

Understanding the cocoa genetic resources in the Pacific to assist producers to supply the growing craft market

Natalie L. Dillon, Dapeng Zhang, Lars Nauheimer, Elison Toramo, Paitia Nagalevu, Marie-Vianney Melteras, Sandrine Wallez, Kuinimeri Finau, Sefanaia Nakidakida, Petelo Lepou & Yan Diczbalis

To cite this article: Natalie L. Dillon, Dapeng Zhang, Lars Nauheimer, Elison Toramo, Paitia Nagalevu, Marie-Vianney Melteras, Sandrine Wallez, Kuinimeri Finau, Sefanaia Nakidakida, Petelo Lepou & Yan Diczbalis (2024) Understanding the cocoa genetic resources in the Pacific to assist producers to supply the growing craft market, New Zealand Journal of Crop and Horticultural Science, 52:4, 306-320, DOI: [10.1080/01140671.2023.2278788](https://doi.org/10.1080/01140671.2023.2278788)

To link to this article: <https://doi.org/10.1080/01140671.2023.2278788>



© 2023 The State of Queensland through the Department of Agriculture and Fisheries. Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 18 Dec 2023.



Submit your article to this journal [↗](#)



Article views: 1453



View related articles [↗](#)



View Crossmark data [↗](#)






Citing articles: 2 View citing articles [↗](#)

RESEARCH ARTICLE



Understanding the cocoa genetic resources in the Pacific to assist producers to supply the growing craft market

Natalie L. Dillon ^a, Dapeng Zhang ^b, Lars Nauheimer ^c, Elison Toramo^d,
Paitia Nagalevu^e, Marie-Vianney Melteras^f, Sandrine Wallez^g, Kuinimeri Finau^h,
Sefanaia Nakidakidaⁱ, Petelo Lepou^h and Yan Diczbalis^j

^aCentre for Tropical Agriculture, Horticulture and Forestry Science, Department of Agriculture and Fisheries, Mareeba, Australia; ^bUnited States Department of Agriculture – Agricultural Research Service (USDA-ARS), Sustainable Perennial Crops Laboratory (SPCL), Beltsville Agricultural Research Centre (BARC), Beltsville, MD, USA; ^cAustralian Tropical Herbarium, James Cook University, Cairns, Australia; ^dMinistry of Agriculture and Livestock, Honiara, Solomon Islands; ^eLand Resource Division, Pacific Community, Suva, Fiji; ^fVanuatu Agriculture Research and Technical Centre (VARTC), Luganville, Vanuatu; ^gAlternative Communities Trade in Vanuatu (ACTIV), Stella Mare Sub-Division, Second Lagoon, Port Vila, Vanuatu; ^hPlant and Postharvest Technologies, Scientific Research Organisation of Samoa (SROS), Apia, Samoa; ⁱWainigata Research Station, Ministry of Agriculture, Savusavu, Fiji; ^jCentre for Wet Tropics Agriculture, Horticulture and Forestry Science, Department of Agriculture and Fisheries, South Johnstone, Australia

ABSTRACT

The Pacific countries of Fiji, Samoa, Solomon Islands, and Vanuatu account for less than 2% of the world's cacao dry bean production. To capitalise on the rapidly expanding origin craft cocoa market, understanding the genetic attributes of *Theobroma cacao* in the Pacific is essential. A six-year ACIAR-funded project collected 1647 dried leaf samples from research stations and smallholder farms to identify the population ancestry of *T. cacao* in the partner countries. Using SNP marker profiling, the study identified germplasm high in Amelonado, which represents the majority of material throughout the Pacific. However, the study also identified samples high in the remaining nine genetic groups, including Criollo, Nanay, IMC (Iquitos), Guiana, Parinari (Marañón), Nacional, Ucayali (Contamana), LCT EEN (Curaray), and Purus, which were distributed throughout the Pacific. The results, if utilised in local selection trials, could reposition growers in the Pacific countries, allowing them to supply the growing origin craft cocoa market, with genetically unique beans. Cocoa genetics was one aspect of a project incorporating production and postharvest research to refocus the commodity based cocoa industry to a supplier of high-quality beans of unique Pacific cocoa origin to the rapidly expanding craft 'bean to bar' cocoa industry.

ARTICLE HISTORY

Received 6 July 2023
Accepted 18 October 2023

HANDLING EDITOR

Tony Conner

KEYWORDS

Cacao; cocoa; chocolate;
Pacific; SNP markers; genetic
attributes; diversity; craft
origin; production; growers

Introduction

Cocoa has a long history in the Pacific with trees introduced to Pacific islands in the late nineteenth and early twentieth centuries. Bartley (2005) reports introductions into Fiji

CONTACT Natalie L. Dillon  natalie.dillon@daf.qld.gov.au

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The State of Queensland through the Department of Agriculture and Fisheries. Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

from Trinidad in 1883 and cultivation that started in Samoa in 1883 with plants introduced from Ceylon (Sri Lanka). Commercial production in Papua New Guinea (PNG) is reported to have commenced in the early twentieth century. Initial introductions of *T. cacao* were from Trinidad, Ceylon (Sri Lanka) and Java (Indonesia) and are believed to have been Criollo or Trinitario types. The dates of cocoa introduction into the various Pacific countries and north Queensland are less than clear and rely on available copies of government and company reports and the literature quoted. Regardless, the Pacific was an important producer of cocoa for much of the nineteenth and well into the twentieth century before West Africa, Malaysia and Indonesia became the dominant competitors.

Many countries produce cocoa in the region collectively known as the 'Pacific'. Papua New Guinea and Bougainville are the largest producers by volume. This paper reports on the cocoa genetics of the traditionally main producing countries to the east of Bougainville (Solomon Islands, Vanuatu, Fiji and Samoa) where cocoa was introduced and established from the late 1800s by colonial settlers. Today, cocoa is still an important contributor to small-holder income in these countries. The report also includes the fledgling new industry developing in the wet tropical coastal region of north Queensland.

More recently the growth of cocoa production in the eastern Pacific has been constrained by the negative impacts of disease, natural disasters, and the subpar performance of many older plantations, primarily due to management factors. However, the potential for improvement through the adoption of selected genetics and the implementation of appropriate production and post-harvest agricultural practices exists. This combination can lead to increased production, better disease control, and a more effective, durable, and sustained farming approach. This becomes particularly significant in the eastern Pacific region where cacao cultivation is predominantly undertaken by small-scale farmers with limited resources, often situated in remote or environmentally fragile areas.

