

Traditional varieties of cacao (*Theobroma cacao*) in Madagascar: their origin and dispersal revealed by SNP markers

Yanmei Li^{1,2}, Dapeng Zhang^{1*}, Lambert A. Motilal³, Philippe Lachenaud⁴, Sue Mischke¹, and Lyndel W. Meinhardt¹

¹ USDA/ARS, NEA, Beltsville Agricultural Research Center, SPCL, 10300 Baltimore Avenue, Bldg. 001, Rm. 223, BARC-W, Beltsville, MD 20705, USA

² Southwest University, No. 2 Tiansheng Road, Beibei District, Chongqing 400715, China

³ Cocoa Research Unit, The University of the West Indies, St. Augustine, Trinidad, Rep. Trinidad and Tobago, West Indies

⁴ Centre de Coopération Internationale en Recherche Agronomique pour le Développement, UPR Bioagresseurs de Pérennes, TA 80/02, 34398 Montpellier Cedex 5, France

* Corresponding author, E-mail: dapeng.zhang@usda.gov

Abstract

Cacao (*Theobroma cacao* L.) is an important Neotropical crop originating in South America and dispersed by European explorers, arriving in Madagascar in the late 19th century. Although Madagascar is an important producer of cocoa for the premium chocolate market, the varietal composition and genetic diversity in cacao germplasm from Madagascar, especially in traditional cacao farms, remains unknown. A total of 190 cacao accessions, including 40 farmer accessions collected from traditional cacao farms in Madagascar, and 150 accessions representing seven reference cacao populations, were analyzed using Single Nucleotide Polymorphism (SNP) markers. Multivariate analysis and Bayesian stratification resulted in the clustering of the 40 farmer accessions into three groups: Criollo, Amelonado and Trinitario. These three traditional varieties were commonly cultivated in tropical America in the 18th century, but most of them have been replaced by improved varieties. The present study demonstrated that Madagascar is distinctive in that all three traditional cacao varieties, Criollo, Amelonado and Trinitario, are still maintained on-farm for cocoa production, as in Mesoamerica and the Caribbean several hundred years ago. Results from the present study are significant in terms of understanding the early dispersal of cacao from tropical America and Asia to Africa, in addition to the well-documented route from Brazil to São Tomé & Príncipe. The results also provide new information for planning future conservation and utilization of cacao germplasm in Madagascar.

Citation: Li Y, Zhang D, Motilal LA, Lachenaud P, Mischke S, et al. 2021. Traditional varieties of cacao (*Theobroma cacao*) in Madagascar: their origin and dispersal revealed by SNP markers. *Beverage Plant Research* 1: 4 <https://doi.org/10.48130/BPR-2021-0004>

INTRODUCTION

Cacao (*Theobroma cacao* L.), a perennial crop native to the South American rainforest, has its center of diversity in the upper Amazon^[1–3]. Cacao cultivation started in Mesoamerica as long as 3,000 years ago. The earliest evidence for cacao use is from ancient Olmec and Maya pottery dating to about 1900–1500 BC^[4–6].

Cultivated cacao is traditionally subdivided into Criollo, Forastero and Trinitario^[7,8]. Criollo was thought to be the only cacao variety cultivated in Mesoamerica before Europeans arrived^[2,9,10]. Forastero cacao encompasses a range of populations from South America^[2,11] and the Trinitario group is believed to be hybrids between Criollo and Forastero germplasm from Venezuela^[2,7]. Today, few cultigens are ancient Criollo, although this type can still be found in some Mesoamerican rainforests^[9,12,13]. Criollo and many Trinitario are renowned for their distinct aroma and flavor, making them preferred raw materials for fine flavored chocolate. The fine flavored farmer varieties are highly valuable for the international premium chocolate market or for future breeding of new cacao varieties with improved quality attributes^[14].

Dispersal of cacao from tropical America to Southeast Asia

started in 1560, when the Spaniards introduced 'Venezuelan Criollo', a fine flavor variety, from tropical America to Indonesia^[15]. Cacao production started in northern Sulawesi where cocoa was processed and consumed locally^[16]. Another introduction of the Criollo variety was from Mexico to Indonesia in 1670, via the Acapulco-Manila galleons^[2,17]. The crop is thought to have then spread to other regions, including the Malay Peninsula and areas of South Asia^[15].

A cacao, with typical Criollo appearance, was introduced to Madagascar in the late 1800's and cultivated in the Sambirano valley^[18]. This first introduced variety was called 'Madagascar Criollo'^[15] and the putative source was Reunion Island^[19]. Fauchère speculated that Reunion had received this variety, similar to 'Old Red Ceylon', from Ceylon (Sri Lanka), where it had been initially called 'Venezuela Criollo'. A second introduction to Madagascar, in the early 1900's, was believed to be a Trinitario type, locally called 'Tamatave', suggesting introduction on the east coast^[19].

