's basic flavors that include acidity, sweetness, astringency, and bitterness. The latent variable "off" refers to deleterious flavors, which are animal and burned. Moreover, global quality was also measured, and the variable "quality" was defined. The basic, fine, and off flavors influenced the global quality (Fig. S2. B). Finally, collinearity based on the variation inflation factor (VIF), the coefficient of determination (R^2), and the significance of each construct in the model (p-value) were indicators of the performance of the model (Hair et al., 2019). This analysis was performed using R v.3.6.0 with the package "semPLS". Additionally, the details about PLS-SEM methodology are available in the Supplementary material.

2.3. Analysis of volatile organic compounds in chocolate

The volatile organic compounds (VOCs) of chocolate samples (2.5 g) were extracted using solid phase microextraction in the headspace (SPME-HS) as described by Assi-Clair et al. (2019), using a 50/30-μm divinylbenzene/carboxene/polydimethylsiloxane (DVB/CAR/PDMS) fiber provided by Supelco. The extraction conditions were previously optimized by combining the exposure time of the fiber at different temperatures. The fiber was pre-conditioned in the chromatograph injector at 250 °C for 3 min and then, to reach equilibrium, it was exposed to the chocolate sample headspace at 50 °C for 45 min.

The volatile compounds were analyzed using an Agilent 6890 N gas chromatography–mass spectrometry (GC–MS) instrument equipped with a Hewlett Packard capillary column DBWAX, 30 m length × 0.25 mm internal diameter × 0.25 μm film thickness (Palo Alto, CA, USA). The GC oven temperature was initially set at 40 °C for 5 min, increased to 140 °C at a rate of 2 °C/min and then increased at a rate of 10 °C/min to 250 °C for 66 min. The carrier gas was high-purity helium at 1 ml min^{-1}. The splitless injection mode was at 250 °C (2 min). The selective mass detector was a quadrupole (Hewlett Packard, Model 5973), with an electronic impact ionization system at 70 eV and at 230 °C (Assi-Clair et al., (2019). VOCs were tentatively identified using the three criteria previously defined by Assi-Clair et al. (2019): (i) comparison of the retention index with those in an in-house database, (ii) probability-based matching of mass spectra with those obtained from the Wiley 275L library of mass spectra, and (iii) in some cases, identification was confirmed using standards of the components. For each chocolate sample, three replicates were carried out. The dispersion coefficients in the headspace and absorption on the fiber were neglected. The moisture content of the chocolates was determined using the oven-drying method according to the International Organization for Standards (ISO 1666:1996) method. The content of each volatile compound was calculated according to Equation (1) (Assi-Clair et al., 2019).

$$q_i(\mu g g^{-1}) = \frac{25 \times A_i}{A_{but} \times m_e \times W} \tag{1}$$

where q_i is the amount of compound i, A_i is the area of compound i, A_{but} is the area of butanol (standard), 25 is the content of butanol in mg kg^{-1}, m_e is the mass of sample introduced into the vial in g and W is water content of the sample (Assi-Clair et al., 2019). Odor activity values (OAVs) were calculated by dividing the detected headspace concentration by the respective odor threshold value (OTV) to estimate the contribution of the volatiles to the overall chocolate aroma. The OTVs in oil media used were from van Gemert (2011).

A principal component analysis of the concentrations of the volatile organic compounds was performed to visualize the behavior of the volatiles on chocolates derived from cocoa beans processed at the three fermentation times. The data were mean-centered and autoscaled. Then, orthogonal partial least squares discriminant analysis (OPLS-DA) was applied by group comparison between 96 h of fermentation and the other fermentation times (72, 120 and 144 h) in order to obtain discriminating VOC quality markers. The discriminant power of each VOC was evaluated by the variable importance in projection (VIP) score. The first 20 VOCs with VIP > 1.0 were considered relevant (Efenberger-Szmechtyk et al., 2018). The quality of the model was evaluated considering its fitting capacity (R^2Y) value, the prediction capacity (Q^2Y), the cross-validated analysis of variance (CV-ANOVA) and the permutation test to evaluate overfitting of the data. The statistical analysis was performed using the software R Studio.

2.4. Untargeted metabolomic analysis in fermented and dried cocoa beans

The extracts derived from fermented and dried cocoa beans were obtained through a procedure described by Santander et al. (2021). A UPLC-ESI-QTOF-MS system (Acquity, Waters, Milford, Massachusetts, USA) was used. Chromatographic separation was performed using a CSH C_{18} column (130 Å, 1.7 μm particle size, 2.1 mm × 100 mm, Waters) coupled to a Vanguard CSH C_{18} precolumn (130 Å, 1.7 μm particle size, 2.1 mm × 5 mm, Waters). The method and conditions of analysis were developed in the same way as previously described by Santander et al. (2021).

For the ESI-MS/MS analyses, an electrospray ionization source and a QTOF analyzer (Waters) were used. Positive ionization mode under a full scan from 100 to 1500 Da was employed in continuous mode. The conditions of the ESI + -QTOF-MS system were the same as those of the protocol of Santander et al. (2021).

For the data processing, Raw files were collected in continuous mode and centered with the lock mass using MassLynx software V3.1 SCN 639 (Waters Inc., Milford, Massachusetts, USA). Then, these data were processed using Progenesis QI software (Waters Inc.), with which the chromatographic peaks were selected and aligned. Subsequently, the effects of the drift on the quality control baseline were corrected using the Metabodrift tool (Thonusin et al., 2017). For this, the data were normalized using the LOESS method with an alpha of 0.5, and the ions were filtered with a relative standard deviation>30% in the quality control samples. The data were treated with the dilution factor that relates the mass of the ground cocoa and the resuspended mass of the sample.

Data tidying was complemented by removing the masses of ions that had a value of zero in 90% of the samples. Then, these zeros were replaced with the minimum value detected in order to represent the baseline signal. Additionally, the logarithmic transformation gLog and autoscaling were applied to the data.

For metabolomic multivariate analysis, a hierarchical cluster analysis (HCA) using the correlation distance followed by multiscale bootstrapping resampling was carried out to differentiate the metabolomic profiles between the time of fermentation and the region. The analysis included the confidence estimates of the clusters based on the Approximately Unbiased p-value (AU-P) and Bootstrap Probability p-value (BP-P) Second, OPLS-DA was performed to identify discriminating metabolites in the fermentation process and associate them to the quality. For this, an independent comparison was made between the different times of fermentation. It was compared 72 h, 96 h, 120 h/144 h among the treatments. Then, R^2Y, Q^2Y, and Cross-Validation test was measured to evaluate the model performance. Finally, the 20 features with the highest VIP were chosen with the aim to be identified. The multivariate analysis was carried out using R with the package “ropls”.

Once the discriminating metabolites were identified by multivariate analysis, MS/MS fragmentation was performed. The mass of each parent ion was compared with the *Theobroma cacao* peptide database published by D'Souza et al. (2018) with a tolerance of 10 ppm, from which the primary structure of the peptide was determined. Subsequently, this peptide was translated into SMILES nomenclature using the PepSMI online tool and an in-house web scraping script. Then, *in silico* fragmentation of the peptide was performed using the CFM-ID 3.0 online tool. Finally, the experimental and in silico fragmentation patterns were compared using our own R script. The tolerance used for each ion was 10 ppm, and the peptides that matched >5 ions were annotated. Additionally, peptides that remained unknown were identified using SearchGUI 4.0.8 based on the methodology described by (D'Souza et al., 2018). Finally, the metabolites were identified using the *Theobroma cacao* library in the Plant Metabolic Network database (https://plantcyc.org/) by comparison of the MS/MS reported spectra or by *in silico* fragmentation.

3. Results and discussion

3.1. Sensory profile variations in the chocolate produced by cocoa beans sampled at different fermentation times

The results of the sensory analysis performed by a highly specialized

and experienced panel of judges were deemed highly reliable (Cronbach's alpha statistic coefficient > 0.9, see details in Table S1). Fig. 1. A-C shows the sensory profiles of the chocolate samples obtained after 72, 96 and 120/144 h of spontaneous fermentation in the three regions of the Colombian Pacific: Cisneros, Bajo Calima and Sabaletas. The most important aspect to be highlighted is that the sensory profile of these chocolates showed the typical characteristics of a fine aroma with relatively high perceptions of the sensory attributes of fruitiness, nuttiness, florality, and spices.