Utilising improved cocoa varieties alongside suitable agricultural techniques and links to markets, the living standards of these producers can be elevated. This, in turn, contributes to a more stable supply of cocoa for the industry, resulting in a mutually beneficial situation for farming families, the surrounding ecosystem and chocolate producers. This approach not only enhances the economic well-being of the farmers but also promotes the conservation of natural resources and the preservation of biodiversity in the region.

Pacific cocoa is gaining traction amongst bean-to-bar buyers with particular reference to the country of origin. This interest has been stimulated by the fact that Pacific beans from PNG, Bougainville, Solomon Islands, Vanuatu, Fiji, Samoa, and Australia; submitted to the 'International Cacao of Excellence' since the competition accepted entries from SE Asia and Oceania in 2010, have been rated in the Top 50 bean and received International Awards of Excellence. Indeed, Dillon et al. (2020) suggests that the fine flavour, unusual genetic resources, and novel 'single origin' branding makes the Pacific islands and north Australian cocoa competitive in specialty, high-value, low-volume markets.

Since Motamayor et al. (2008) seminal paper on the geographic and genetic population differentiation of cocoa in Central and South America we have a better understanding of cocoa population diversity within the centres of origin of cocoa. Cocoa diversity was originally believed to revolve around three types of cocoa: Criollo and Forastero types with their hybrid Trinitario being the mainstay of cocoa production globally. We now understand that there are a minimum of ten genetic groups, with nine of them originally being grouped within the Forastero category.

Criollo cocoa is renowned for its delicate and nuanced flavours, often characterised by floral and fruity notes, with hints of jasmine, berries, wood, nuts, and caramel (Ascrizzi et al. 2017). The taste is often described as smooth and creamy, with a pleasant balance of sweetness and bitterness. Criollo cocoa is generally regarded as having a superior flavour profile but is also more susceptible to pests and diseases and has lower productivity compared to other cacao varieties. Forastero, a broad term that now describes nine distinct genetic groups, cacao is the most widely cultivated variety and typically exhibits a stronger and more robust flavour profile, with fruity, earthy, and bitter notes (Aprotosoie et al. 2016). Within the Forastero group there are family groupings, such as National and Scavina, that have unique and sought after flavour profiles (Bartley 2005). Generally, the taste of Forastero can be described as having a pleasant balance of sweetness and acidity, with hints of red fruit, apricot, and spices. Forastero cacao has a relatively lower cocoa butter content compared to other varieties, which can affect the texture and mouthfeel of chocolate made from this cacao, as cocoa butter contributes to the smoothness and creaminess. Trinitario cacao is a hybrid of Criollo and Forastero, combining some of the flavour complexity of Criollo with the disease resistance and higher productivity of Forastero. Criollo and many Trinitario varieties have gained a reputation for their superior aroma and flavour, making them highly sought-after as raw materials in the production of fine flavoured chocolate. These varieties hold significant value in the global premium chocolate market, as well as in the development of future cacao cultivars with enhanced quality attributes (CocoaNet 2012).

The unique cocoa aroma and flavour emerge through intricate biochemical and chemical reactions occurring during the postharvest processing of raw cocoa beans. This aroma is influenced by various factors, including the cocoa genotype, chemical composition of the raw seeds, environmental conditions, farming practices, as well as fermentation, drying and storage practices.

The great bulk of cocoa produced in the Pacific (Solomon Islands and east to Samoa) is from seedlings, initially believed to be Criollo types but then with subsequent introductions of Trinitario and Forastero (subtype–Amelonado) material. An improved understanding of cocoa genetics available in each of the Pacific countries will allow future cocoa researchers to select material of interest in terms of genetic diversity, yield, and potential links to organoleptic qualities. In essence, there is an opportunity for each Pacific nation producers to develop a range of clones unique to the country by virtue of their genetic population ancestry, terroir (environmental factors), and post-harvest practices.

Materials and methods

Materials – sample collection

Cacao leaf samples for DNA extraction were collected from research stations and farmer fields between 2014 and 2019, on the Pacific islands of Fiji, Samoa, Solomon Islands, Vanuatu, and in far-north Queensland, Australia. Selection of trees was initially based on productivity and morphological characteristics as reported by Marfu (2012) and then from interaction with farmers as to production characteristics and bean cotyledon colour. Where possible, healthy, young, flushing leaf buds were collected from individual

trees. Where leaf buds were not available, the next youngest leaf material available at the time was collected. Leaf samples were dried on silica gel for long-term preservation. Where possible morphological (pod colour (immature and mature), pod weight, pod shape, bean number, wet bean weight) and agronomic traits (high, medium, or low yielding and peak season) were recorded, particularly when pods (immature and mature) were present.

In addition to local samples, DNA fingerprint profiles for 169 reference samples were provided by collaborators at the Sustainable Perennial Crops Laboratory (SPCL), Beltsville, MD, USA, and were used for comparison with cacao samples from the Pacific Islands and Australia (Table 1). Reference samples included individual genotypes representative of 10 previously defined genetic populations or germplasm groups including Criollo, Trinitario, Amelonado, Upper Amazon Forastero accessions (Motamayor et al. 2008).

DNA extraction

DNA extraction was performed on 10–20 mg samples of leaf tissue using the DNeasy® Plant extraction kit (Qiagen GmbH, Germany). Extractions were carried out according to manufacturer's instructions with one and half times the first two buffers (AP1, P3) and RNase used in the procedure. DNA was quantified on NanoDrop™ 2000 UV-Vis Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). DNA integrity was assessed by 1% (w/v) horizontal gel electrophoresis. Extracted DNA was stored at 4°C and diluted to a final concentration of 10 ng.µL⁻¹ prior to single nucleotide polymorphism (SNP) analysis.

SNP genotyping

SNP loci were discovered through the analysis of expressed sequence tags (ESTs) derived from diverse cacao plant parts, ensuring a comprehensive representation of the cacao transcriptome (Argout et al. 2008, 2011). For the present study, a panel of 96 SNP loci was selected based on the screening results of 1536 SNPs using Illumina's GoldenGate Assay (Michel Boccara, unpublished data), as well as their previous utilisation in documented research (Fang et al. 2014; Lukman et al. 2014; Li et al. 2021). Sequence information of the 96 SNPs is as described by Li et al. (2021). The methodology for SNP genotyping of cacao is as described by Wang et al. (2009).