In 1910, Madagascar only produced 'Madagascar Criollo', exporting 28 tons of cacao beans^[15]. In 2017, Madagascar had an annual cocoa production of 8,000 metric tons^[20]. Despite relatively small-scale production, the superior quality of the Madagascar cacao bean is recognized among chocolate

experts worldwide. So far a series of studies on assessment of genetic diversity have been carried out in West and Central African countries, but little information is available in Madagascar, Malawi and East African countries, such as Tanzania and Uganda^[21]. In this study, we used 96 SNP markers to genotype farmer varieties collected from traditional farms in Madagascar. Our objective was to assess genetic identity of these fine-flavored varieties and elucidate their origin. Our results will help establish baselines for sustainable conservation and utilization of Madagascar cacao germplasm.

MATERIALS AND METHODS

Plant materials

A total of 190 samples, including 40 farmer clones sampled from four traditional farms in the Sambirano valley, in northwest Madagascar (Fig. 1), and 150 reference clones were used (Table 1).

For the Madagascar farmer varieties, four healthy young leaves were collected from each tree. Dry samples were prepared using silica gel and the dried leaves were sent to the USDA Beltsville Agricultural Research Center, Maryland, USA. The 150 reference samples were collected from various *ex situ* national and international genebanks.

SNP genotyping

The DNeasy® Plant Mini kit (Qiagen Inc., Valencia, CA, USA) was used to extract DNA from the dried cacao leaves following the specified kit instruction. The dry leaf tissue was placed in a 2-mL microcentrifuge tube with one ¼-inch ceramic sphere and 0.15 g garnet matrix (Lysing Matrix A; MP Biomedicals, Solon, OH, USA). Tissue disruption was performed in two 1 minute high-speed (30 Hz) shaking steps in a TissueLyser II (Qiagen Inc.). After that, lysis solution (DNeasy® kit buffer AP1 containing 25 mg mL⁻¹ polyvinylpo-

lypyrrolidone) and RNase A were added to the powdered leaf samples and the mixture was incubated at 65 °C. The remainder of the extraction method followed the manufacturer's suggestions. DNA was eluted from the silica column with two washes of 50 µL Buffer AE, which were pooled, resulting in 100 µL DNA solution. DNA concentration was determined using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE, USA) with absorbance at 260 nm. DNA purity was estimated by the absorbance of 260:280 ratio and the 260:230 ratio.

SNP loci were identified from expressed sequence tags (ESTs) from a wide range of cacao plant parts that displayed a good representation of the cacao transcriptome^[22,23]. The selection of 96 SNP loci for the panel used in the current study was chosen based on the screened result of 1536 SNPs using Illumina's Golden Gate Assay (Michel Boccara, unpublished data) and their application in previously reported research^[24–27]. The detailed sequence information of the 96 SNPs is presented in Supplemental Table S1. The protocol for SNP genotyping of cacao used the Fluidigm 96.96 Dynamic Array™ (Fluidigm, San Francisco, CA). The genotyping followed the procedure of a 96.96 Dynamic Array IFC, as described by Wang et al.^[28]. The genotyping results were acquired on an EP1™ imager (Fluidigm Corp.), recorded with Fluidigm Genotyping Analysis Software (Fluidigm, San Francisco, CA) and exported for data analysis.

Data analysis

The exported raw data of SNP locus and sample calls was first organized in Microsoft Excel 2007. Methods of multivariate analysis, as implemented in GenAlEx 6.5^[29,30], was used to assess the relationship among the individual cacao samples from Madagascar, as well as their relationship with reference cacao groups. The DISTANCE procedure of GenAlEx 6.5^[29,30] was performed to compute pairwise genetic distances. Principal Coordinates Analysis (PCoA), as imple-