For each region, significant differences between the attributes of the chocolate flavor profile as a function of fermentation time were found (Fig. S3. A-I). In general, an increase in the intensity of the fine attributes (fruitiness, nuttiness, florality and spices) was observed between 72 and 96 h (Fig. S3. A, D, G). After 96 h, the sensory perception of almost all of these attributes tended to decrease. Notably, the chocolate samples from each of the three regions had the highest scores in global quality at 96 h. Therefore, in all regions the fine attributes and global quality are correlated and appear to show the same dynamic. In terms of undesirable quality attributes, astringency and acidity perception were mostly present in the two chocolate samples from Bcal and Sab after 120 h of fermentation (Fig. S3. E, H). The same results were obtained for the chocolate samples from Bcal, Sab and Cis after 72 h, but these samples were also linked to the attribute of being burned (burnt-toasted) (Fig. S3. C, F,I).

A structural equation model was built by PLS-SEM (Fig. 1) to objectively explain the influences of the sensory profile attributes on global quality. The model quantified the weights of the three latent variables that encompass basic, fine, and off-flavor attributes at different process times. The model quality indicators were statistically acceptable. The R^2 values of each coefficient for each construct were > 0.4 (Fig. 1. D-F) and the noncollinearity was validated with the inner VIF < 3 (Tables S2-S3 and Fig. S4) and outer VIF (Table S4) (Hair et al., 2019). The p-value calculated for the general model in which processing time was modeled showed that the *basic* and *fine flavors* contribute significantly to the quality. An extended explanation of the PLS-SEM model and the quality indicators are presented in the section "Extended methodology of PLS-SEM for sensory quality" in the Supplementary data.

The data from the three regions were used to evaluate the PLS-SEM model's behavior based on the time of fermentation (Fig. 1. D-F). The results showed that the vectors from the *basic* and *off-flavors* were negatively correlated with *quality*, indicating that when the sensory perception of the *basic* or *off* attributes is high, the perception of *quality* is lower. On the other hand, the *fine* variable weight was positive, indicating that the higher the perception of this attributes was, the greater the perception of the *quality* of the chocolate would be. As previously observed at 72 h, the *basic* vector had a greater absolute value than that of either the *fine* or *off* vectors, indicating a greater influence of the former group of attributes on chocolate quality. The contribution of the *fine* flavors at 72 h was lower than the absolute values of both *basic* and *off*. Therefore, after 72 h of fermentation, chocolate presents a stronger *basic* flavor perception. At 96 h, the higher absolute value of the weight vector was for the *fine* flavors, which also had a positive sign indicating a direct relationship, while the *off* flavors had a very low and nonsignificant contribution. Thus, for the chocolate obtained at 96 h, the *fine* attributes are those that have the most significant influence on *quality*. At 120/144 h, the absolute values of the *basic* and *fine* flavor vectors were equal, indicating that the perception of *quality* at this processing time was influenced by both flavors, with the *off* flavors showing no influence. Additionally, at all time points, the variable *off* had no significant contribution to quality in our case. Therefore, in all cases, 96 h was the optimal time of fermentation. At 72 h, the weight of the *basic* vector was more negative than at 96 and 120/144 h and at 96 h, the *fine* vector had a greater contribution, with the chocolate after 120/144 h of processing being influenced by both *fine* and *basic* flavors.

When the model was run separately for each region and fermentation time, equivalent behavior was observed (Fig. S5). In all regions at 96 h, the perception of fine attributes had a greater influence on the *quality*

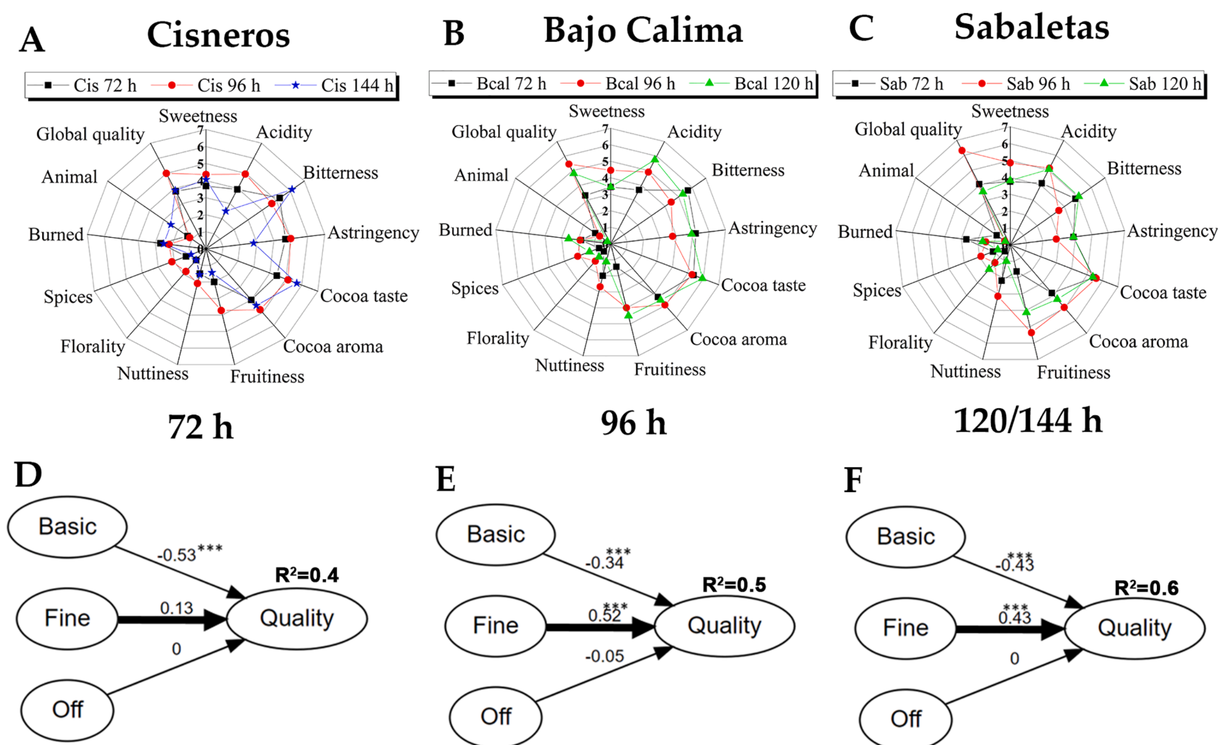


Fig. 1. Sensory profiles of chocolates made by fermented cocoa at 72, 96 and 120/144 h, of mixtures of cocoa materials from: **A)** Cisneros, **B)** Bajo Calima, and **C)** Sabaletas. PLS-SEM formative approach of cocoa flavor for all regions at different fermentation times: **D)** 72 h, **E)** 96 h and **F)** 120/144 h. *** p-Value < 0.005. The vector of *fine* variable is highlighted in order to emphasize the change of fine attributes over time.

followed by 120/144 h, while at 72 h, there was a greater influence from the basic attributes.

3.2. Volatile organic compounds

A total of 83 volatile organic compounds (VOCs) were identified, and their semiquantitative headspace concentrations in chocolate samples are reported in Table S5. Each VOC was characterized by aroma descriptors, as reported in the scientific literature. To analyze the differences between the processing time points according to the families of volatile compounds detected in the chocolate samples, PCA was performed by loading the concentrations of all detected VOCs as variables and the fermentation time points as the observations. Fig. 2 shows clear differentiation at all fermentation time points (72, 96, and 120/144 h) in each region based on the first two principal components of the PCA model. This indicates that the VOC profile in the derived chocolate varies depending on the fermentation time. These findings demonstrate that there is a dynamic change in the composition of the cocoa seeds as the fermentation time progresses.

The odor activity values (OAVs) of the volatiles were calculated to indicate whether a compound contributes to the flavor profile, as well as to link and highlight their relationship with the findings of the sensory analysis of the chocolate. The OAV results for the majority of the volatiles are given in Table S6. A volatile compound with an OAV ≥ 1 may be considered odor-active, which suggests that it is relevant for the aroma of chocolate.