Statistical data analysis and software

Descriptive statistics measuring informativeness and quality of the 96 SNP loci were calculated using GenAlEx 6.5 (Peakall and Smouse 2006, 2012). The key descriptive statistics calculated were observed heterozygosity (H_O), within-population gene diversity (H_S), overall gene diversity (H_T), gene diversity amongst samples (H_{ST} ; $H_T - H_S$), fixation index (F_{ST} ; D_{ST}/H_T), and the inbreeding coefficient (F_{IS} ; $1 - H_O/H_S$) as an average across all loci for all sampling sites. The observed heterozygosity (H_O), within-population gene diversity (H_S), and inbreeding coefficient (F_{IS}) were also calculated as an average across loci for each country separately.

Table 1. List of 169 cacao reference selections representing 10 genetic clusters of origin.

Accession #	Germplasm ID	Accession #	Germplasm ID	Accession #	Germplasm ID	Accession #	Germplasm ID
EET 103	Nacional hybrids	IMC 31	IMC	NA 702	Nanay	Ucayali 7	Ucayali
EET 183	Nacional hybrids	IMC 33	IMC	NA 756	Nanay	BC32 Belize	Criollo
EET 19	Nacional hybrids	IMC 38	IMC	NA 79	Nanay	PR127	Criollo
EET 399	Nacional hybrids	IMC 47	IMC	NA 916	Nanay	PR129	Criollo
EET 400	Nacional hybrids	IMC 50	IMC	POUND 1/ B	Nanay	PR139	Criollo
EET 48	Nacional hybrids	IMC 55	IMC	POUND 2/ A	Nanay	PR145	Criollo
EET 62	Nacional hybrids	IMC 58	IMC	PA 107	Parinari	PR154	Criollo
EET 95	Nacional hybrids	IMC 6	IMC	PA 12	Parinari	PR155	Criollo
EET 96	Nacional hybrids	IMC 71	IMC	PA 120	Parinari	PR156	Criollo
Las Brisas 14	Ancient Nacional	IMC 77	IMC	PA 121	Parinari	PR46	Criollo
Las Brisas 16	Ancient Nacional	IMC 78	IMC	PA 124	Parinari	PR71	Criollo
Las Brisas 30	Ancient Nacional	IMC 87	IMC	PA 126	Parinari	BCCN 1 Belize	Amelonado
Las Brisas 7	Ancient Nacional	IMC 94	IMC	PA 136	Parinari	ES 01 Belize	Amelonado
Gloria 1	Ancient Nacional	IMC 96	IMC	PA 149	Parinari	ES 02 Belize	Amelonado
Gloria 12	Ancient Nacional	LCT EEN 376	LCT EEN	PA 150	Parinari	ES 03 Belize	Amelonado
Gloria 13	Ancient Nacional	LCT EEN 386	LCT EEN	PA 156	Parinari	SC 02 Belize	Amelonado
Gloria 2	Ancient Nacional	LCT EEN 389	LCT EEN	PA 169	Parinari	UFC 01 Belize	Amelonado
Gloria 22	Ancient Nacional	LCT EEN 390	LCT EEN	PA 191	Parinari	UFC 02 Belize	Amelonado
Gloria 7	Ancient Nacional	LCT EEN 395	LCT EEN	PA 195	Parinari	PR151	Amelonado
Gloria 9	Ancient Nacional	LCT EEN 399	LCT EEN	PA 2	Parinari	PR76	Amelonado
GU 114/P	GU	LCT EEN 401	LCT EEN	PA 289	Parinari	PR8	Amelonado
GU 175/P	GU	LCT EEN 403	LCT EEN	PA 299	Parinari	REDAMEL 1/27	Amelonado
GU 199/A	GU	LCT EEN 406	LCT EEN	PA 30	Parinari	REDAMEL 1/3	Amelonado
GU 216/A	GU	LCT EEN 408	LCT EEN	PA 300	Parinari	DR2	Trinitario
GU 219/P	GU	LCT EEN 409	LCT EEN	PA 46	Parinari	PR10	Trinitario
GU 230/A	GU	LCT EEN 411	LCT EEN	PA 70	Parinari	PR122	Trinitario
GU 241/P	GU	LCT EEN 412	LCT EEN	Ucayali 11	Ucayali	PR122	Trinitario
GU 254/A	GU	LCT EEN 413	LCT EEN	Ucayali 15	Ucayali	PR152	Trinitario
GU 261/P	GU	LCT EEN 414	LCT EEN	Ucayali 19	Ucayali	PR157	Trinitario
GU 265/P	GU	LCT EEN 415	LCT EEN	Ucayali 24	Ucayali	PR28	Trinitario

(Continued)

Table 1. Continued.

Accession #	Germplasm ID	Accession #	Germplasm ID	Accession #	Germplasm ID	Accession #	Germplasm ID
GU 277/G	GU	LCT EEN 427	LCT EEN	Ucayali 28	Ucayali	PR9	Trinitario
GU 300/P	GU	LCT EEN 429	LCT EEN	Ucayali 35	Ucayali	RIM 189	Trinitario
GU 310/P	GU	LCT EEN 432	LCT EEN	Ucayali 36	Ucayali	RIM 189	Trinitario
GU 322/P	GU	LCT EEN 434	LCT EEN	Ucayali 37	Ucayali	ICS 6	Trinitario
KER 5	GU	NA 13	Nanay	Ucayali 38	Ucayali	ICS 40	Trinitario
ELP16 A	GU	NA 32	Nanay	Ucayali 39	Ucayali	ICS 5	Trinitario
ELP7 A	GU	NA 33	Nanay	Ucayali 4	Ucayali	ICS 6	Trinitario
IMC 10	IMC	NA 149	Nanay	Ucayali 41	Ucayali	ICS 83	Trinitario
IMC 103	IMC	NA 286	Nanay	Ucayali 43	Ucayali	ICS 89	Trinitario
IMC 105	IMC	NA 399	Nanay	Ucayali 5	Ucayali	ICS 93	Trinitario
IMC 14	IMC	NA 670	Nanay	Ucayali 6	Ucayali		
IMC 20	IMC	NA 7/10	Nanay	Ucayali 66	Ucayali		
IMC 20	IMC	NA 70	Nanay	Ucayali 69	Ucayali		

Distance-based multivariate analysis was used to assess the relationship among the individual varieties, as well as their relationship with the reference germplasm. Pairwise genetic distances (Peakall and Smouse 2012) were computed using the DISTANCE procedure implemented in GenALEX 6.5. The same programme was then used to perform Principal Coordinates Analysis (PCoA), based on the pairwise distance matrix.

ONCOR was used for genetic mixture analysis and assignment tests among sampled cacao from the Pacific Islands and far-north Queensland, Australia (Kalinowski et al. 2007). This analysis identifies the origin of individual samples in a mixture of individuals from genetically distinct populations.