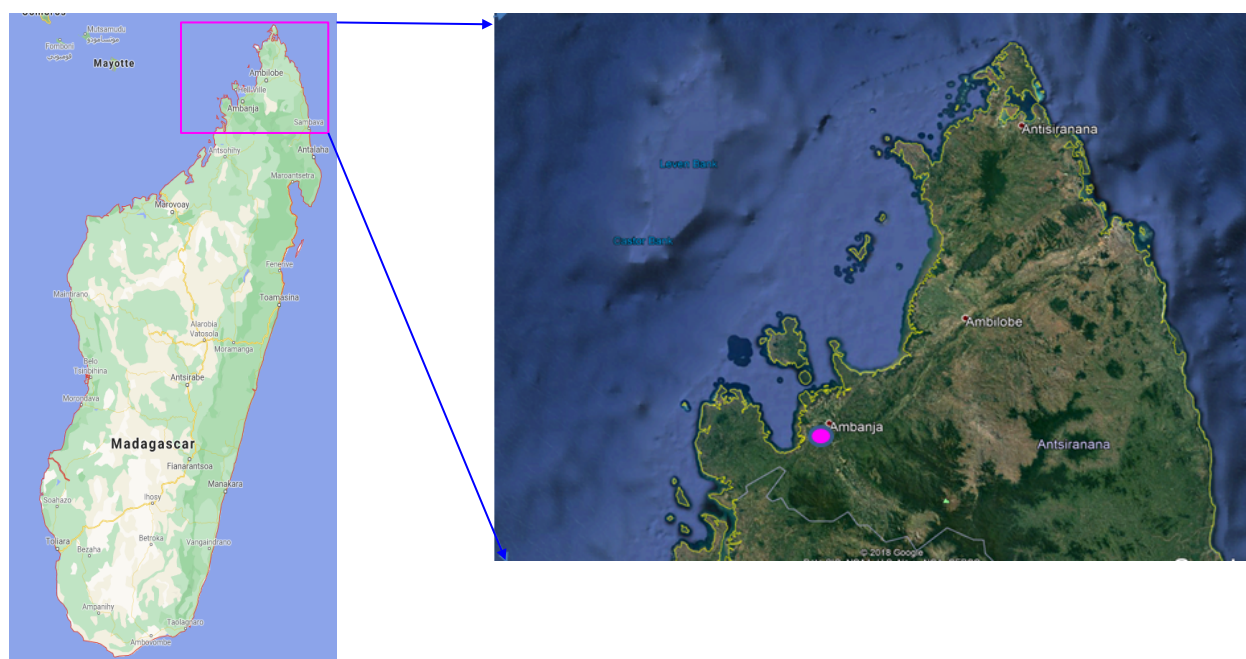


Fig. 1 Map showing sample collecting site of traditional farmer clones in the Sambirano valley, in northwest Madagascar.

Table 1. List of 190 cacao accessions, including 40 farmer clones from Madagascar and 150 reference accessions used in SNP genotyping.

#	Accession	Type of variety*
1	MA01	Criollo
2	MA02	Amelonado
3	MA03	Criollo
4	MA04	Amelonado
5	MA05	Amelonado
6	MA06	Criollo
7	MA07	Criollo
8	MA08	Amelonado
9	MA10	Amelonado
10	MA11	Criollo
11	MA12	Criollo
12	MA13	Amelonado
13	MA14	Trinitario
14	MA15	Criollo
15	MA16	Trinitario
16	MA17	Trinitario
17	MA18	Amelonado
18	MA19	Trinitario
19	MA20	Amelonado
20	Izaka1	Unknown
21	Izaka2	Unknown
22	Izaka3	Unknown
23	Izaka4	Unknown
24	Izaka5	Unknown
25	Izaka7	Unknown
26	Izaka8	Unknown
27	Izaka9	Unknown
28	Izaka10	Unknown
29	Somia11	Trinitario
30	Somia12	Amelonado
31	Somia13	Amelonado
32	Somia14	Trinitario
33	Somia15	Criollo
34	Somia16	Trinitario
35	Somia17	Criollo
36	Somia18	Trinitario
37	Somia19	Trinitario
38	Somia20	Amelonado
39	Somia21	Amelonado
40	Somia22	Trinitario

Reference accessions

Genetic group	Sample size	Country of origin	Source
IMC	20	Peru	ICGT, Trinidad
Nanay	20	Peru	ICGT, Trinidad
Parinari	20	Peru	ICGT, Trinidad
Nacional	20	Ecuador	INIAP, Ecuador
Scavina	20	Peru	ICGT, Trinidad
Amelonado	20	Puerto Rico, Nicaragua, Costa Rica, Trinidad	ICGT, Trinidad; CATIE, Costa Rica; SPCL, USDA
Criollo	20	Puerto Rico, Nicaragua, Costa Rica, Trinidad	ICGT, Trinidad; CATIE, Costa Rica; SPCL, USDA
Trinitario	10	Trinidad	ICGT, Trinidad
Total	190		

* Perceived type of variety by local farmers based on morphological characteristics.

mented in the same program, was applied based on the pairwise distance matrix. The standardized option was taken for both distance and covariance, as described by Peakall and Smouse^[29,30].

To analyze population structure of the 190 cacao samples, the model-based Bayesian cluster analysis software STRUCTURE v2.3.4^[31] was used. The 40 cacao samples were analyzed together with the 150 reference samples representing seven distinctive cacao germplasm groups, which were selected based on known passport data of germplasm collection, as well as the history of cacao germplasm utilization in Africa and Asia^[2]. The analysis used an admixed model with the number of clusters (K value) set to 7, corresponding to the possible cacao genetic groups presenting in the Madagascar cacao accessions. For K = 7, ten independent runs were performed using 200,000 iterations after a burn-in period of 100,000. The result was visualized using the program CLUMPP 1.1.2^[32] and DISTRUCT1.1^[33].

RESULTS AND DISCUSSION

The genetic relationships among the farmer accessions and reference clones were shown by Principal Coordinates Analysis (Fig. 2). The plane of the first three main PCO axes accounted for 40.5%, 30.2% and 5.7% of total variation respectively. The 190 samples were grouped into eight clusters representing seven original cacao germplasm groups of Criollo, Amelonado, Nacional, Iquitos Mixed Calabacillo (IMC), Scavina (SCA), Nanay (NA), and Parinari (PA), as well as 10 reference hybrid clones between Criollo and Amelonado (Trinitario). Out of the 40 Madagascar farmer accessions, nine were grouped with reference Criollo, 12 were grouped with Trinitario and 19 were grouped with Amelonado. No accessions from Madagascar were grouped with the cluster of Upper Amazon Forastero (Fig. 2).

