



Proceedings of the 16th Weurman Flavour Research Symposium

Edited by

Elisabeth Guichard

Jean-Luc Le Quéré

National Research Institute for Agriculture, Food and Environment (INRAE)

Centre for Taste, Smell and Feeding Behaviour

Centre des Sciences du Goût et de l'Alimentation (CSGA)

Dijon, France



Genetic bases of fruity notes (fresh and dried) of the Nacional cocoa variety

KELLY COLONGES^{1,2,3}, Juan-Carlos Jimenez⁴, Alejandra Saltos⁴, Edward Seguin⁵,
Rey Gastón Loor Solorzano⁴, Olivier Fouet¹, Xavier Argout¹, Sophie Assemet^{2,3}, Fabrice
Davrieux^{2,3}, Eduardo Morillo⁴, Renaud Boulanger^{2,3}, Emile Cros^{2,3}, Claire Lanaud¹

¹ CIRAD, UMR AGAP, F-34398 Montpellier, France. AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France. Kelly.colonges@cirad.fr

² CIRAD, UMR QUALISUD, Montpellier, France. QualiSud, Univ Montpellier, CIRAD, Montpellier SupAgro, Univ Avignon, Univ Réunion, Montpellier, France.

³ Qualisud, Univ Montpellier, Avignon Université, Cirad, Institut Agro, IRD, Université de La Réunion, Montpellier, France.

⁴ Instituto Nacional de Investigación Agropecuarias, INIAP, Ecuador

⁵ Guittard, United-States

Abstract

Theobroma cacao is the only source of cocoa. Cocoa is classified into two types of products: bulk cocoa and fine flavour cocoa. Contrary to bulk cocoa, fine aromatic cocoa is characterized by its floral and fruity aromas [1]. In order to understand the genetic determinism of the formation of these aromas in cocoa beans, a genetic study using the Genetic Wide Association Study (GWAS) method was undertaken. It was carried out on 158 clones belonging to a population of Nacional tree type cultivated in Ecuador, whose volatile compound concentrations and sensory profiles were characterized for its diversity. This study revealed areas of correlation between, on the one hand, the genetic diversity of this population, revealed by molecular marker alleles and the volatile compounds detected in the different clones, and on the other hand, between this same genetic diversity and their sensory profiles. These correlation zones, also called associations, are therefore linked to one of these traits, but also in some cases to both types of traits. Thanks to these associations, which correspond to a restricted area of the cocoa genome, and the knowledge of its complete sequence [2], candidate genes have been brought to light. Some of them are known and identified in biosynthesis pathways of volatile compounds, which are themselves known to have a fruity note. In a preliminary study, a difference in the expression of these genes was identified between four genotypes (two floral and two fruity genotypes) during different stages of development and fermentation of the beans. The results showed that the candidate genes tended to be activated during fermentation and not during the maturation stages of the pods.

Keywords: cocoa, fruity aroma, genetics

Introduction

There are two types of cocoa: "Standard or bulk" cocoa, which has a strong cocoa taste, and aromatic fine cocoa, which is characterised by floral and fruity notes [1]. The Nacional variety of cocoa is classified as a fine variety, characterised by floral and spicy notes [3] known as the "ARRIBA" flavour. At present, the trees grown as Nacional (called Nacional modern in this study) are the result of several generations of crosses between the ancestral Nacional and Trinitarios (themselves hybrids between the Criollo and Amelonado varieties)[4]. Criollo is also a fine aromatic cocoa variety characterised by fruity notes [5]. While Amelonado is known for its strong cocoa aroma. The floral aroma of Nacional has been studied and two main biosynthetic pathways have been identified as responsible for this aroma: the terpene biosynthetic pathway and the L-phenylalanine degradation pathway [6, 7]. In this study, part of the deciphering of the fruity flavour in cocoa from trees of the Nacional modern variety will be presented. A GWAS (Genome Wide Association Study) was conducted to find out which areas of the genome are responsible for this fruity flavour. The GWAS study was carried out using phenotyping data including the determination of volatile compounds related to the fruity taste as well as sensory analyses. To further investigate the genomic determinants of the fruity aroma of Nacional modern, candidate genes in the biosynthetic pathways of the identified fruity compounds were searched for in the identified association areas.