Results

Sample collection and DNA extraction

Cacao leaf samples were collected from research stations and farmer fields across the Pacific islands of Fiji, Samoa, Solomon Islands, Vanuatu, and in far-north Queensland, Australia. Sample collection was limited to the countries incorporated within the project scope. A total of 1647 genotypes were collected of which 308 were from Fiji, 141 from Samoa, 117 from the Solomon Islands and 658 from Vanuatu, and 423 from far-north Queensland, Australia. High molecular weight genomic DNA was successfully isolated from all samples collected and the allelic composition and genetic structure were assessed.

Data analysis

The key descriptive statistics are shown in Table 2 as an average across all loci and as the mean for each country. For all loci, the observed heterozygosity (H_O) was 0.23, which is inversely proportional with inbreeding. The total gene diversity (H_T) was 0.179 and can be separated into the within-population gene diversity (H_S ; 0.17), representing the

Table 2. Key descriptive statistics calculated as an average for all loci across all samples or within one country.

	Mean across all loci	Australia	Fiji	Samoa	Solomon Islands	Vanuatu
n. ind.	1647	423	308	141	117	658
H _O	0.2300	0.20	0.23	0.26	0.24	0.23
H _S	0.1735	0.15	0.17	0.19	0.17	0.18
H _T	0.1786	NA	NA	NA	NA	NA
D _{ST}	0.0051	NA	NA	NA	NA	NA
F _{ST}	0.0284	NA	NA	NA	NA	NA
F _{IS}	−0.3252	−0.13	−0.16	−0.24	−0.18	−0.17

Notes: Number of individuals (n. ind), observed heterozygosity (H_O), within-population gene diversity (H_S), overall gene diversity (H_T), gene diversity amongst samples (H_{ST}), fixation index (F_{ST}), and the inbreeding coefficient (F_{IS}), not applicable (NA).

proportion of heterozygosity expected under random mating and proportional with genetic variability within the population, and the gene diversity amongst populations (D_{ST}; 0.005). The fixation index (F_{ST}) was 0.028 and is a measure of population differentiation due to genetic structure. A fixation index value approximating zero is expected under random mating. The inbreeding coefficient (F_{IS}) was −0.33 indicating high heterozygosity across the Pacific and Australian cacao populations.

Phylogenetic analysis

The phylogenetic analysis (Figure 1) clearly groups the reference samples that included individual genotypes representative of 10 previously defined genetic populations or germplasm groups including Criollo, Trinitario, Amelonado, Nacional, Parinari (Marañón), Ucayali (Contamana), LCT EEN (Curaray), Guiana, Iquitos (IMC), Nanay and Purus (Motamayor et al. 2008). The Pacific islands and Australia have representatives of all the 10 defined genetic *T. cacao* populations (Figure 2).

Australia

Australia, particularly the far-north Queensland region, has recently emerged as a small-scale cocoa producer. The genetic source for cocoa cultivation in the area primarily relies on SG2 hybrid seeds, obtained either from, what was the former PNG Coconut and Cocoa Research Institute (CCRI) or from earlier introductions from Papua New Guinea (PNG). The SG2 hybrids were bred by CCRI in PNG and were originally released as a poly-cross hybrid of 15 different crosses and were reported to be the most widely grown varieties in PNG (Efron et al. 2003a, 2003b). Five SG2 hybrids were introduced into Australia to establish trials to examine the feasibility of cocoa production in northern Australia (Diczbalis et al. 2010). Through seedling selection samples conducted as part of the project, it has been observed that cocoa genetics exhibit a highly diverse nature. The samples represent various genetic backgrounds, including Amelonado, Guiana, Parinari (Marañón), Purus, IMC (Iquitos), Criollo, LCT EEN (Curaray), Nacional, and Ucayali (Contamana). The Principal Coordinates Analysis (PCoA) shows the genetic relationship between the Australian samples and the reference clones (Figure 3). Variation is 17% and 10.4% on axis 1 and 2, respectively.

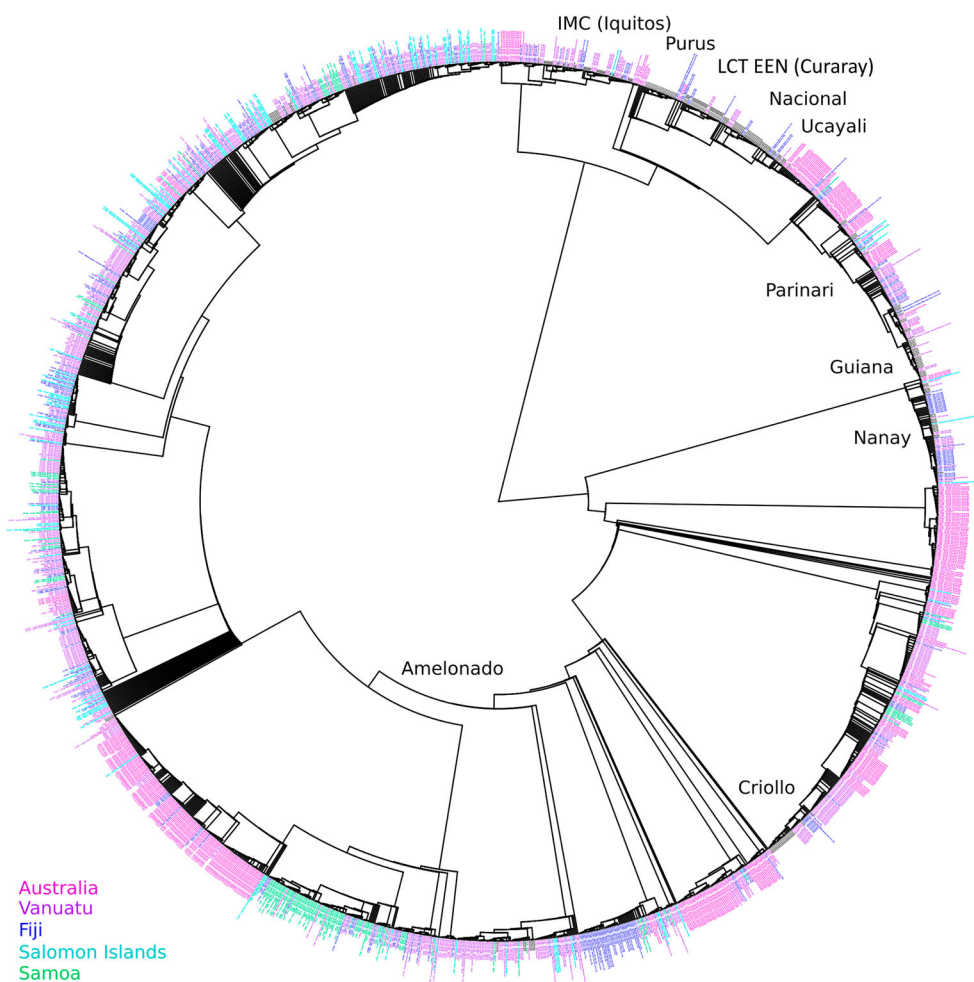


Figure 1. Phylogenetic analysis of *T. cocoa* samples from Australia, Fiji, Samoa, Solomon Islands, and Vanuatu compared to 169 references selections representing 10 genetic clusters of origin.