The Bayesian clustering analysis generated a consistent result as with the PCoA, which separated the 190 samples into eight respective genetic groups: IMC, Nanay, SCA, PA, Nacional, Amelonado Criollo and Trinitario (Fig. 3). These eight groups represented all known cacao germplasm that

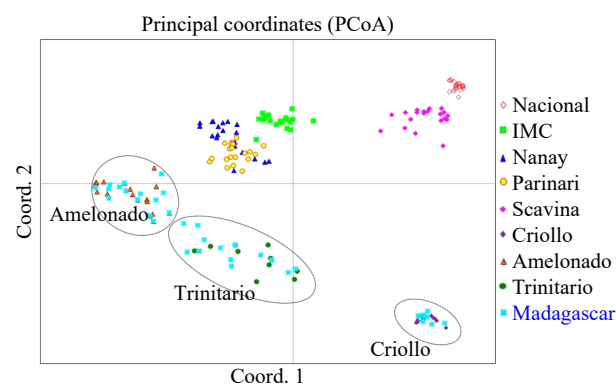


Fig. 2 PCoA plot of 190 cacao accessions, including 40 farmer varieties from Madagascar and 150 reference clones from the International Genebanks. The plane of the first three main PCO axes accounted for 76.3% of total variation. First axis = 40.5% of total information, the second = 30.16% and the third = 5.72%. All accessions correspond to the sample list in Table 1.

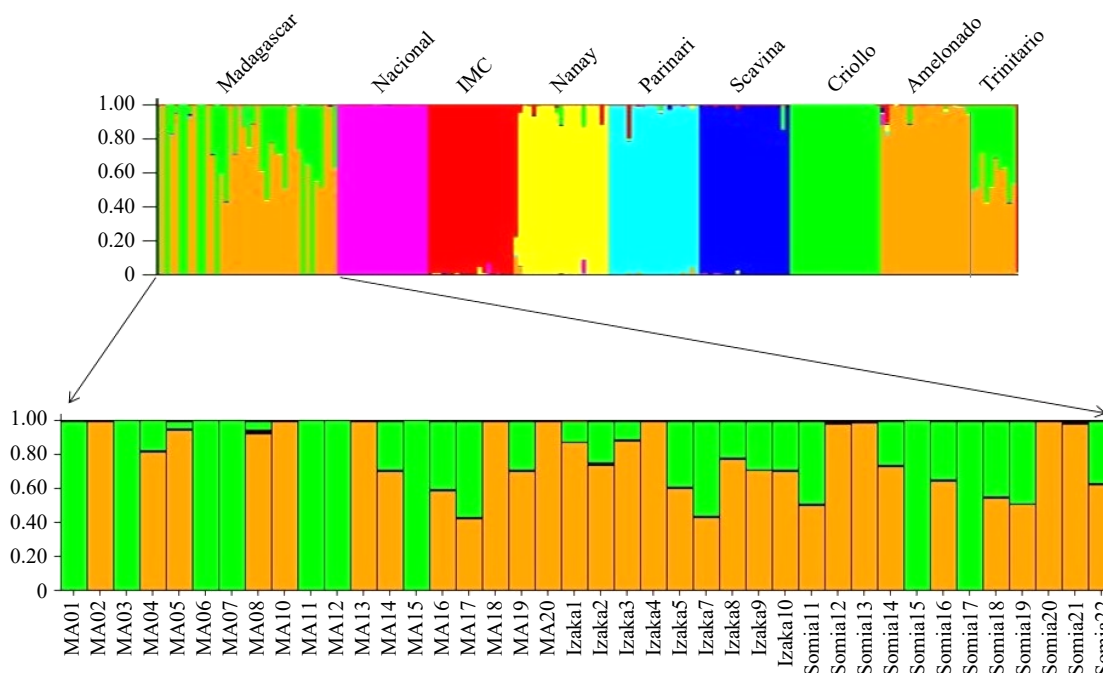


Fig. 3 Inferred clusters in the 40 Madagascar farmer selections and 150 reference international clones using STRUCTURE ($K = 7$), where K is the potential number of genetic clusters that may exist in the overall sample of individuals. Clusters of individuals are represented by colors. Individuals with multiple colors have admixed genotypes from multiple clusters. Each color represents the most likely ancestry of the cluster from which the genotype or partial genotype was derived.

were distributed to Asia and Africa before the 1990s^[34]. Therefore, they covered all possible sources of ancestry contribution to the current farmer varieties in Madagascar. We used the assignment coefficient (Q -value) as a criterion to classify the analyzed accessions. If an accession had an assignment coefficient ≥ 0.80 for a specific genetic group, this accession was classified as belonging to a single genetic group or having a single ancestral origin. On the other hand, if a genotype had a Q value below 0.80 for a specific genetic group, then this accession was denoted as a hybrid. The ancestral origin of a hybrid accession could be from two or more germplasm populations. Based on this criterion, there are nine Criollo, 16 Amelonado and 15 Trinitario in the 40 Madagascar samples (Fig. 3, Table 2). The 15 Trinitario accessions from Madagascar were assigned as hybrids with their admixed ancestry coming from Criollo and Amelonado (Fig 3, Table 2). Again, no ancestry of Upper Amazon Forastero was detected among the Madagascar samples. On average, the reference clusters of Criollo and Amelonado have a coefficient of membership (Q value) of 0.988 and 0.948 respectively.