Experimental

Plant material

The plant material used for these experiments was composed of a collection of 151 cocoa trees from Ecuador conserved in the Pichilingue experimental station of the "Instituto Nacional de Investigaciones Agropecuarias" (INIAP) and the "Colección de Cacao de Aroma Tenguel" (CCAT) of Tenguel. This population represents the Nacional variety currently grown in Ecuador and has been described by Loor *et al* [8].

Biochemical analysis

Cocoa beans samples were all fermented and dried with the same method and in the same place of Pichilingue. A part of cocoa beans samples was also roasted. Volatile compounds analysis was carried out on dried fermented beans and roasted beans. The SPME extraction fibre and GC/MS analysis were conducted according to the conditions described by Assi Clair *et al* [9] with Agilent 6890N gas chromatography-mass spectrometer (GC-MS) equipped with a Hewlett Packard capillary column DBWAX, 60 m length \times 0.25 mm internal diameter \times 0.25 μ m film thickness (Palo Alto, CA, USA).

Sensory analysis

146 individuals were characterized by sensory analyses based on blind tastings carried out on 3 repetitions per sample. The tastings were carried out on cocoa liquor. Thirteen fruity notes were judged with a score ranging from zero (no fruity notes detected) to ten. We used the average of the three replicates for the phenotype used to carry out GWAS.

Genetic analysis

GWAS was performed with SNP and SSR markers associated with biochemical (146 accessions \times 5195 markers) and sensory (144 accessions \times 5195 markers) traits using TASSEL v5. Two models were used: mixed linear model (MLM) and a fixed effect model (GLM). In both cases, a structure matrix was determined by running a PCA. Candidate genes were identified through association zones and their annotated function in the cocoa genome.

Expression analysis

The expression of four genes (two coding for alpha-beta hydrolases, one for carboxylesterase and one for GDSE esterase/lipase) was studied during pod ripening at 18, 20 and 22 weeks of development as well as 24 hours after the start of fermentation. These studies were conducted on four different clones, known for their fine flavour, characterised by a fruity and floral aroma: EET103, EET19, EET575 and EET62.

Results and discussion

GWAS analyses are used to statistically determine which areas of the genome are linked to the traits being tested.

Sensory traits

Thanks to this genetic analysis method, 22 areas of associations related to sensory data were detected. Of the 13 fruit sensory categories, associations were detected for 3 of them: "Fruity-Dark Tree Fruit", "Fruity-Dried fruit" and "Fruity-Berries". The strongest areas of association were detected for the note "Fruity-Dark Tree Fruit" on chromosome 1 (Figure 1).

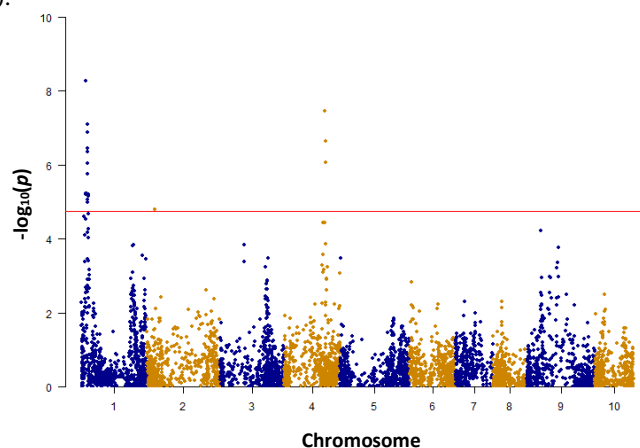


Figure 1: Manhattan plot representing all markers tested along the 10 cocoa chromosomes for associations with the Fruity Dark tree fruit note.