Fiji

Cocoa genetics were evaluated in various locations including the Ministry of Agriculture (MOA) research stations at Naduraloulou, Wainigata, and Dobuilevu, as well as in farmers' fields situated in the primary cocoa-producing regions of Viti Levu and Vanua Levu. The Principal Coordinates Analysis (PCoA) shows the genetic relationship between the Fijian samples and the reference clones (Figure 4). Variation is 16.6% and 15.9% on axis 1 and 2, respectively. The research clonal collections (Naduraloulou, Wainigata, and Dobuilevu) encompass representatives from eight out of the 10 genetic populations, namely, Amelonado, IMC (Iquitos), Nanay, Parinari (Marañón), Guiana, Nacional, Ucayali (Contamana), and Purus. On the other hand, selections obtained from farmers' fields primarily consist of Amelonado, while also including family attributes of Criollo, IMC (Iquitos), Nanay, and Parinari (Marañón). The MOA is currently promoting Trinitario types of cocoa clones, which possess a Criollo content ranging from 50% to 60%.

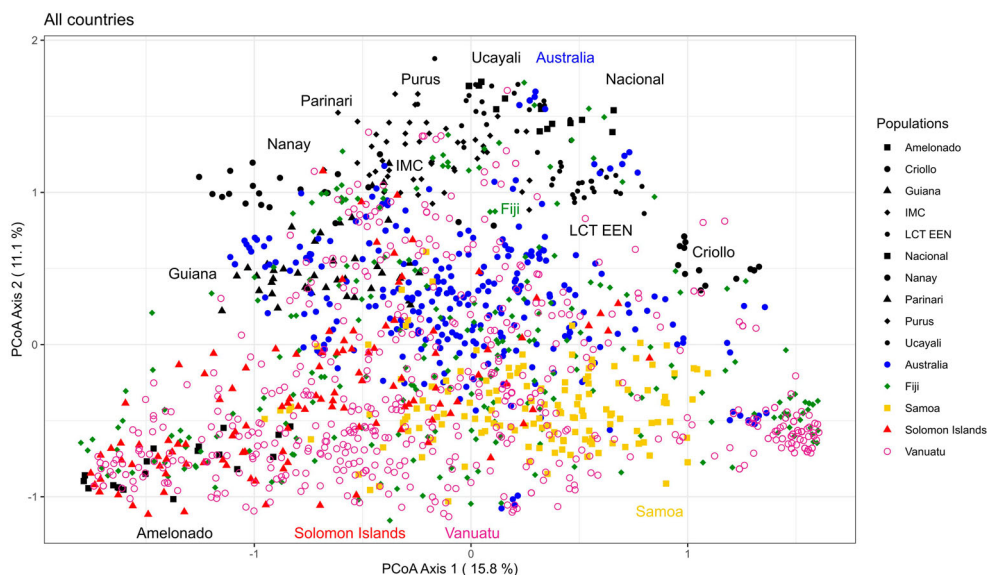


Figure 2. PCoA of *T. cacao* samples from Australia, Fiji, Samoa, the Solomon Islands, and Vanuatu compared to 169 references selections representing 10 genetic clusters of origin.

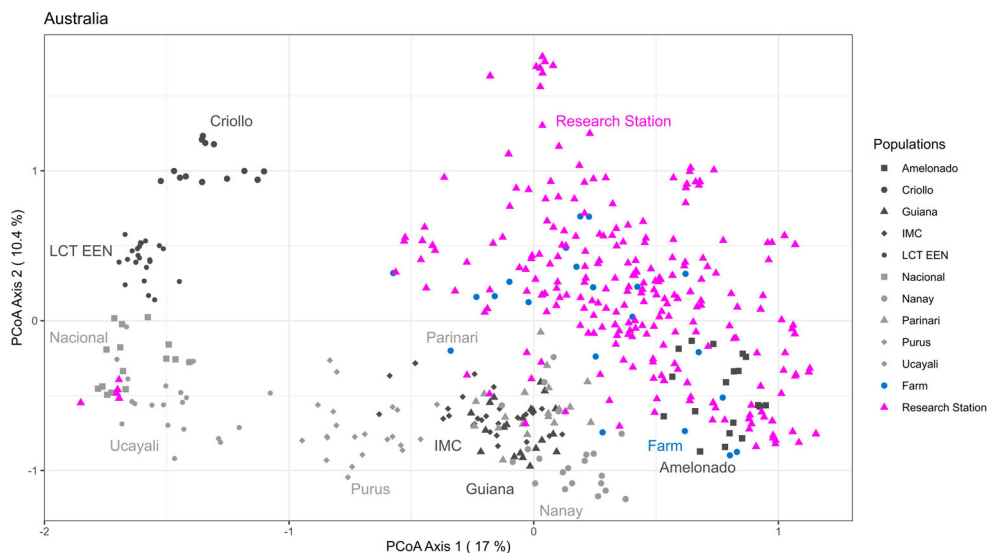


Figure 3. PCoA of *T. cacao* samples from Australia.