SNP genotyping of the Madagascar farmer accessions revealed a clear germplasm group composition of three clusters corresponding to ancient Criollo, Amelonado and Trinitario. Despite there being over 100 years of cultivation since the early introduction of cacao into this island, ancient Criollo, Amelonado and Trinitario are still maintained in Madagascar. Results from the present study clarified the genetic identity of the so called 'Madagascar Criollo'^[15] which was cultivated as far back as the late 19th century on the island. Apparently Criollo was the first introduced group in Madagascar, which had previously been transferred from

Mesoamerica to the Philippines and Indonesia. From there, Criollo was transferred to the island of Mauritius at the beginning of the 18th century and to Madagascar in late 19th century via the island of La Reunion. Another account suggests that Criollo could have come to Madagascar from Venezuela by way of Trinidad, Ceylon (Sri Lanka) and La Reunion Island.

Our results also confirmed that different cacao germplasm groups were introduced into Madagascar later and suggest that the possible source of introduction may be East Africa. Cacao was cultivated in Zanzibar, Tanzania as early as 1891. In 1895, British Central Africa (now Malawi) imported 100 cacao pods from Grenada^[2]. These Grenada accessions might have been dispersed to Madagascar via East Africa. Given their original source in Mesoamerica and the Caribbean in the late 19th century, it's not a surprise that Amelonado and Trinitario were among the introduced germplasm. The current result is compatible with the local grower's observation that the so called 'Forastero' cacao in Madagascar has oval-shaped green pods with shallow rows, dense and hard shells and small flat purple beans, typical characteristics of Amelonado cacao. Varieties of the Amelonado type have better disease resistance and productivity than Criollo in Madagascar. It thus may have replaced the majority of the Madagascar Criollo in production. Moreover, it's worth noting that the classification of the collected varieties by local researchers, based solely on morphological characteristics (Table 1) is highly consistent with the result revealed by SNP analysis (Table 2). This consistency suggested that in Madagascar, the three traditional varieties, Criollo, Amelonado and Trinitario, have a clear difference in terms of morphological characteristics thus it's reliable to manage local germplasm using morphological

Table 2. Assigned parentage for 40 Madagascar cacao accessions based on seven reference germplasm groups, using Bayesian clustering analysis.

Name of accessions		Assigned parentage (Q-value) based on seven reference populations						Result of assignment
Name	Nacional	IMC	Nanay	Parinari	Scavina	Criollo	Amelonado	
MA01	0.000	0.001	0.000	0.000	0.000	0.997	0.000	Criollo
MA02	0.001	0.001	0.001	0.001	0.001	0.001	0.995	Amelonado
MA03	0.000	0.000	0.000	0.000	0.000	0.998	0.000	Criollo
MA04	0.006	0.023	0.023	0.012	0.002	0.118	0.816	Amelonado
MA05	0.001	0.001	0.001	0.001	0.001	0.047	0.948	Amelonado
MA06	0.000	0.000	0.000	0.000	0.000	0.998	0.000	Criollo
MA07	0.001	0.000	0.000	0.000	0.000	0.997	0.000	Criollo
MA08	0.001	0.004	0.005	0.009	0.008	0.047	0.926	Amelonado
MA10	0.000	0.001	0.001	0.002	0.001	0.001	0.995	Amelonado
MA11	0.000	0.000	0.000	0.000	0.000	0.998	0.000	Criollo
MA12	0.000	0.000	0.000	0.000	0.000	0.998	0.000	Criollo
MA13	0.001	0.001	0.001	0.002	0.001	0.001	0.994	Amelonado
MA14	0.001	0.001	0.001	0.001	0.001	0.291	0.705	Trinitario
MA15	0.000	0.000	0.000	0.000	0.000	0.997	0.000	Criollo
MA16	0.001	0.001	0.001	0.001	0.001	0.403	0.591	Trinitario
MA17	0.001	0.002	0.002	0.002	0.001	0.569	0.424	Trinitario
MA18	0.000	0.001	0.001	0.001	0.000	0.000	0.998	Amelonado
MA19	0.001	0.001	0.001	0.001	0.001	0.289	0.707	Trinitario
MA20	0.001	0.002	0.002	0.001	0.000	0.001	0.993	Amelonado
Izaka1	0.016	0.010	0.003	0.001	0.014	0.083	0.875	Amelonado
Izaka2	0.001	0.001	0.005	0.002	0.001	0.248	0.743	Trinitario
Izaka3	0.001	0.001	0.001	0.001	0.002	0.110	0.885	Amelonado
Izaka4	0.000	0.002	0.001	0.001	0.000	0.001	0.995	Amelonado
Izaka5	0.002	0.002	0.003	0.001	0.001	0.387	0.604	Trinitario
Izaka7	0.001	0.001	0.002	0.002	0.001	0.562	0.433	Trinitario
Izaka8	0.004	0.003	0.003	0.002	0.001	0.186	0.802	Amelonado
Izaka9	0.001	0.001	0.003	0.001	0.000	0.286	0.707	Trinitario
Izaka10	0.001	0.001	0.001	0.002	0.001	0.290	0.705	Trinitario
Somia11	0.001	0.001	0.001	0.002	0.001	0.488	0.507	Trinitario
Somia12	0.001	0.001	0.001	0.001	0.001	0.010	0.985	Amelonado
Somia13	0.000	0.002	0.007	0.002	0.001	0.001	0.986	Amelonado
Somia14	0.001	0.001	0.001	0.001	0.001	0.260	0.736	Trinitario
Somia15	0.000	0.000	0.000	0.000	0.000	0.998	0.000	Criollo
Somia16	0.001	0.001	0.002	0.002	0.001	0.344	0.650	Trinitario
Somia17	0.000	0.001	0.000	0.000	0.000	0.997	0.000	Criollo
Somia18	0.001	0.002	0.003	0.002	0.001	0.445	0.547	Trinitario
Somia19	0.001	0.001	0.001	0.001	0.001	0.487	0.507	Trinitario
Somia20	0.000	0.001	0.001	0.001	0.000	0.000	0.997	Amelonado
Somia21	0.001	0.005	0.005	0.002	0.001	0.002	0.982	Amelonado
Somia22	0.001	0.001	0.002	0.007	0.001	0.366	0.622	Trinitario