Table 1 shows the discriminating 20 volatiles that were defined by the OPLS-DA model by comparing 96 h with the other fermentation times. All validity tests were deemed statistically acceptable (Fig. S6; $R^2Y = 0.949$, $Q^2Y = 0.914$, and for the permutation test $n = 1000$, pR^2Y and pQ^2Y were below 0.05). From these data, we found that 2 volatiles, namely, 3-methyl butanal and 2-phenylethyl acetate, which were also reported by Magagna et al. (2017) and Utrilla-Vázquez et al. (2020), were considered as key aroma markers since their presence was fundamental for the chocolate aroma properties, as they are related to desirable aromatic notes such as fruity, flowery and cocoa-chocolate. Additionally, 2-furanmethanol, furfural and 2-phenyl-but-2-enal, which are associated with the fine flavor characteristics of flowery, sweet, chocolate and caramel, were discriminating, and these VOCs are of special interest because they are considered technological markers due to their susceptibility to change during the processing stages (Magagna et al., 2017).

The discriminating volatiles were monitored according to fermentation time (72, 96, 120/144 h) to visualize the most characteristic volatiles from each treatment. This may contribute to defining the aromatic quality of chocolates made by fermented cocoa after the three processing times. The volatile compounds associated with desirable aroma descriptors (i.e., fruity, flowery, chocolate and almond), such as hexanal, benzeneacetaldehyde, 3-methyl butanal, furfural, 2-phenyl-but-2-enal, 2-acetylpyrrole, benzaldehyde, 5-methyl furfural, 2-furanmethanol, nonan-2-ol, ethyl octanoate, benzyl alcohol, nonan-2-one, *cis*-linalool oxide and 2-phenylethyl acetate, were found predominantly in chocolates after 96 h, and the majority of these compounds exhibited their highest OAVs at that point during fermentation. They could be classified as high-quality key aroma markers. In contrast, ethanol is a low-quality key aroma marker because it is related to the undesirable alcoholic aroma that was mainly present in the chocolate samples from 72 h. These findings confirm that after 96 h of fermentation, there is a significant improvement in the development of several key fine cocoa aroma volatiles associated with higher quality.

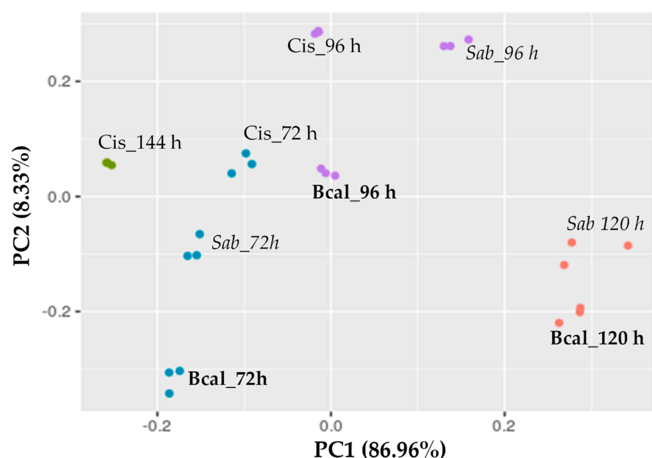


Fig. 2. PCA-score plot from volatile profile data of chocolates samples produced from 72 (blue points), 96 (violet points) and 120 (red points) / (green points)144 h fermented beans in Cisneros (Cis), Bajo Calima (Bcal) and Sabaletas (Sab) regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Regarding volatile acids, acid perception in chocolates (regardless of the region) increased as the fermentation process progressed to 120 h (Fig. S3. B, E, H). This perception was confirmed by the increasing concentration of residual acetic acid in the chocolates (showing a significant OAV > 1 in all samples; Table S6) as fermentation progressed. This undesirable volatile that gives a sour and vinegar-like aroma (Frauendorfer & Schieberle, 2019) is mostly volatilized during roasting, but evidently some traces remained in the chocolate bars. Additionally, the highest amount of isobutyric acid (Table S5) was observed in the chocolate samples from the 144 h-fermented beans. This compound is mostly linked with overfermentation and generates hammy and putrid off flavors (Rodriguez-Campos et al., 2012). Isobutyric acid is correlated with the highest score for animal descriptors in the chocolate from Cis 144 h (Fig. S3. C).

High concentrations of some alcohols are favorable in order to obtain cocoa products with flowery, fruity and sweet flavors (Rodriguez-Campos et al., 2012). Heptan-2-ol, characterized by citrus notes, is an active aroma that was predominant in the Bcal and Sab chocolate samples at 96 h. 2-Phenylethanol, reported to have an aroma with flowery and spice notes, also stands out due its significant contribution to the overall flavor in Cis 96 h (Table S6). Furthermore, linalool is a terpene alcohol that is also relevant and may be mainly responsible for fruity and flowery aromas due to its considerable aroma activity (OAV > 1) in the chocolates after 96 h of fermentation.

Esters and ketones are of special interest because they are associated with fruity and flowery aromas. As seen in Table S5, the contents of these groups of aromatic compounds increase as the processing time progresses (Barišić et al., 2019). The highest scores for fruity and floral attributes in the chocolate produced from beans fermented for 96 h and 120 h (Fig. S3. A, D, G) are related to aromatically active esters (OAV > 1), such as methyl acetate, ethyl acetate, 2-phenylethyl acetate, butyl acetate and isoamyl benzoate. Actually, these esters may particularly affect chocolate aroma quality due to the induction of desirable flavor aromatic notes such as fruity, winey, flowery and rose (Nielsen et al., 2005). These results were similar to those previously reported by Assi-Clair et al. (2019), who suggested that four fermentation days, instead of six fermentation days, significantly improved the development of key fruity aroma compounds. In addition, despite the unknown odor threshold in oil, the most abundant ester was 2,3-butanediol monoacetate, which was previously reported in dark chocolate samples (Deuscher et al., 2019).

The fruity and floral sensory attributes specifically detected in the chocolates samples from 96 h and 120 h were also linked to the highest production of ketones at these processing times (Table S5 and Fig. S3. A, D, G). Our findings are in accordance with the previous results of Assi-Clair et al. (2019) and Barišic (2019), who reported that the main ketones present in chocolate are heptan-2-one, pentan-2-one, nonan-2-one, acetophenone, and acetoin. Nonan-2-one was present in a considerable concentration that influences the overall aroma, imparting a

Table 1
Discriminating key-aroma compounds.

VOC	VIP	Aroma descriptor	Relative concentration*								
			Cis_72	Bcal_72	Sab_72	Cis_96	Bcal_96	Sab_96	Bcal_120	Sab_120	Cis_144
Ethanol	2.20	Alcoholic	0.32	0.74	0.52	-2.07	0.13	-1.24	0.84	1.04	-0.29
Hexanal	2.16	Green, tallowy, leaf-like	-0.36	-0.42	-0.48	2.13	0.73	0.48	-1.04	-1.26	0.22
Benzeneacetaldehyde	2.14	Sweet	-1.13	-0.71	-0.59	0.74	1.14	1.49	-0.27	-0.57	-0.10
3-Methyl butanal	2.09	Malty, chocolate, cocoa	-1.23	-1.40	-0.65	0.71	0.79	1.60	-0.63	0.05	0.75
Furfural	1.93	Caramel, burnt, ethereal, almond	-0.77	-0.83	-1.82	1.36	0.43	1.14	-0.04	0.33	0.21
Dimethylsulfide	1.80	Vegetable, tomato, garlic	0.65	0.36	-0.68	-0.31	-0.05	-2.06	1.78	0.22	0.08
2-Phenyl but-2-enal	1.79	Flowery, cocoa, sweet, roasted, chocolate	-0.63	-1.42	-1.17	0.20	0.75	1.71	-0.26	1.06	-0.25
2-Acetylpyrrole	1.74	Chocolate, hazelnut	-0.73	-0.58	-0.45	0.14	0.98	1.73	-0.06	0.71	-1.75
Benzaldehyde	1.67	Bitter, almond, grass, chocolate, cocoa	-0.86	-1.47	-0.67	-0.24	0.78	1.96	-0.27	0.76	0.02
5-Methyl furfural	1.63	Almond, burnt sugar, caramel	-0.09	-0.57	-1.18	0.45	0.44	1.71	0.21	0.73	-1.71
Isobutyl acetate	1.61	Fruit, apple, banana	1.27	-0.41	-1.04	-0.64	-1.16	-0.88	1.03	0.42	1.40
2-Furanmethanol	1.53	Sweet, alcoholic, bitter	-0.86	0.06	0.75	-0.24	0.84	1.82	-0.87	0.08	-1.58
Nonan-2-ol	1.35	Fruity	0.39	-1.08	-1.84	1.17	1.28	-0.08	0.65	-0.60	0.11
Pentanol	1.35	Pungent	0.98	0.06	0.01	1.53	0.61	-0.14	-1.33	-1.63	-0.09
Ethyl octanoate	1.33	Fruity, flowery	0.90	-1.53	-1.72	1.50	0.37	0.13	0.27	-0.05	0.13
Benzyl alcohol	1.32	Sweet. flowery	-0.78	-1.49	-0.88	-0.28	0.76	1.63	0.61	0.92	-0.49
Nonan-2-one	1.27	Fruity	1.00	-1.47	-1.44	1.45	0.78	-0.22	0.16	-0.59	0.32
(E)-3-Penten-2-one	1.27	Fruity	0.73	0.23	-0.03	-1.46	-0.17	-0.19	0.23	-0.19	0.85
Cis-Linalool oxide	1.27	Sweet, flowery, earthy, woody	0.75	-1.26	-1.41	0.79	0.36	0.45	-0.61	0.34	0.59
2-Phenylethyl acetate	1.22	Rose, honey, tobacco, flowery, fruity, sweet	0.73	-1.47	-1.62	1.10	0.46	0.37	0.70	0.64	-0.91