Samoa

Genetic sampling conducted at the Nu'u Research station and in farmers' fields reveals that the cocoa seedling population demonstrates a significant level of genetic homogeneity and is predominantly of Trinitario origin (Figure 5). All the samples obtained fall

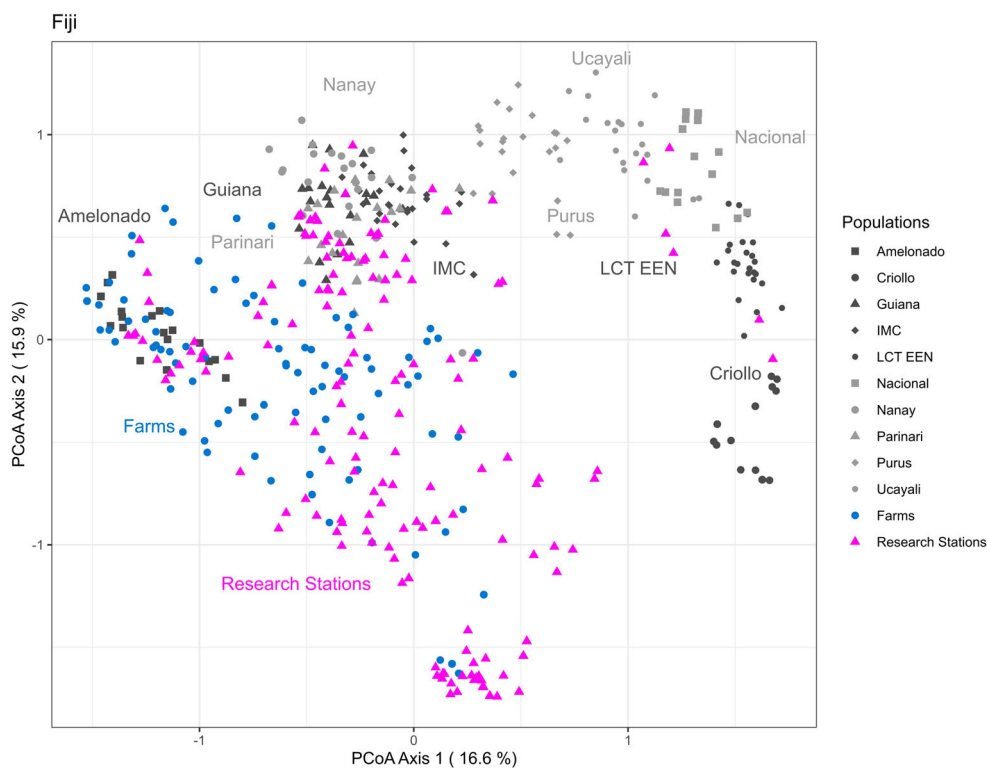


Figure 4. PCoA of *T. cacao* samples from Fiji.

within between the Amelonado and Criollo families. Some samples exhibit a combination of IMC (Iquitos), Nanay, or Parinari (Marañón), with Criollo and/or Amelonado in their genetic makeup. The Principal Coordinates Analysis (PCoA) shows the genetic relationship between the Samoan samples and the reference clones. Variation is 18.5% and 10% on axes 1 and 2, respectively.

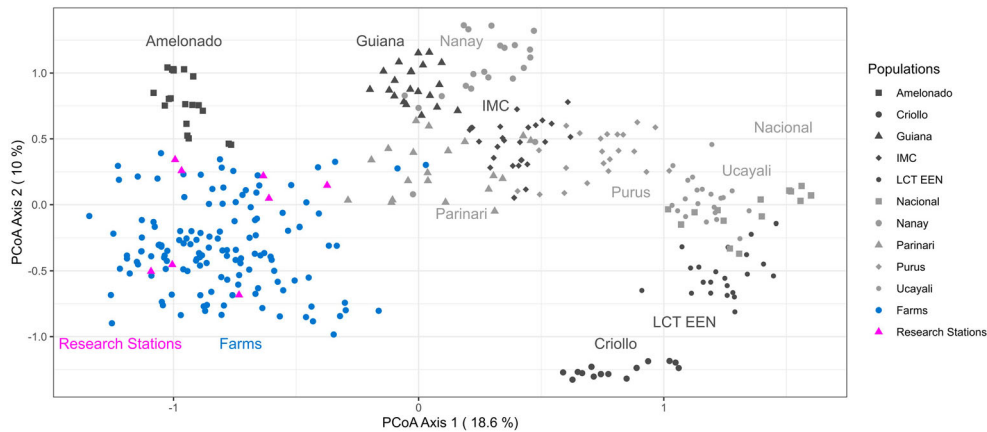


Figure 5. PCoA of *T. cacao* samples from Samoa.

Solomon Islands

The sampling process has been limited to farmers' fields in the primary cocoa-producing provinces, namely Guadalcanal, Makira, Malaita, Western, and Santa Isabel. The majority of samples obtained from farmers' fields consist of Amelonado cocoa, while some examples of Parinari (Marañón) and Nanay genetics are also present (Figure 6). A significant number of samples exhibit a substantial Criollo content, indicating Trinitario types, along with other genetic backgrounds such as National. The Principal Coordinates Analysis (PCoA) shows the genetic relationship between the Samoan samples and the reference clones. Variation is 22.2% and 8.1% on axes 1 and 2, respectively.

Vanuatu

Cocoa genetics were assessed through extensive sampling conducted at the main research collection at Vanuatu Agricultural Research and Technical Centre (VARTC) on Espiritu Santo Island, as well as in farmers' fields located in the primary cocoa-producing regions. The clonal collection at the research station boasts a diverse genetic composition, incorporating representatives from all 10 genetic family groupings, namely Amelonado, Criollo, IMC (Iquitos), Nanay, Parinari (Marañón), Guiana, LCT EEN (Curaray), National, Ucayali (Contamana), and Purus (Figure 7). Conversely, samples collected from farmers' fields in Aore, Efate, Epi, Malekula, and Malo predominantly consist of Amelonado cocoa, while also containing examples of Criollo, Nanay, IMC (Iquitos), Guiana, and Parinari (Marañón). The Principal Coordinates Analysis (PCoA) shows the genetic relationship between the Samoan samples and the reference clones. Variation is 21.1% and 9.4% on axes 1 and 2, respectively.

Discussion

The improved understanding of *T. cacao* genetics available in each of the investigated Pacific countries presents an exciting opportunity for cocoa researchers and producers to select and develop unique *T. cacao* clones that are tailored to specific regions. By studying the genetic diversity of *T. cacao* varieties within a country, cocoa researchers can identify material of interest that possesses desirable traits such as yield potential and organoleptic qualities.

The genetic population ancestry of *T. cacao* plays a crucial role in determining the flavour profiles and other characteristics of the resulting chocolate (Sukha and Butler

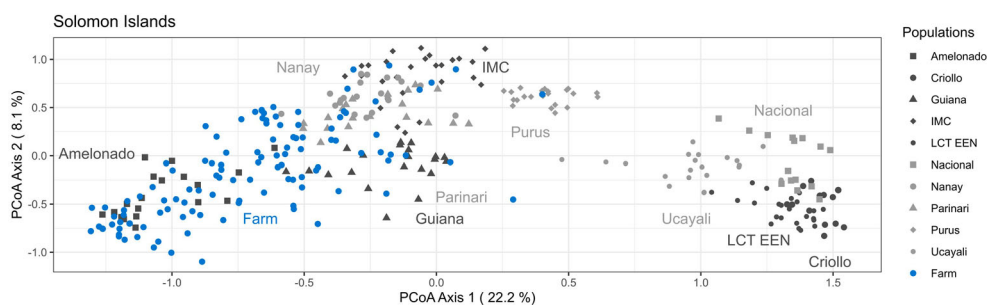


Figure 6. PCoA of *T. cacao* samples from Solomon Islands.