Biochemical traits

The GWAS analysis allowed the detection of 480 areas of associations related to the concentrations of volatile compounds. These association zones were detected in relation to 3 acids, 12 alcohols, 7 aldehydes, 22 esters, 8 ketones, 1 lactone, 11 pyrazines and 5 terpenes. According to these results, five biosynthetic pathways appear to

play a role in the synthesis of fruity aroma in cocoa: the monoterpene biosynthetic pathway, the L-phenylalanine degradation pathway, the simple sugar degradation pathway, the fatty acid degradation pathway and the pyrazine biosynthetic pathway. The study of associations related to the monoterpene biosynthetic pathway and the L-phenylalanine degradation pathway have been extensively studied previously [7].

Areas of common associations between sensory and biochemical traits were first sought. These areas have a higher probability to explain the fruity aroma of cocoa. Six co-locations between associations related to pyrazines and those related to sensory traits were identified. Four of them are present on chromosome 1, one on chromosome 4 and one on chromosome 10. Eight areas of co-locations between associations related to sensory traits and compounds involved in the degradation pathways of fatty acids and simple sugars were observed: four co-locations are on chromosome 1, two on chromosome 2, one on chromosome 4 and one on chromosome 10.

Candidate genes

In all the areas of association detected, candidate genes were sought. A first search was carried out by looking for genes coding for proteins with essential enzymatic functions for the different biosynthesis pathways identified.

In cocoa, it is well known that pyrazines are synthesized through the Maillard reaction [10]. It is a set of chemical reactions that take place during the process. In cocoa processing, the Maillard reaction occurs mainly during roasting, but it can also occur during drying. The Maillard reaction allows the synthesis of pyrazine by combining free amino acids and reducing sugars. In this study, we therefore looked for genes coding for enzymes involved in protein degradation or in the synthesis of reducing sugars.

In the 277 association areas related to pyrazines, 213 candidate genes involved in the production of precursors of the Maillard reaction were identified. These genes have mainly peptidase or protease functions.

Preliminary expression studies of some of these genes were conducted. The expression of two genes coding for alpha-beta hydrolases was studied. The two genotypes EET103 and EET19 express alpha-beta hydrolase 1 equally or more during fermentation. The EET575 genotype expresses it more at 18 weeks of maturity and the EET62 genotype expresses it more at 18 and 20 weeks of maturity. Alpha-beta hydrolase 2 is most expressed at 20 weeks of pod development for genotype EET103, at 22 weeks of development for genotypes EET19 and EET575, at 18 and 20 weeks of development for genotype EET62 (Figure 2). Some genotypes appear to express more alpha-beta hydrolase 1 during fermentation than during pod maturity. It is at this time that alpha-beta hydrolases would break down proteins into amino acids through their peptidase functions. Cocoa beans would then be richer in amino acids, which would allow more pyrazine synthesis through the Maillard reaction during drying and roasting. Reineccius (2006) observed a higher concentration of pyrazine in well-fermented beans [11]. The expression of alpha-beta hydrolase 2 seems to be more intense during pod maturation than during the fermentation. Pod ripening also seems to play a role in aroma synthesis.

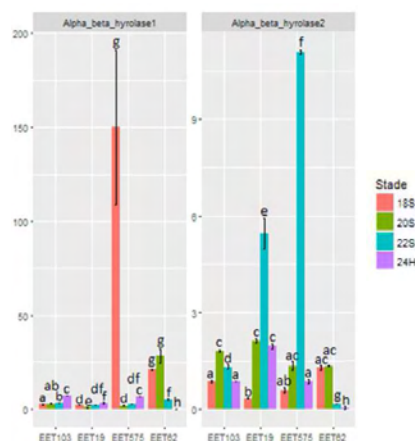


Figure 2: Histogram of the expression profile of the genes coding for Alpha-Beta Hydrolase at different stages of bean maturity and during fermentation. Histograms with the same letter are not significantly different at the 1% threshold.