*Colour intensity goes from green (minimum) to red (maximum).

fruity flavor due to a high OAV (Table S6). Previous works have also reported that nonan-2-one highly contributes to fruity aromatic notes (Rottiers et al., 2019; Tuenter et al., 2020).

The aldehyde group-containing carbonylic compounds are also vital for the development of a desirable cocoa aroma. Ten aldehydes were identified, and certain increases of these compounds were observed in chocolate samples from the seeds fermented for 72 to 96 h in all regions. Strecker aldehydes such as benzeneacetaldehyde and benzaldehyde, obtained from 2-phenylalanine and phenylalanine, respectively, exhibit sweet and fruity notes, and their contribution to the overall flavor is likely preponderant due to their OAVs > 1 (Table S6). 2-Phenyl-but-2-enal is considered the key component of chocolate aroma related to chocolate and floral aromatic notes (Wang et al., 2019).

The presence of benzaldehyde, 2-methyl propanal (derived from valine), 3-methyl butanal (from leucine), and 2-methyl butanal (from isoleucine), which are three Strecker aldehydes present at higher concentrations in the chocolates from cocoa beans fermented for 96 h, confer a strong chocolate and cocoa flavor, as observed in Fig. S3. A, D, G (Barišić et al., 2019). The results are in line with the findings reported by Menezes et al. (2016), who obtained a high concentration of the aldehydes 2-methylbutanal and 2-phenyl-2-butenal that favor chocolate quality, and nonanal that confers the cocoa with fruity notes.

Pyrazines are the key type of nitrogenous heterocyclic volatiles and the main components forming the cocoa basic aroma. They exhibit cocoa, earthy, nutty and roasted aromas, so these volatiles are desirable in chocolate (Tuenter et al., 2020). The total concentration of pyrazines increased as fermentation time progressed (Table S5). Notably, the beans fermented for 96 h and 120 h were richer in pyrazines than those fermented for 72 h. These results are related to the cocoa aroma attributes of the sensory evaluation, as a slight increase in this perception was observed from 72 h to 96 h in samples from all three regions (Fig. S3. A, D, G). The major pyrazine detected was 2,3,5-trimethylpyrazine, which is aromatically active (OAV > 1). Similar results were obtained by Barišić et al., (2019) and Hinneh et al. (2018). They reported that trimethylpyrazine is one of the most important compounds for the development of nutty flavor. In contrast, tetramethylpyrazine was not odor-active, and its contribution to flavor was negligible due to its low OAV < 1.

The contents of furans, furanones and pyrroles increased until 96 h of fermentation (Table S5) and then drastically decreased. As they are characterized by desirable caramel and sweet-like flavors, this behavior was correlated with a slightly higher sweetness perception in the chocolate samples at 96 h (Fig. S3. B, E, H). The most representative compounds of this group, due their higher concentrations in chocolate samples, were furfural, 5-methyl furfural, 2-furanmethanol, 2-acetylpyrrole and 2,3-dihydro-3,5-dihydroxy-6-methyl-1-4H-pyran-4-one (Table S5). The presence of these same furans was reported in chocolates in previous works (Hinneh et al., 2020; Tuenter et al., 2020), and specifically, 2-acetylpyrrole has been reported as a key odorant of dark chocolate due to its powerful aromatic nature (Deuscher et al., 2020).

Our findings confirm that the chocolates produced from cocoa seeds fermented for 96–120 h present volatiles that are highly correlated with good-quality aroma descriptors. Consequently, better scores for the intensities of the desirable sensory attributes and overall quality were observed at these fermentation times. The decrease in fine flavor sensory attributes before and after this time window could be explained by the changes in the amounts and/or types of precursors of volatile compounds formed during fermentation, which react during the roasting stage when Maillard reactions take place (Assi-Clair et al., 2019). The hypothesis related to the quantity of flavor precursors was suggested by Marseglia, Musci, Rinaldi, Palla, & Caligiani (2020), who stated that a higher total sum of volatile compounds was obtained from higher amounts of flavor precursors as peptides and free amino acids. Likewise, other studies have reported that VOCs are also influenced by the biological activity of microorganisms during spontaneous fermentation (Assi-Clair et al., 2019).

3.3. Metabolomic fingerprint and discriminating metabolites as potential process biomarkers

Hierarchical cluster analysis (HCA) was applied to visualize patterns that showed the changes and differentiation in the chemical constituents

of the cocoa beans according to fermentation time. As seen in Fig. 3, the first separation showed clear discrimination between the group of metabolites that conformed to the metabolomic fingerprint of the 72 h treatments on the main left branch, while those located on the main right branch were composed of those from the 96 and 120/144 h treatments. These results are in accordance with the sensory analysis, where there was also a clear distinction between the 72 h treatment and the 96 and 120/144 h treatments. Afterward, OPLS-DA was applied to the left branch built by the 72 h treatments versus the right branch composed of the other treatments to determine the differential metabolites that could be potential biomarkers associated with quality. The potential fermentation biomarkers were selected using the VIP from the discriminant analysis, choosing the twenty biomarkers with the highest VIP scores.

In Table 2, we report the 20 features with the highest VIP scores as potential biomarkers for the 72 h treatment, which were not differentiated by fine attributes but were rated as intermediate-low quality. Then, the 20 features with the highest VIP scores for the 96 h treatment were rated as potential biomarkers of high quality, according to sensory analysis (PLS-SEM), in which the fine attributes characterized the sensory profile. Additionally, the 120 h treatments were compared with the other treatments to extract the features with a greater VIP score, and these were classified as intermediate quality potential biomarkers. Finally, there was a specific cluster for Cisneros 144 h, which was the sample with the poorest quality in the sensory profile and was placed in an independent branch of the HCA.

From the eighty features statistically deemed to be discriminating metabolites (the most discriminant first twenty from each fermentation stage), forty metabolites were tentatively selected. The other forty features that were not selected are presented in Table S7. Thirty-three of the selected features corresponded to peptides with lengths varying between 2 and 12 residues. The remaining seven corresponded to other discriminating metabolites classified as products of peptides, hydroxy amino acids, one oligosaccharide, and one flavonoid (Table 2).

Peptides can be precursors of the aroma and taste of roasted cocoa beans and chocolate. The discriminant peptides found in this study have also been reported in other studies (D'Souza et al., 2018; Scollo et al., 2020). According to (D'Souza et al., 2018), cocoa aroma is mainly associated with methionine (M), leucine (L), valine (V), isoleucine (I), phenylalanine (F), and alanine (A) at the *N*-terminus because these amino acids are more reactive with the reducing sugars in the initial stages of the Maillard reaction. In our study, only at 96 h of fermentation did most of the peptides annotated present these amino acids at their *N*-terminus (first residue of the peptide). In addition, peptides with tyrosine (Y), glutamine (Q) and glycine (G) at the C-terminus are selectively substrates for exopeptidases, and these peptides are also present in higher concentrations at 96 h. On the other hand, peptides with threonine (T) and valine (V) at the C-terminus have been reported to be unstable with a short lifespan. In this work, these peptides were mostly found after 144 h of fermentation.