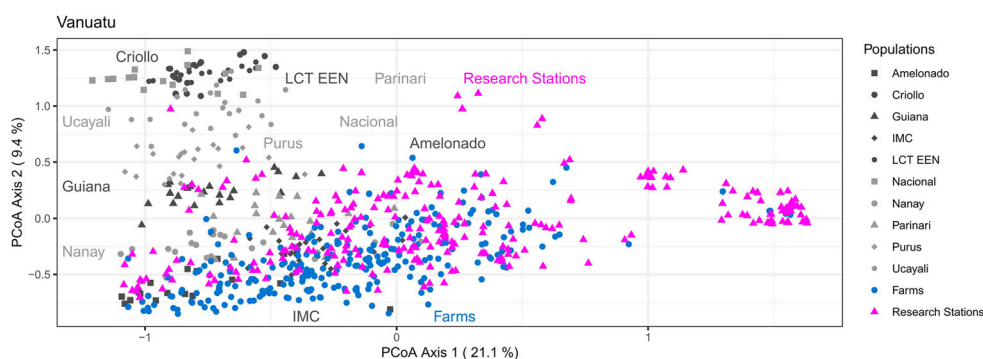


Figure 7. PCoA of *T. cacao* samples from Vanuatu.

2006; Sukha 2008; Sukha et al. 2017; Colonges et al. 2022). By selecting specific genetic ancestry, researchers can target desired flavour profiles and work towards developing unique and distinct flavours that are representative of the country's cocoa production.

Furthermore, the terroir, or environmental factors, of each Pacific nation can have a significant impact on cocoa flavour. Factors such as soil composition, climate, and altitude contribute to the development of specific flavour compounds in cocoa beans. By considering the interplay between genetics and terroir, researchers can identify cocoa clones that are best suited for a particular region and have the potential to showcase unique flavour profiles that reflect the country's environmental conditions. Implementing specific post-harvest techniques, such as fermentation and drying, producers can further enhance and refine the flavour characteristics of their cocoa beans.

In general, fine flavour cocoa is produced from Criollo or Trinitario varieties with bulk cocoa produced from Forastero varieties (ICCO 2023). Preliminary sensory analysis on samples from the Pacific has identified fruity, malty, buttery, and sweet flavour nuances (data not shown), based on the current genetic structure, indicative of fine flavour Trinitario type cacao (Diczbalis et al. 2023; P. Tyler, pers. comm).

Cocoa beans undergo a complex post-harvest process that alters both their physical and chemical properties, developing individual flavour characteristics based on their background genetics, growing conditions, and processing techniques. Within these genetic groups, there is further diversity that contributes to variations in flavour profiles. Factors such as specific gene expressions, genetic mutations, and the interaction of different genes influence the development of flavour compounds in cocoa beans. The presence and concentration of flavour compounds, such as volatile aromas and polyphenols, can vary depending on the genetic makeup of the *T. cacao* variety. Additionally, environmental factors, such as soil composition, climate, and altitude, can interact with the genetics of cacao trees and further shape flavour profiles. The same cacao variety grown in different regions can exhibit variations in flavours due to these environmental influences.

Post-harvest processing and chocolate-making techniques also contribute to the final flavour of chocolate. A shorter fermentation time is usually required for fine flavour cocoa compared to that for bulk cocoa (Afoakwa et al. 2008). Factors like fermentation, drying, roasting, and conching can all impact the development and retention of specific flavour compounds. Poor fermentation may result in the loss of previously present fine

flavour aroma. However, the genetics of the cocoa bean are the foundation stone in the pyramid of variables that influence cocoa and hence chocolate flavour (Kongor et al. 2016). Pacific Island nations producing cocoa can utilise the genetic variability that exists within their cocoa collections, both clonal and seedling selections, to create unique lines and fine flavour components linked to their origins.

Conclusions

The early introduction of superior *T. cacao* from Java and Ceylon, and the subsequent introductions to build the *T. cacao* collections at the research stations, into the Pacific region has led to the rich collection of varieties found today. Overall, the combination of improved cocoa genetics understanding, consideration of terroir, and careful post-harvest practices offers an opportunity for Pacific nation producers to develop a range of unique *T. cacao* clones that are representative of their country's genetic diversity and environmental conditions. This allows for the creation of distinct flavours that can contribute to the development of a thriving cocoa industry and promote the recognition and appreciation of each country's cocoa heritage.

Acknowledgements

The authors would like to acknowledge the farmers in Australia, Fiji, Samoa, Solomon Islands and Vanuatu who generously allowed us access to their properties to collect *T. cacao* samples. The authors acknowledge the managers, field and nursery staff at each of the research stations for their assistance to collect samples. The authors acknowledge Louise Hucks (DAF) for technical assistance and Stephen Pinney (ARS-USDA) for performing SNP genotyping.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Australian Centre for International Agriculture Research (ACIAR) under grant HORT/2014/078 Aligning genetic resources, production, and postharvest systems to market opportunities for Pacific Island and Australian cocoa and under grant PC/2014/032 Evaluation of molecular marker technology for identification of elite cocoa germplasm in the South Pacific.

ORCID

Natalie L. Dillon  <http://orcid.org/0000-0002-9818-6354>
Dapeng Zhang  <http://orcid.org/0000-0001-8212-6114>
Lars Nauheimer  <http://orcid.org/0000-0002-2847-0966>

References

Afoakwa EO, Paterson A, Fowler M, Ryan A. 2008. Flavor formation and character in cocoa and chocolate: a critical review. *Critical Reviews in Food Science and Nutrition*. 48(9):840–857. doi: 10.1080/10408390701719272.