Acids, alcohols, ketones and esters are mainly synthesised as a result of the degradation of fatty acids and/or simple sugars. Candidate genes encoding enzymes involved in these degradation pathways were searched for in all association zones of all volatile compounds.

In the 480 association areas related to compounds involved in simple sugars degradation, 125 candidate genes were identified. These genes have mainly hydrolase or esterase functions. In the association areas related to compounds involved in fatty acid degradation, 217 candidate genes were identified. These genes have mainly lipase or esterase functions. There are common candidate genes between these simple sugar degradation pathway and fatty acid degradation pathway.

Preliminary expression studies of some of these genes were conducted. The first one coding for a carboxylesterase and the second for a GDSL esterase/lipase. The carboxylesterase gene is more expressed during fermentation in genotypes EET19 and EET575, more expressed after 20 weeks of seed development in genotype EET103 and similarly expressed at 18 and 22 weeks of seed development in genotype EET62. The GDSL esterase lipase gene is more expressed during fermentation in all genotypes except EET103. In EET103, it is more expressed at 20 weeks of development (Figure 3).

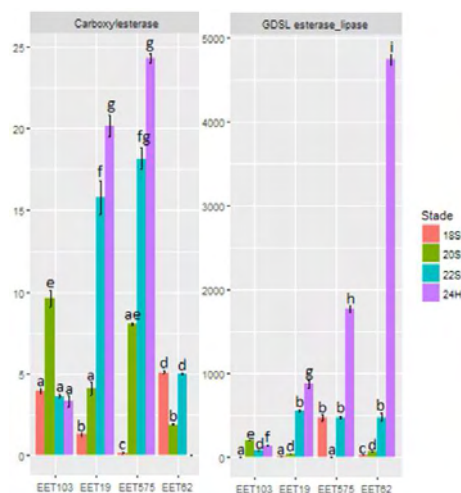


Figure 3: Histogram of the expression profile of the genes coding for Carboxylesterase and GDSL esterase/lipase at different stages of bean maturity and during fermentation. Histograms with the same letter are not significantly different at the 1% threshold.

In most cases, the genes seem to be more strongly expressed during fermentation. During this processing stage, cocoa beans are exposed to fermenting micro-organisms. These exponentially growing micro-organisms are detected by the beans, which then set up a defence system. A large number of volatile compounds have been identified as being used by plants to defend themselves against micro-organism attacks [12–14].

Of the 67 volatile compounds, association zones could not be detected for 13 of them. These compounds belong to different chemical families (two acids, three alcohols, one aldehyde, five esters and two ketones). As most of these different compounds do not seem to be genetically related to the cocoa trees in this population, it is very likely that they were synthesised by the microorganisms present during the fermentation [15–17]. In addition, it is possible that the microorganisms present during fermentation synthesise some intermediate compounds. The biosynthesis of these compounds could be a synergy between the enzymatic actions of the fermenting microorganisms and the cocoa tree.

The Nacional fruity aroma could also include compounds belonging to the L-phenylalanine degradation pathway, such as benzaldehyde, benzyl alcohol and 1-phenylethyl acetate. Associations have been detected for these three compounds and reported in a previous study [7]. These compounds are also precursors of compounds known to have floral tastes. The enzymatic activity allowing the degradation of benzoic acid into benzaldehyde or benzyl alcohol is one of the keys to the increased presence of compounds with fruity tastes (Figure 4).

Conclusion

The study of aroma determinism is complex. Flavours depend on different factors such as genetics, growing environment and processing. In this study, the objective was to investigate the genetic part of the fruity notes of cocoa (fresh and dry). Using GWAS, we were able to begin to determine how the fruit notes were synthesised by the cocoa beans. Two types of fruity notes were observed: fresh fruity notes, which are mainly, composed of esters and terpene compounds. In this case, three metabolic pathways seem to be involved: the degradation of fatty acids, the degradation of simple sugars and the monoterpene biosynthetic pathway. Similarly, dried fruit aromas were observed, mainly composed of pyrazines. Pyrazines were synthesised by the Maillard reaction during the drying or roasting process. The concentration of pyrazine depends on the synthesis of the Maillard precursors: amino acids and reducing sugars.