The length of the peptide also plays a major role because smaller peptides are potentially more reactive in the Maillard reaction, and it has been shown that these peptides have taste-modulating properties (Salger et al., 2019). Di- or tripeptides appear to exhibit higher intensities after 144 h of fermentation (WF, LT, CAM, IPH and the glutamyl peptide QIR), and they may have a role in the generation of off flavors. This hypothesis is aligned with the previous findings of Salger, Stark, & Hofmann (2019), who isolated and reported a large list of short peptides (2–5 amino acids) from overfermented cocoa beans that also showed attributes associated with off flavors for chocolate, such as umami (i.e., SPV), salty (i.e., pEEE, VP, SP, PS), bitter taste (i.e., PA, SP, RP, PP, and KP), and kokumi perception (i.e., VPA), which seems to enhance and prolong the mouthfeel of off-flavors.

It was also reported by D'Souza et al., (2018) that as fermentation time progressed, the length of the peptides decreased and the diversity of the peptides increased. In this study, we also found the same phenomenon with some of the peptides annotated as potential biomarkers. For example, after 96 h of fermentation, we found the larger peptide VDNIFNNPDESY. Then, at 120 h, we observed a product of its degradation, the shorter fragment IFNNPDESYF. Additionally, at 120 h, smaller peptides such as AM and AMF were present, and at 144 h, oxidized dimers of the amino acid alanine were found. These results are aligned with the findings presented by Kumari et al. (2018), who concluded that the peptide profile can represent the level of physicochemical transformation of cocoa seeds.

Some oligopeptides observed in this study have also been reported as relevant in previous works, including AGGGGL (from albumin) and FGVPSK and IFVPHYN (from vicilin). GAGGGGL was identified as a peptide of considerable interest because it is one of the 37 most discriminant metabolites due to its variation in the time of fermentation

Fig. 3. Cluster dendrogram from the Hierarchical Cluster Analysis of the metabolomics data.

Table 2
Discriminating metabolites as potential biomarkers of quality.

Assigned identity	Observed m/z	Adduct	Z	M	Formula	Theoretical m/z	Δm/z (ppm)	RT	VIP	Quality - Treatment	Relative intensity*				Ontology	Amino acid region
											72 h	96 h	120 h	144 h		
Peptides																
WF	334.1544	$[M+H]^+$	1	333.1477	$C_{20}H_{19}N_3O_2$	334.1550	1.8	12.41	1.79	Intermediate - low (72h)					Albumin	152-154
LT	264.1809	$[M+NH_4]^+$	1	247.1527	$C_{10}H_{21}N_3O_4$	264.1792	-6.4	8.17	1.77	Intermediate - low (72h)					Vicilin	253-255
SPG	331.1591	$[M+H]^+$	1	330.1539	$C_{13}H_{22}N_4O_6$	331.1612	6.3	0.86	1.75	Intermediate - low (72h)					Vicilin	457-462
CAM	170.0609	$[M+H+NH_4]^+$	1	323.0968	$C_{11}H_{21}N_3O_4S_2$	170.0619	5.9	7.5	1.74	Intermediate - low (72h)						
KLVD	237.6453	$[M+2H]^+$	2	473.2844	$C_{21}H_{39}N_5O_7$	237.6498		8.49	1.74	Intermediate - low (72h)					Vicilin	524-527
PFLEDK	748.3885	$[M+H]^+$	1	747.3796	$C_{35}H_{53}N_7O_{11}$	748.3876	1.2	11.65	1.73	Intermediate - low (72h)						
IPH	366.2116	$[M+H]^+$	1	365.2058	$C_{17}H_{27}N_5O_4$	366.2136	-5.5	4.5	1.72	Intermediate - low (72h)						
DNSAGKW	389.1776	$[M+2H]^+$	2	776.3447	$C_{33}H_{48}N_{10}O_{12}$	389.1799	-5.9	0.56	1.71	Intermediate - low (72h)					Albumin	128-134
FGVPSK	317.6811	$[M+2H]^+$	2	633.3481	$C_{30}H_{47}N_7O_8$	317.6816	-1.6	6.48	2.77	High (96h)					Vicilin	519-524
NHR	444.2291	$[M+NH_4]^+$	1	426.1969	$C_{16}H_{26}N_8O_6$	444.2314	5.2	8.33	2.56	High (96h)						
LPSVK	272.1773	$[M+2H]^+$	2	542.3422	$C_{25}H_{46}N_6O_7$	272.1787	-5.1	4.6	2.50	High (96h)						
AGGGGL	431.2284	$[M+H]^+$	1	430.2170	$C_{17}H_{30}N_6O_7$	431.2249	8.1	9.09	2.47	High (96h)					Albumin	54-59
LLC	394.2389	$[M+NH_4]^+$	1	377.2091	$C_{15}H_{31}N_5O_4S$	394.2357	-8.1	10.56	2.46	High (96h)						
LKY	212.1348	$[M+2H]^+$	2	422.2524	$C_{21}H_{34}N_4O_5$	212.1335	-6.2	3.84	2.36	High (96h)						
VDNIFNNPDESY	713.8104	$[M+2H]^+$	2	1425.6043	$C_{62}H_{87}N_{15}O_{24}$	713.8097	1.0	10.58	2.35	High (96h)					Vicilin	526-537
IFNNPDESYF	623.276	$[M+2H]^+$	2	1244.5344	$C_{58}H_{76}N_{12}O_{19}$	623.2748	1.9	10.92	2.26	Intermediate (120h)					Vicilin	529-538
FFF	460.2227	$[M+H]^+$	1	459.2153	$C_{27}H_{29}N_3O_4$	460.2231	0.9	11.47	2.25	Intermediate (120h)						
ETVFN	609.2822	$[M+H]^+$	1	608.2800	$C_{27}H_{40}N_6O_{10}$	609.2879	-9.4	8.03	2.24	Intermediate (120h)					Vicilin	292-296
ENSPPLK	392.7125	$[M+2H]^+$	2	783.4121	$C_{34}H_{57}N_9O_{12}$	392.7136	-2.8	4.39	2.22	Intermediate (120h)					Vicilin	171-177
LW	300.1699	$[M+H-H_2O]^+$	1	317.1734	$C_{17}H_{23}N_3O_3$	300.1706	2.3	11.75	2.22	Intermediate (120h)						
AMF	368.1632	$[M+H]^+$	1	367.1560	$C_{17}H_{25}N_3O_4S$	368.1639	-1.9	9.03	2.20	Intermediate (120h)					Vicilin	185-187
SYYGAF	707.3051	$[M+H]^+$	1	706.2957	$C_{35}H_{42}N_6O_{10}$	707.3035	2.3	10.14	2.20	Intermediate (120h)					Vicilin	281-286
AM	221.0962	$[M+H]^+$	1	220.0876	$C_8H_{16}N_2O_3S$	221.0954	3.6	0.86	2.18	Intermediate (120h)					Vicilin	185-187
VDNIFNNPDESYF	787.3448	$[M+2H]^+$	2	1572.6727	$C_{71}H_{96}N_{16}O_{25}$	787.3439	1.1	11.52	2.17	Intermediate (120h)					Vicilin	526-538
VT	201.1242	$[M+H-H_2O]^+$	1	218.1261	$C_9H_{18}N_2O_4$	201.1234	-4.0	6.42	2.17	Intermediate (120h)						
DNK	376.1857	$[M+H]^+$	1	375.1748	$C_{14}H_{25}N_5O_7$	376.1827	-8.0	5.41	2.16	Intermediate (120h)					Peroxidase	260-262
IFVPHYN	445.2357	$[M+2H]^+$	2	888.4488	$C_{44}H_{60}N_{10}O_{10}$	445.2320	8.3	6.46	2.15	Intermediate (120h)						
IFVPHYN	445.2357	$[M+2H]^+$	2	888.4488	$C_{44}H_{60}N_{10}O_{10}$	445.2320	8.3	6.46	2.80	Low (144h)					Vicilin	394-400
KDK	407.2629	$[M+NH_4]^+$	1	389.2269	$C_{16}H_{31}N_5O_6$	407.2613	-3.9	3.56	2.80	Low (144h)						
SQSPV	517.2757	$[M+H]^+$	1	516.2538	$C_{21}H_{36}N_6O_9$	517.2617		3.72	2.70	Low (144h)					Vicilin	354-358
LHV	184.6191	$[M+2H]^+$	2	367.2220	$C_{17}H_{29}N_5O_4$	184.6183	-4.5	3.42	2.69	Low (144h)						
PQP	341.1851	$[M+H]^+$	1	340.1741	$C_{15}H_{24}N_4O_5$	341.1819	-9.4	11.31	2.67	Low (144h)						
QRI	399.2340	$[M+H-H_2O]^+$	1	416.2378	$C_{17}H_{32}N_6O_6$	399.2350	2.5	4.49	2.60	Low (144h)					Vicilin	496-498
Other metabolites																
Hydroxyleucine	186.1286	$[M+K]^+$	1	147.0895	$C_6H_{13}NO_3$	186.1282	-2.1	7.5	1.73	Intermediate - low (72h)					130.07 (100); 131.07(48);147.09(1)	
A-type procyanidin dimer arabinoside	709.2697	$[M+H]^+$	1	708.1684	$C_{26}H_{34}O_7$	709.1763		11.15	1.77	Intermediate - low (72h)					123.0446(88); 573.8192(20); 439.2071(33); 151.0503 (38); 133.0888(21)	
b-D-Xylopyranosyl-a-L-rhamnopyranosyl-D-fucose	487.9098	$[M+2Na]^+$	1	442.1681	$C_{17}H_{30}O_{13}$	488.1476		6.54	2.37	High (96h)					147.0792(20); 165.0425(10); 423.2266(20);294.1125 (22); 311.1538(16); 461.5550 (100)	
Hydroxyisoleucine	130.0873	$[M+H-H_2O]^+$	1	147.0906	$C_6H_{13}NO_3$	130.0862	-8.5	0.71	2.18	Intermediate (120h)					84.0823(35); 130.0873(100)	
Adenine	136.0632	$[M+H]^+$	1	135.0539	$C_5H_5N_5$	136.0617		2.46	2.71	Low (144h)					119.0375(100); 136.0623(67); 137.0600(1)	
Ala-Ala-OH	307.0319	$[M+H]^+$	1	268.069	$C_{11}H_{12}N_2O_6$	307.0327	2.6	4.59	2.70	Low (144h)					161.0746(100); 143.0606(70); 115.0868(81)	
Kynurenine	247.1647	$[M+K]^+$	1	208.0879	$C_{10}H_{12}N_2O_3$	247.1791		3.65	2.58	Low (144h)					120.08(100); 148.10(100); 74.02(20); 192.07(15);	