- Aprotosoie AC, Luca SV, Miron A. **2016**. Flavor chemistry of cocoa and cocoa products – an overview. *Comprehensive Reviews in Food Science and Food Safety*. 15(1):73–91. doi: [10.1111/1541-4337.12180](https://doi.org/10.1111/1541-4337.12180).
- Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X, Risterucci AM, Da Silva C, Cascardo J, Allegre M, et al. **2008**. Towards the understanding of the cocoa transcriptome: production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. *BMC Genomics*. 9:512. doi: [10.1186/1471-2164-9-512](https://doi.org/10.1186/1471-2164-9-512).
- Argout X, Salse J, Aury JM, Gaultinan MJ, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova SN, et al. **2011**. The genome of *Theobroma cacao*. *Nature Genetics*. 43:101–108. doi: [10.1038/ng.736](https://doi.org/10.1038/ng.736).
- Ascrizzi R, Flamini G, Tessieri C, Pistelli L. **2017**. From the raw seed to chocolate: volatile profile of *Blanco de Criollo* in different phases of the processing chain. *Microchemical Journal*. 133:474–479. doi: [10.1016/j.microc.2017.04.024](https://doi.org/10.1016/j.microc.2017.04.024).
- Bartley BGD. **2005**. The genetic diversity of cacao and its utilization. Wallingford: CABI Publishing.
- CocoaNet. **2012**. A global strategy for the conservation and use of cacao genetic resources, as the foundation for a sustainable cocoa economy. B. Laliberté, compiler. Montpellier (France): Bioversity International.
- Colonges K, Solorzano RG, Jimenez J-C, Lahon M-C, Seguíne E, Calderon D, Subia C, Sotomayor I, Fernández F, Lebrun M, et al. **2022**. Variability and genetic determinants of cocoa aromas in trees native to South Ecuadorian Amazonia. *Plants People Planet*. 4:618–637. doi: [10.1002/ppp3.10268](https://doi.org/10.1002/ppp3.10268).
- Diczbalis Y. **2023**. Aligning genetic resources, production and post-harvest systems to market opportunities for Pacific Island and Australian cocoa (HORT/2014/078). Australian Centre for International Agricultural Research (ACIAR) Final Report.
- Diczbalis Y, Lemin C, Richards N, Wicks C. **2010**. Producing cocoa in Northern Australia. Rural Industries Research and Development Corporation. Publication Number 09/092.
- Dillon N, Diczbalis Y, Oakeshott J, Nagalevu P. **2020**. Aligning Pacific cocoa genetics to productivity and quality for the craft speciality chocolate market. In: *Proceedings of the Third International Tropical Agriculture Conference (TROPAG 2019)*; 2019 Nov 11–13; Brisbane, Australia.
- Efron Y, Epaina P, Marfu J. **2003b**. Breeding strategies to improve cocoa production in Papua New Guinea. In: *International workshop on cocoa breeding for improved production systems*; 2003 Oct 19–21; Accra, Ghana. p. 12–32.
- Efron Y, Epaina P, Taisa S. **2003a**. Analysis of the factors affecting yield and yield variability in the SG2 cocoa hybrid variety in Papua New Guinea. In: *International workshop on cocoa breeding for improved production systems*; 2003 Oct 19–21; Accra, Ghana.
- Fang W, Meinhardt LW, Mischke S, Bellato CM, Motilal L, Zhang D. **2014**. Accurate determination of genetic identity for a single cacao bean, using molecular markers with a nanofluidic system, ensures cocoa authentication. *Journal of Agricultural and Food Chemistry*. 62(2):481–487. doi: [10.1021/jf404402v](https://doi.org/10.1021/jf404402v).
- International Cocoa Organization (ICCO). **2023**. Fine Flavor Cocoa. [accessed 2023 June 2]. <http://www.icco.org/fine-or-flavor-cocoa.html>.
- Kalinowski ST, Manlove KR, Taper ML. **2007**. ONCOR: software for genetic stock identification. Bozeman: Montana State University.
- Kongor JE, Hinneh M, Van de Walle D, Afoakwa EO, Boeckx P, Dewettinck K. **2016**. Factors influencing quality variation in cocoa (*Theobroma cacao*) bean flavour profile – a review. *Food Research International*. 82:44–52. doi: [10.1016/j.foodres.2016.01.012](https://doi.org/10.1016/j.foodres.2016.01.012).
- Li Y, Zhang D, Motilal LA, Lachenaud P, Mischke S, Meinhardt LW. **2021**. Traditional varieties of cacao (*Theobroma cacao*) in Madagascar: their origin and dispersal revealed by SNP markers. *Beverage Plant Research*. 1:4. doi:[10.48130/BPR-2021-0004](https://doi.org/10.48130/BPR-2021-0004).
- Lukman, Zhang D, Susilo AW, Dinarti D, Bailey B, Mischke S, Meinhardt LW. **2014**. Genetic identity, ancestry and parentage in farmer selections of cacao from Aceh, Indonesia revealed by

- single nucleotide polymorphism (SNP) markers. *Tropical Plant Biology*. 7(3–4):133–143. doi: [10.1007/s12042-014-9144-6](https://doi.org/10.1007/s12042-014-9144-6).
- Marfu J. 2012. Review of Genetic Resources available in national collections for selection and distribution to farmers in Samoa, Fiji and Vanuatu. A report compiled by the PNG Cocoa Coconut Institute Limited for the Secretariat of the Pacific Community Land Resources Division. Unpublished.
- Motamayor JC, Lachenaud P, da Silva E, Mota JW, Loo R, Kuhn DN, Brown JS, Schnell RJ. 2008. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE*. 3(10):e3311. doi: [10.1371/journal.pone.0003311](https://doi.org/10.1371/journal.pone.0003311).
- Peakall R, Smouse PE. 2006. GenAEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 6:288–295.
- Peakall R, Smouse PE. 2012. GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 28(19):2537–2539. doi: [10.1093/bioinformatics/bts460](https://doi.org/10.1093/bioinformatics/bts460).
- Sukha DA. 2008. The influence of processing location, growing environment and pollen donor effects on the flavour and quality of selected cacao (*Theobroma cacao* L.) genotypes [PhD thesis]. St Augustine, Trinidad and Tobago: Department of Chemical Engineering of the University of the West Indies; p. 283.
- Sukha DA, Butler DR. 2006. The CFC/ICCO/INIAP Cocoa Flavour Project – investigating the spectrum of fine flavour within genotypes and between origins. *INGENIC Newsletter*. 10:22–25.
- Sukha DA, Umaharan P, Butler DR. 2017. Evidence for applying the concept of “Terroir” in cocoa (*Theobroma cacao* L.) flavour and quality attributes. 2017 International Symposium on Cocoa Research (ISCR); Nov 13–17; Lima, Peru.
- Wang J, Lin M, Crenshaw A, Hutchinson A, Hicks B, Yeager M, Berndt S, Huang WY, Hayes RB, Chanock SJ, et al. 2009. High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. *BMC Genomics*. 10:561. doi: [10.1186/1471-2164-10-561](https://doi.org/10.1186/1471-2164-10-561).