A large number of genes potentially involved in the synthesis of fruity aromas have been identified. Preliminary study of the expression of candidate genes shows that the synthesis of enzymes responsible for the production of certain volatile compounds takes place during bean development but also during fermentation, before the death of

the embryo. The hypothesis is that cocoa beans detect fermenting microorganisms and trigger defence mechanisms.

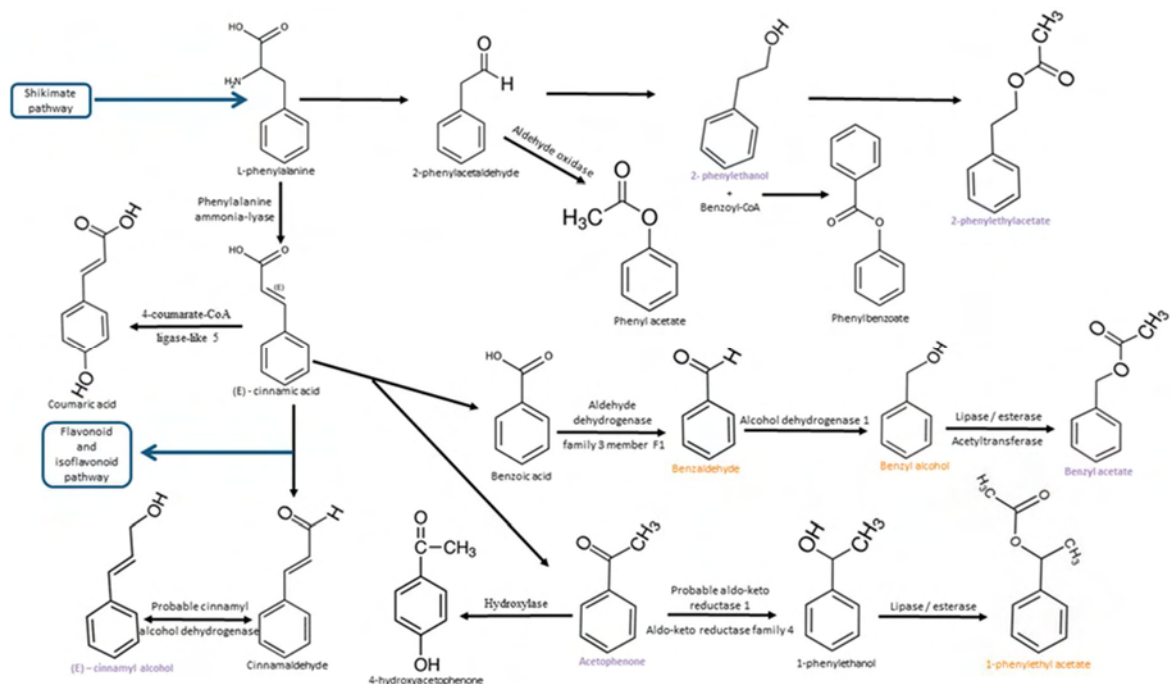


Figure 4: Degradation pathway of L-phenylalanine identified in cocoa. Compounds known to have a floral taste are noted in purple and compounds known for have a fruity taste are noted in orange.

The exact role of these candidate genes in this synthesis is still difficult to determine. A study of the enzymatic activity of the different enzymes encoded by these candidate genes could complement this study. This could also help to identify more precisely the roles of each enzyme.

A complementary analysis, made by Gas-Chromatography-Olfactometry (GCO), comparing fruity, floral and standard genotypes of cocoa, would allow identifying key compounds of the fruity aroma of cocoa.