*Color intensity goes from green (minimum) to red (maximum).

(Marseglia et al., 2014; Mayorga-Gross et al., 2016). Moreover, the peptides FGVPSK and IFVPHYN were identified, which have been considered relevant peptides as precursors of a strong cocoa-specific aroma after *in vitro* reactions between proteolysis products derived from the interaction of proteases and vicilin-class (7S) globulin and reducing sugars (Voigt et al., 2018). Additionally, we found that as a putative biomarker of fermentation time, the tripeptide SPG from the 47 kDa vicilin fragment appeared at higher intensities at the late times of fermentation (120–144 h) and may have a prominent role in flavor formation. We interpret that this peptide might come from the precursor PVNSPGKY, which was reported by John, (2019), who observed its higher intensity at later stages of fermentation and showed its greatest decrease in intensity upon roasting. Therefore, they concluded that its loss during roasting suggests that it participates directly in the Maillard reaction to generate flavor compounds or be further broken down into smaller reactive peptide fragments, such as SPG.

In this study, we found that after the fermentation process, peptides from vicilin were more abundant than peptides from albumin at a ratio of 5 to 1 (15 peptides from vicilin and 3 peptides from albumin). These results are in accordance with those of Scollo et al. (2020), who demonstrated that compared with albumin, vicilin is more extensively degraded during fermentation, and its peptides are more abundant. Moreover, it is worth highlighting that the amino acid region of the vicilin peptides found in this study were between residues 171 and 538. Similarly, Kratzer et al., (2009) found that the amino acid region of vicilin in which cocoa aroma precursors are derived is between residues 131 and 566.

Other metabolites were also tentatively identified (Table 2). Two hydroxy amino acids, hydroxy-leucine and hydroxy-isoleucine, are secondary products derived from the oxidation pathway of leucine and isoleucine. Both amino acids suffer deamination by amino acid dehydrogenases. The free hydrophobic amino acids leucine and isoleucine are products released by the Ehrlich pathway (Smit et al., 2009). For isoleucine, the derived volatiles are 2-methylbutanal, 2-methylbutanol and 2-methylbutanoic acid. In this work, these aldehydes and the aforementioned alcohols were detected in the 72 h and 144 h chocolate samples.

Another discriminating metabolite was the A-type procyanidin dimer arabinoside, which is a glycosylated flavonoid that was also reported in chocolate samples by various authors (Cambrai et al., 2017; Fayeulle et al., 2019). Notably, Fayeulle et al., (2019) found that the A-type procyanidin dimer O-pentoside is one of the 50 most relevant flavonoids for discriminating chocolate samples into four sensory poles/groups with different qualities. Cambrai et al. (2017) also found that the presence of this glycosylated derivative of flavan-3-ols and other flavonoids in the chocolate matrix allows their discrimination by geographical and varietal origin. In general, when establishing a link between the flavonoids present in the beans and their sensory qualities, flavonoids are mainly involved in the perceptions of bitterness and stringency. O-glycosides have been reported to cause velvety, silky, astringent and mouth-coating sensations (Fayeulle et al., 2019). The above reports may explain the high bitterness of the chocolates after 144 h of fermentation considering that this dimer exhibited a higher intensity at this fermentation time. The same result was observed at 120 h. The relatively higher intensity of this flavonoid in the beans after 120 h is also related to the high scores of astringency and bitterness attributes perceived by the judges. In contrast, the cocoa beans at 96 h showed the lowest relative intensity of this compound and exhibited the lowest astringency and bitterness scores.

The oligosaccharide b-D-xylopyranosyl-a-L-rhamnopyranosyl-D-fucose exhibited higher relative intensities after 96 and 120 h of fermentation and decreased after 144 h. This suggests that the presence of this carbohydrate could be a product of the enzymatic degradation of a higher molecular weight polysaccharide. Actually, Megias-Perez et al., (2020) reported the absence of disaccharides and a lower content of oligosaccharides during the last fermentation stage. However, the

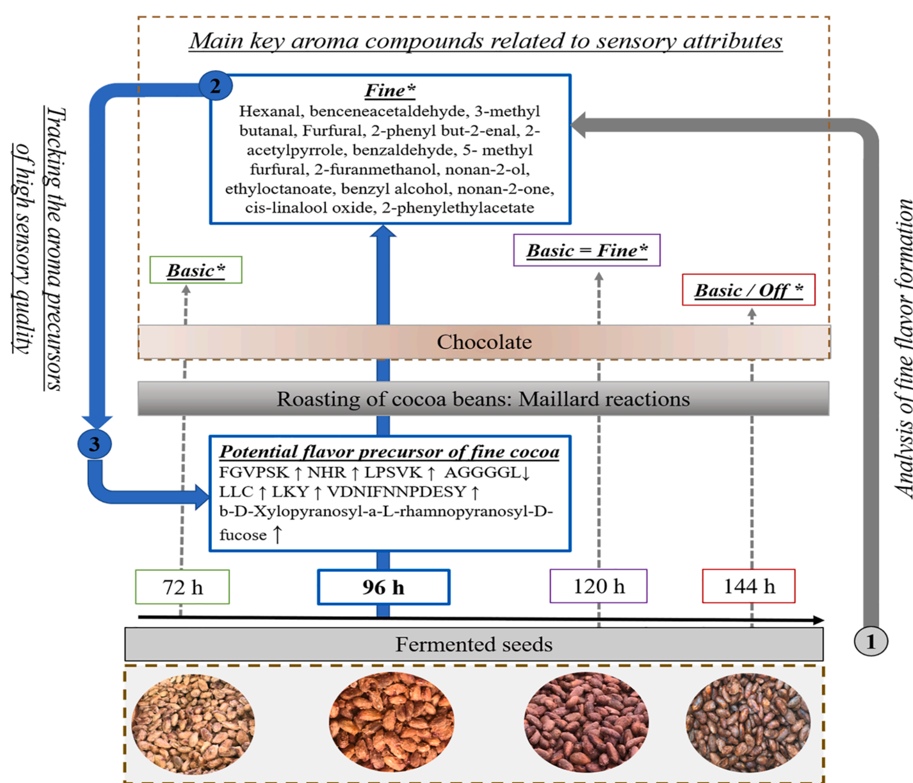


Fig. 4. Scheme of tracking of potential process biomarkers of fine quality. Numbers show the main steps followed to define biomarkers. The arrows in the potential flavor precursors indicate whether their relative intensity increases or decreases at 96 h. *Table 1 presents in more detail the volatiles related to the sensory attributes.

degradation mechanism of low molecular weight carbohydrates during the fermentation of cocoa is still unknown.