descriptors.

The origin of the Trinitario cacao in Madagascar has not been fully elucidated. The present results suggest one possible explanation being that Trinitario cacao was introduced to the island. In terms of their SNP profiles, the Trinitario clones in Madagascar were very similar to the Trinitario varieties collected from traditional farms in Nicaragua^[24]. The present study therefore suggests that the original forms of Trinitario, Amelonado and Criollo were transferred from Mesoamerica and the Caribbean region to Asia and the Pacific and then from there to Africa. This type of Trinitario can be referred to as 'classical Trinitario' because they only have parentage from Criollo and Amelonado. In contrast, some of the Trinitario cacao stayed in Mesoamerica and the Caribbean region was later influenced by other Upper Amazon Forastero germplasm and became a complex of hybrids, with parentage contributed by at least three original introduction events^[35]. Nonetheless, according to

Fabrice et al.^[36], Trinitario cacao in Madagascar was developed on the island from the 1960s to 1980s, by crossing the introduced Criollo and Amelonado varieties, by researchers in the National Center for Applied Research in Malagasy Rural Development, Madagascar.

Results from the present study are significant in terms of understanding the early dispersal of cacao from Mesoamerica to Africa, in addition to the well-documented route of Bahia, Brazil to São Tomé & Príncipe. Due to the large size of Madagascar and its isolated geographic setting, this island is home to many threatened and endangered species. About 80 percent of the plants and animals found there are found nowhere else in the world. The geographical isolation apparently played a significant role in maintaining the original genetic identities of the traditional varieties once commonly found in Mesoamerica. Although these varieties often have lower yield and are more vulnerable to a variety of biotic and abiotic threats, they often have preferred fine flavor for the

craft chocolate industry, which are increasingly sought after by gourmet specialty markets. The gourmet specialty markets and supply chains are providing opportunities to differentiate the small volumes of fine flavor cocoa from the rest of the bulk cocoa, so there is a price premium to offset the reduced yields. Conservation through use of threatened cacao diversity, therefore may provide incentives for a demand-driven on-farm conservation. Knowledge of on-farm genetic diversity will provide a sound scientific basis for logical decision-making and increase the credibility of the planning process for in-situ conservation. It's necessary to collect a set of representative samples of these traditional varieties and add them into the *ex situ* genebank. Meanwhile, selecting high yielding and fine flavor clones will promote the use of elite clones by farmers. The results also suggests a lack of host resistance to cacao diseases in the Madagascar cacao germplasm. Genetic improvement is needed to combine traits of fine flavor, diseases resistance and high yield, to provide a genetic foundation for sustainable production of fine flavor cocoa in Madagascar.

Madagascar is a relatively small cocoa producer. Approximately 33,000, mainly smallholder cocoa farmers, produce approximately 8,000 metric tons of cocoa beans per year, accounting for less than 0.2% of world production^[20]. Nonetheless, Madagascar is the only African country that is classified as a fine cocoa producer according to the ICCO. Malagasy cocoa production has been classified as 100 % fine cocoa by the ICCO since 2016^[37]. The cocoa sector is recognized to have the potential to boost economic growth and producers' income for local people^[38]. Cocoa production in Madagascar is concentrated in the Sambirano valley within the Ambanja district, where most cacao plantations are small farms. In recent years, the international cocoa price has seen large price fluctuation for bulk cocoa. The price fluctuation is mostly related to the trade of bulk cacao grown in the large producing countries, such as Ghana, Ivory Coast and Indonesia, which make up the majority of the world cacao supply. On the other hand, demand for fine flavor cocoa beans has been rapidly increasing in the international market^[39]. Promoting production of fine flavor cocoa that can achieve a premium price in high-end markets will likely create sustainable livelihoods for small holder farmers and provide local communities with job opportunities. Moreover, while cocoa production in other countries is largely grown in monocultures, requiring the removal of all surrounding trees, farmers in the Sambirano, Madagascar plant cacao in an agroforestry system. As a tropical understory crop, cacao is suitable for intercropping with many perennial species. The agroforestry system not only results in higher productivity through increased soil fertility and reduced risk of pests, diseases, and weed outbreaks, it also provides additional sources of income to small cocoa farmers. This system also initiates the process of agroecological restoration which improves carbon storage and nutrient cycling. Therefore, cocoa cultivation in the agroforestry system is beneficial for both environment and the people of Madagascar.