References

1. Sukha DA, Butler DR, Umaharan P, Boulton E. The use of an optimised organoleptic assessment protocol to describe and quantify different flavour attributes of cocoa liquors made from Ghana and Trinitario beans. *Eur Food Res Technol*. 2008;226(3):405-13.
2. Argout X, Martin G, Droc G, Fouet O, Labadie K, Rivals E, et al. The cacao Criollo genome v2.0: an improved version of the genome for genetic and functional genomic studies. *BMC Genomics*. 2017;18:730.
3. Luna F, Crouzillat D, Cirou L, Bucheli P. Chemical composition and flavor of Ecuadorian cocoa liquor. *J Agric Food Chem*. 2002;50(12):3527-32.
4. Llor RG, Risterucci AM, Courtois B, Fouet O, Jeanneau M, Rosenquist E, et al. Tracing the native ancestors of the modern *Theobroma cacao* L. population in Ecuador. *Tree Genet Genomes*. 2009;5(3):421-33.
5. Lachenaud P, Motamayor JC. The Criollo cacao tree (*Theobroma cacao* L.): a review. *Genet Resour Crop Evol*. 2017;64(8):1807-20.
6. Ziegler G. Linalool contents as characteristic of some flavor grade cocoas. *Z Lebensm-Unters Forsch*. 1990;191(4-5):306-9.
7. Colonges K, Jimenez JC, Saltos A, Seguíne E, Solorzano RL, Fouet O, et al. Two main biosynthesis pathways involved in the synthesis of the floral aroma of the Nacional cocoa variety. *Front Plant Sci*. 2021; Submitted.
8. Llor S. RG. Contribution à l'étude de la domestication de la variété de cacaoyer Nacional d'Equateur : recherche de la variété native et de ses ancêtres sauvages. INIAP Archivo Historico; 2007. 406 p.
9. Assi-Clair BJ, Koné MK, Kouamé K, Lahon MC, Berthiot L, Durand N, et al. Effect of aroma potential of *Saccharomyces cerevisiae* fermentation on the volatile profile of raw cocoa and sensory attributes of chocolate produced thereof. *Eur Food Res Technol*. 2019;245(7):1459-71.
10. Afoakwa EO, Paterson A, Fowler M, Ryan A. Flavor Formation and Character in Cocoa and Chocolate: A Critical Review. *Crit Rev Food Sci Nutr*. 2008;48(9):840-57.

11. Reineccius GA. Flavor Chemistry and Technology [Internet]. CRC Press; 2006 [cited 19 March 2021]. Available at: <https://www.taylorfrancis.com/books/flavor-chemistry-technology-gary-reineccius/10.1201/9780203485347>.
12. Pichersky E, Lewinsohn E, Croteau R. Purification and Characterization of S-Linalool Synthase, an Enzyme Involved in the Production of Floral Scent in *Clarkia breweri*. *Arch Biochem Biophys*. 1995;316(2):803-7.
13. Parent GJ, Giguère I, Mageroy M, Bohlmann J, MacKay JJ. Evolution of the biosynthesis of two hydroxyacetophenones in plants. *Plant Cell Environ*. 2018;41(3):620-9.
14. Palmqvist E, Almeida JS, Hahn-Hägerdal B. Influence of furfural on anaerobic glycolytic kinetics of *Saccharomyces cerevisiae* in batch culture. *Biotechnol Bioeng*. 1999;62(4):447-54.
15. Soles RM, Ough CS, Kunkee RE. Ester Concentration Differences in Wine Fermented by Various Species and Strains of Yeasts. *Am J Enol Vitic*. 1982;33(2):94-8.
16. Abbas CA. Production of Antioxidants, Aromas, Colours, Flavours, and Vitamins by Yeasts. In: *Yeasts in Food and Beverages* [Internet]. Springer, Berlin, Heidelberg; 2006. p. 285-334. Available at: https://link.springer.com/chapter/10.1007/978-3-540-28398-0_10.
17. Ho VTT, Zhao J, Fleet G. Yeasts are essential for cocoa bean fermentation. *Int J Food Microbiol*. 2014;174:72-87.