Adenine was also found to be a discriminant metabolite presenting a higher relative intensity after 72 h of fermentation. This metabolite belongs to the purine chemical class and has been reported as the most efficient precursor of xanthines (theobromine and caffeine) through the xanthosine pathway (Ashihara et al., 2017). Similar to kynurenine, which was reported as a tryptophan derivative present in cacao (Yilmaz & Gökmen, 2018), although these compounds were highlighted during our fermentation process, there are still no reports of their influence on chocolate flavor. Only Yilmaz & Gökmen (2018) found that kynurenic acid, a compound derived from the kynurenine transformation pathway, was present in cocoa powder.

4. Conclusions

Sensory analysis is the cornerstone that supports the differences between fine flavor and bulk cocoa production. In this sense, we proposed, for the first time, a series of statistical analyses that allowed us to validate the panel reliability and the chocolate flavor profile of the different treatments. Moreover, it was possible to determine the most significant sensory features of the chocolate profile that allow discrimination between cocoa bean fermentation times. In addition, statistical analysis and validation tests enabled us to compare the proposed treatments and define which chocolates presented superior characteristics regarding their fine flavor attributes.

We would like to highlight that this research is pioneering, showing that the flavor profile of chocolate varies according to the fermentation time of cocoa beans between 72 h (3 days) and up to 144 h (6 days). An integral approach considering highly reliable sensory evaluation, the volatile profiles of the chocolate samples and metabolomic data allowed the ideal time of fermentation to be defined to obtain cocoa with fine flavor attributes such as fruitiness, spices, nuttiness and florality. We found that for the three different regions studied, a fermentation time of 96 h (4 days) was optimum. At 120 h (5 days), the quality was slightly reduced, and at 144 h (6 days), off flavors appeared. On the other hand, chocolate produced with cocoa beans fermented for only 72 h (3 days) did not have fine aroma quality attributes but only the quality of a bulk cocoa product.

Considering that the optimum quality is reached at 96 h, we can propose some potential precursor compounds/process biomarkers that also peaked at 96 h. These process biomarkers could potentially be associated with fine attributes and related to key aromatic compounds (VOCs). Fig. 4 summarizes how this approach may allow tracking to fine flavor precursor compounds. However, in order to confirm these assumptions, it will be necessary to test these pure compounds *in vitro* by inducing the Maillard reaction and analyzing the aroma compounds released. This will be a fundamental step to elucidate and understand cocoa flavor formation and help to optimize postharvest operations to improve quality.

Funding statement

This was supported by Bancoldex, the Andes University, Buenaventura Chamber of Commerce, and the Colombian Ministry of Agriculture and Rural Development (MADR). This work was also supported by the Corporación Colombiana de Investigación Agropecuaria AGROSAVIA, through grant No. 1001134 as part of the internal projects of the research agenda entitled 'Cacao in postharvest processes'.

CRediT authorship contribution statement

Sebastián Escobar: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Margareth Santander:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Martha Zuluaga:** Data curation, Formal analysis, Investigation, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Iván Chacón:** Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Jader Rodríguez:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Fabrice Vaillant:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge the generous assistance provided by cocoa farmers associations of Buenaventura: Asociación de Productores Campesinos de Cisneros, Asociación de Agricultores y Cacaoteros de Bajo Calima, and Asociación de Productores y Transformadores de Zacarías. We also thank the Andes University, especially Professor Bart Van Hoof, Maite Rosales, Juanita Duque, and Jesús Riascos researchers, Bancoldex and the Buenaventura Chamber of Commerce for contributing resources for the development of this research project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2021.130627.

References

Abbott, P. C., Benjamin, T. J., Burniske, G. R., Croft, M. M., Fenton, M. C., Kelly, C., Lundy, M., Rodríguez-Camayo, F., & Wilcox, M. (2017). Un análisis de la cadena productiva del cacao en Colombia [An analysis of the cocoa production chain in Colombia (p. 86). USAID Colombia.

Ashihara, H., Mizuno, K., Yokota, T., & Crozier, A. (2017). Xanthine Alkaloids: Occurrence, Biosynthesis, and Function in Plants. In A. D. Kinghorn, H. Falk, S. Gibbons, & J. Kobayashi (Eds.), Progress in the Chemistry of Organic Natural Products (pp. 105:1-88). Springer. 10.1007/978-3-319-49712-9_1.

Assi-Clair, B. J., Koné, M. K., Kouamé, K., Lahon, M. C., Berthiot, L., Durand, N., … Guéhi, T. S. (2019). Effect of aroma potential of Saccharomyces cerevisiae fermentation on the volatile profile of raw cocoa and sensory attributes of chocolate produced thereof. *European Food Research and Technology, 245*(7), 1459–1471. https://doi.org/10.1007/s00217-018-3181-6.

Barišić, V., Kopjar, M., Jozinović, A., Flanjak, I., Ačkar, Đ., Miličević, B., … Babić, J. (2019). The chemistry behind chocolate production. *Molecules, 24*(17), 3163. https://doi.org/10.3390/molecules24173163.

Cambrai, A., Marchioni, E., Julien-David, D., & Marcic, C. (2017). Discrimination of Cocoa Bean Origin by Chocolate Polyphenol Chromatographic Analysis and Chemometrics. *Food Analytical Methods, 10*(6), 1991–2000. https://doi.org/10.1007/s12161-016-0744-7.

D'Souza, R. N., Grimbs, A., Grimbs, S., Behrends, B., Corno, M., Ullrich, M. S., & Kuhnert, N. (2018). Degradation of cocoa proteins into oligopeptides during spontaneous fermentation of cocoa beans. *Food Research International, 109*, 506–516. https://doi.org/10.1016/J.FOODRES.2018.04.068.

Deuscher, Z., Andriot, I., Sémon, E., Repoux, M., Preys, S., Roger, J. M., … Le Quéré, J. L. (2019). Volatile compounds profiling by using proton transfer reaction-time of flight-mass spectrometry (PTR-ToF-MS). The case study of dark chocolates organoleptic differences. *Journal of Mass Spectrometry, 54*(1), 92–119. https://doi.org/10.1002/jms.4317.

Deuscher, Z., Gourrat, K., Repoux, M., Boulanger, R., Labouré, H., & Quéré, L.e. (2020). Key Aroma Compounds of Dark Chocolates Differing in Organoleptic Properties : A GC-O Comparative Study. *Molecules, 1–33*. https://doi.org/10.3390/molecules25081809.

Efenberger-Szmechtyk, M., Nowak, A., & Kregiel, D. (2018). Implementation of chemometrics in quality evaluation of food and beverages. *Critical Reviews in Food Science and Nutrition, 58*(10), 1747–1766. https://doi.org/10.1080/10408398.2016.1276883.

Fayeulle, N., Meudec, E., Boulet, J. C., Vallverdu-Queralt, A., Hue, C., Boulanger, R., … Sommerer, N. (2019). Fast Discrimination of Chocolate Quality Based on Average-Mass-Spectra Fingerprints of Cocoa Polyphenols. *Journal of Agricultural and Food Chemistry, 67*(9), 2723–2731. https://doi.org/10.1021/acs.jafc.8b06456.

Frauendorfer, F., & Schieberle, P. (2019). Key aroma compounds in fermented Forastero cocoa beans and changes induced by roasting. *European Food Research and Technology, 245*(9), 1907–1915. https://doi.org/10.1007/s00217-019-03292-2.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 1–24. https://doi.org/10.1108/EBR-11-2018-0203.

Hinneh, M., Abotsi, E. E., Van de Walle, D., Tzompa-Sosa, D. A., De Winne, A., Simonis, J., … Dewettinck, K. (2020). Pod storage with roasting: A tool to diversifying the flavor profiles of dark chocolates produced from 'bulk' cocoa beans? (Part II: Quality and sensory profiling of chocolates). *Food Research International, 132*, 109116. https://doi.org/10.1016/j.foodres.2020.109116.