CONCLUSION

Using DNA fingerprinting with SNP markers, we showed that Madagascar is distinctive in that all three traditional

cacao varieties, Criollo, Amelonado and Trinitario, are still maintained on-farm for cocoa production, the same as was seen in Mesoamerica and the Caribbean several hundred years ago. Results from the present study are significant in terms of understanding the early dispersal of cacao from tropical America to Africa, in addition to the well-documented route from Brazil to São Tomé & Príncipe. The results also provide new information for planning sustainable conservation and utilization of traditional cacao varieties in Madagascar, especially for expanding the production of fine flavor cocoa for gourmet specialty markets. We acknowledge that the present study only analyzed samples from the Sambarino valley, where most traditional varieties exist. In the future, it's necessary to understand the genetic diversity in other regions of Madagascar. A full-scale survey of genetic diversity will be useful to support the development of the cocoa industry in Madagascar.

ACKNOWLEDGMENTS

The authors thank Brett Beach, Tim McCollum and Nate Eagle of Beyond Good Chocolate Company (www.beyondgood.com) for providing cacao leaf samples and passport information. The authors would also like to give special thanks to Stephen Pinney and Stephen Emche (USDA, ARS) for assistance with SNP genotyping using the nanofluidic array system and Sue Mischke for the editing of this manuscript. Mention of trade names or commercial products in this publication is solely for providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (<http://www.maxapress.com/article/doi/10.48130/BPR-2021-0004>)

Dates

Received 18 February 2021; Accepted 25 May 2021;
Published online 10 June 2021

REFERENCES

1. Cuatrecasas J, 1964. Cacao and its allies. A taxonomic revision of the genus *Theobroma*. *Contributions from the United States National Herbarium*. 35: 1–614. Washington, DC: Smithsonian Institution Press <http://www.jstor.org/stable/23493192>
2. Bartley BGD. 2005. *The genetic diversity of cacao and its utilization*. Wallingford, Oxfordshire: CAB International, CABI Publishing <https://www.cabi.org/bookshop/book/9780851996196>
3. Ogata N, Gómez-Pompa A, Taube KA. 2006. The Domestication and Distribution of *Theobroma Cacao* L. in the Neotropics. In *Chocolate in Mesoamerica: A Cultural History of Cacao*, Ed. McNeil CL. Gainesville, FL: University Press Florida. pp. 69–89 <http://www.doi.org/10.5744/florida/9780813029535.003.0003>
4. Gómez-Pompa A, Flores JS, Fernández MA. 1990. The sacred cacao groves of the Maya. *Latin American Antiquity* 1:247–57

5. Henderson JS, Joyce RA, Hall GR, Hurst WJ, McGovern PE. 2007. Chemical and archaeological evidence for the earliest cacao beverages. *Proc. Natl. Acad. Sci. U. S. A.* 104:18937–40
6. Powis TG, Cyphers A, Gaikwad NW, Grivetti L, Cheong K. 2011. Cacao use and the San Lorenzo Olmec. *Proc. Natl. Acad. Sci. U. S. A.* 108:8595–600
7. Cheesman E. 1944. Notes on the nomenclature, classification and possible relationships of cocoa populations. *Tropical Agriculture* 21:144–59
8. Toxopeus H. 1985. Botany, types and populations. In *Cocoa*, eds. Wood GAR, Lass RA. Oxford: Blackwell Science. pp. 17–38 <https://doi.org/10.1002/9780470698983.ch2>
9. Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A, et al. 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89:380–86
10. Motilal LA, Zhang D, Umaharan P, Mischke S, Mooleedhar V, et al. 2010. The relic Criollo cacao in Belize—genetic diversity and relationship with Trinitario and other cacao clones held in the International Cocoa Genebank, Trinidad. *Plant Genetic Resources* 8:106–15
11. Motamayor JC, Lachenaud P, da Silva e Mota JW, Loo R, Kuhn DN, et al. 2008. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE* 3:e3311
12. Mooleedhar V, Maharaj WA, O'Brien H. 1995. The collection of Criollo cacao germplasm in Belize. *Cocoa Grower's Bulletin* 49:26–40
13. Lachenaud P, Motamayor JC. 2017. The Criollo cacao tree (*Theobroma cacao* L.): a review. *Genetic Resources and Crop Evolution* 64:1807–20
14. CacaoNet. 2012. A Global Strategy for the Conservation and Use of Cacao Genetic Resources, as the Foundation for a Sustainable Cocoa Economy (B. Laliberté, compiler). Bioversity International, Montpellier, France
15. Hall CJ van. 1932. *Cocoa*. London: Macmillan. <https://mpl.biblio-commons.com/item/show/1157877075>
16. Susilo A, Zhang D, Motilal L, Mischke S, Meinhardt LW. 2011. Assessing genetic diversity in Java fine-flavor cocoa (*Theobroma cacao* L.) Germplasm by simple sequence repeat (SSR) markerseat (SSR) markers. *Tropical Agriculture and Development* 55:84–92
17. Blanco M. 1837. *Flora de Filipinas*. Manila: Imprenta de Santo Thome. 887 pp
18. Jumelle H. 1900. *Le cacaoyer*. Paris: Sa culture et son exploitation dans tous les pays de production Challamel Editions. 210 pp
19. Fauchère A. 1906. *Culture pratique du cacaoyer et préparation du cacao*. Paris: Challamel Eds. 175 pp
20. International Cocoa Organization. 2017. February 2017 Quarterly Bulletin of Cocoa Statistics. <https://www.icco.org/february-2017-quarterly-bulletin-of-cocoa-statistics>
21. Aikpokpodion P. 2012. Defining Genetic Diversity in the Chocolate Tree, *Theobroma cacao* L. Grown in West and Central Africa. In *Genetic Diversity in Plants*, ed Çalişkan M. Croatia: InTech. 498 pp. <http://www.doi.org/10.5772/33101>
22. Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, 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
23. Argout X, Salse J, Aury JM, Guitinan MJ, Droc G, et al. 2011. The genome of *Theobroma cacao*. *Nature Genetics* 43:101–9
24. Ji K, Zhang D, Motilal LA, Boccara M, Lachenaud P, et al. 2013. Genetic diversity and parentage in farmer varieties of cacao (*Theobroma cacao* L.) from Honduras and Nicaragua as revealed by single nucleotide polymorphism (SNP) markers. *Genetic Resources and Crop Evolution* 60:441–53
25. Lindo AA, Robinson DE, Tennant PF, Meinhardt LW, Zhang D. 2018. Molecular Characterization of Cacao (*Theobroma cacao*) Germplasm from Jamaica Using Single Nucleotide Polymorphism (SNP) Markers. *Tropical Plant Biology* 11:93–106
26. Cosme S, Cuevas HE, Zhang D, Oleksyk TK, Irish BM. 2016. Genetic diversity of naturalized cacao (*Theobroma cacao* L.) in Puerto Rico. *Tree Genetics & Genomes* 12:88
27. Arevalo-Gardini E, Meinhardt LW, Zuñiga LC, Arévalo-Gardni J, Motilal L, et al. 2019. Genetic identity and origin of 'Piura Porcelana' – A fine-flavored traditional variety of cacao (*Theobroma cacao*) from the Peruvian Amazon. *Tree Genetics & Genomes* 15:11
28. Wang J, Lin M, Crenshaw A, Hutchinson A, Hicks B, et al. 2009. High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. *BMC Genomics* 10:561
29. Peakall R, Smouse PE. 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–95
30. Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–39
31. Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure from multilocus genotype data. *Genetics* 155:945–59
32. Jakobsson M, Rosenberg NA. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–6
33. Rosenberg NA. 2004. DISTRICT: a program for the graphical display of population structure results. *Molecular Ecology Notes* 4:137–38
34. Zhang D, Motilal L. 2016. Origin, dispersal, and current global distribution of cacao genetic diversity. In *Cacao Diseases*, eds. Bailey BA, Meinhardt LW. Switzerland: Springer. pp. 3–31. https://doi.org/10.1007/978-3-319-24789-2_1
35. Yang J, Scascitelli M, Motilal LA, Sveinsson S, Engels JMM, et al. 2013. Complex origin of Trinitario-type *Theobroma cacao* (Malvaceae) revealed using plastid genomics. *Tree Genetics & Genomes* 9:829–40
36. Fabrice D, Jean-Jacques R, Sophie A, Raherinandrasana NL, Staub I, et al. 2017. Adaptation of fermentation process to local constraints: the case of cocoa in Sambirano valley, Madagascar. In: *Proceedings of the first International Symposium on Cocoa Research ISCR November 13–17 2017*. Lima, Péru (In French with English abstract). <https://agritrop.cirad.fr/586695/>
37. International Cocoa Organization. 2016. ICCO Panel recognizes 23 countries as fine and flavour cocoa exporters. <https://www.icco.org/icco-panel-recognizes-23-countries-as-fine-and-flavour-cocoa-exporters/>
38. World Bank. 2017. Excellence: a winning formula for Madagascar cocoa production. <https://www.worldbank.org/en/news/feature/2017/10/02/excellence-a-winning-formula-for-madagascar-cocoa-production>
39. International Cocoa Organization. 2019. Technical information dossier for the assessment of production and export of fine flavor cocoa – APRIL 2019 ADHOC PANEL MEETING. <https://www.icco.org/fine-or-flavor-cocoa/>



Copyright: © 2021 by the author(s). Exclusive Licensee Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.