Hinneh, M., Semanhyia, E., Van de Walle, D., De Winne, A., Tzompa-Sosa, D. A., Scalone, G. L. L., … Dewettinck, K. (2018). Assessing the in fluence of pod storage on sugar and free amino acid profiles and the implications on some Maillard reaction related flavor volatiles in Forastero cocoa beans. *Food Research International, 111*, 607–620. https://doi.org/10.1016/j.foodres.2018.05.064.

International Organization for Standardization. (2016). *ISO 13299:2016* (p. 41). Sensory analysis Methodology: General guidance for establishing a sensory profile.

John, W. (2019). Exploring Metabolomic Flux and Achieving Prediction Capability in Cocoa Bean Fermentation using Model Systems (p. 124). https://d-nb.info/1188886983/34.

Kratzer, U., Frank, R., Kalbacher, H., Biehl, B., Wöstemeyer, J., & Voigt, J. (2009). Subunit structure of the vicilin-like globular storage protein of cocoa seeds and the origin of cocoa- and chocolate-specific aroma precursors. *Food Chemistry, 113*(4), 903–913. https://doi.org/10.1016/j.foodchem.2008.08.017.

Kumari, N., Grimbs, A., D'Souza, R. N., Verma, S. K., Corno, M., Kuhnert, N., & Ullrich, M. S. (2018). Origin and varietal based proteomic and peptidomic fingerprinting of Theobroma cacao in non-fermented and fermented cocoa beans. *Food Research International, 111*, 137–147.

Magagna, F., Guglielmetti, A., Liberto, E., Reichenbach, S. E., Allegrucci, E., Gobino, G., … Cordero, C. (2017). Comprehensive Chemical Fingerprinting of High-Quality Cocoa at Early Stages of Processing: Effectiveness of Combined Untargeted and Targeted Approaches for Classification and Discrimination. *Journal of Agricultural and Food Chemistry, 65*(30), 6329–6341. https://doi.org/10.1021/acs.jafc.7b0216710.1021/acs.jafc.7b02167.s001.

Marseglia, A., Musci, M., Rinaldi, M., Palla, G., & Caligiani, A. (2020). Volatile fingerprint of unroasted and roasted cocoa beans (Theobroma cacao L.) from different geographical origins. Food Research International, 132(October 2019), 109101. 10.1016/j.foodres.2020.109101.

Marseglia, A., Sforza, S., Faccini, A., Bencivenni, M., Palla, G., & Caligiani, A. (2014). Extraction, identification and semi-quantification of oligopeptides in cocoa beans. *Food Research International, 63*, 382–389. https://doi.org/10.1016/j.foodres.2014.03.046.

Mayorga-Gross, A. L., Quirós-Guerrero, L. M., Fourny, G., & Vaillant, F. (2016). An untargeted metabolomic assessment of cocoa beans during fermentation. *Food Research International, 89*, 901–909. https://doi.org/10.1016/j.foodres.2016.04.017.

Megias-Perez, R., Moreno-Zambrano, M., Behrends, B., Corno, M., & Kuhnert, N. (2020). Monitoring the changes in low molecular weight carbohydrates in cocoa beans during spontaneous fermentation: A chemometric and kinetic approach. *Food Research International, 128*, 108865. https://doi.org/10.1016/j.foodres.2019.108865.

Menezes, A. G. T., Batista, N. N., Ramos, C. L., Silva, A. R.d. A.e., 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. https://doi.org/10.1016/j.foodres.2015.12.036.

Nielsen, D. S., Hønholt, S., Tano-Debrah, K., & Jespersen, L. (2005). Yeast populations associated with Ghanaian cocoa fermentations analysed using denaturing gradient gel electrophoresis (DGGE). *Yeast, 22*(4), 271–284. https://doi.org/10.1002/yea.v22:410.1002/yea.1207.

Ríos, F., Rehpani, C., Ruiz, A., & Lecaro, J. (2017). Estrategias país para la oferta de cacaos especiales - Políticas e iniciativas privadas exitosas en el Perú, Ecuador, Colombia y República Dominicana (p. 136). Swisscontact Colombia Foundation.

Rodriguez-Campos, J., Escalona-Buendía, H. B., Contreras-Ramos, S. M., Orozco-Avila, I., Jaramillo-Flores, E., & Lugo-Cervantes, E. (2012). Effect of fermentation time and drying temperature on volatile compounds in cocoa. *Food Chemistry, 132*(1), 277–288. https://doi.org/10.1016/j.foodchem.2011.10.078.

Rottiers, H., Tzompa Sosa, D. A., De Winne, A., Ruales, J., De Clippeleer, J., De Leersnyder, I., … Dewettinck, K. (2019). Dynamics of volatile compounds and flavor precursors during spontaneous fermentation of fine flavor Trinitario cocoa beans. *European Food Research and Technology, 245*(9), 1917–1937. https://doi.org/10.1007/s00217-019-03307-y.

Salger, M., Stark, T. D., & Hofmann, T. (2019). Taste Modulating Peptides from Overfermented Cocoa Beans [Research-article]. *Journal of Agricultural and Food Chemistry, 67*(15), 4311–4320. https://doi.org/10.1021/acs.jafc.9b0090510.1021/acs.jafc.9b00905.s001.

Santander, M., Rodríguez, J., Vaillant, F. E., & Escobar, S. (2019). An overview of the physical and biochemical transformation of cocoa seeds to beans and to chocolate: Flavor formation. *Critical Reviews in Food Science and Nutrition*, 1–21. https://doi.org/10.1080/10408398.2019.1581726.

Santander, M., Vaillant, F., Sinuco, D., Rodríguez, J., & Escobar, S. (2021). Enhancement of fine flavour cocoa attributes under a controlled postharvest process. *Food Research International, 143*, 110236. https://doi.org/10.1016/j.foodres.2021.110236.

Scollo, E., Neville, D. C. A., Oruna-Concha, M. J., Trotin, M., Umaharan, P., Sukha, D., … Cramer, R. (2020). Proteomic and peptidomic UHPLC-ESI MS/MS analysis of cocoa beans fermented using the Styrofoam-box method. *Food Chemistry, 316*, 126350. https://doi.org/10.1016/j.foodchem.2020.126350.

Smit, B. A., Engels, W. J. M., & Smit, G. (2009). Branched chain aldehydes: Production and breakdown pathways and relevance for flavour in foods. *Applied Microbiology and Biotechnology, 81*(6), 987–999. https://doi.org/10.1007/s00253-008-1758-x.

Thonusin, C., IglayReger, H. B., Soni, T., Rothberg, A. E., Burant, C. F., & Evans, C. R. (2017). Evaluation of intensity drift correction strategies using MetaboDrift, a normalization tool for multi-batch metabolomics data. *Journal of Chromatography A, 1523*, 265–274. https://doi.org/10.1016/j.chroma.2017.09.023.

Tuenter, E., Delbaere, C., De Winne, A., Bijttebier, S., Custers, D., Foubert, K., … Pieters, L. (2020). Non-volatile and volatile composition of West African bulk and Ecuadorian fine-flavor cocoa liquor and chocolate. *Food Research International, 130*, 108943. https://doi.org/10.1016/j.foodres.2019.108943.

Utrilla-Vázquez, M., Rodríguez-Campos, J., Avendaño-Arazate, C. H., Gschaedler, A., & Lugo-Cervantes, E. (2020). Analysis of volatile compounds of five varieties of Maya cocoa during fermentation and drying processes by Venn diagram and PCA. Food Research International, 129(November 2019), 108834. 10.1016/j.foodres.2019.108834.

van Gemert, L. J. (2011). Odour Thresholds: Compilations of odour threshold values in air, water and other media (Second ed.). Oliemans, Punter & Partners BV.

Voigt, J., Textoris-Taube, K., & Wöstemeyer, J. (2018). pH-Dependency of the proteolytic formation of cocoa- and nutty-specific aroma precursors. *Food Chemistry, 255*, 209–215. https://doi.org/10.1016/j.foodchem.2018.02.045.

Wang, P., Ma, X., Wang, W., Xu, D., Zhang, X., Zhang, J., & Sun, Y. (2019). Characterization of flavor fingerprinting of red sufu during fermentation and the comparison of volatiles of typical products. *Food Science and Human Wellness, 8*(4), 375–384. https://doi.org/10.1016/j.fshw.2019.11.004.

Yilmaz, C., & Gökmen, V. (2018). Determination of tryptophan derivatives in kynurenine pathway in fermented foods using liquid chromatography tandem mass spectrometry. *Food Chemistry, 243*, 420–427. https://doi.org/10.1016/j.foodchem.2017.10.004.