BIOSCIENCE
JOURNAL

GENETIC VARIABILITY OF COCOA FROM THE AMAZON BASIN: IMPLICATIONS FOR GENETIC CONSERVATION

Maria Eduarda da Silva GUIMARÃES[1], Luiz Antônio dos Santos DIAS[1],
Caio Márcio Vasconcelos Cordeiro DE ALMEIDA[2], Carlos Alberto Spaggiari SOUZA[3],
Odimar Ferreira ALMEIDA[1], Thais Roseli CORRÊA[4]

[1] Departament of Agronomy, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.
[2] Comitê Executivo do Plano da Lavoura Cacaueira, Porto Velho, Rondônia, Brazil.
[3] Comitê Executivo do Plano da Lavoura Cacaueira, Linhares, Espírito Santo, Brasil.
[4] Departament of Agronomy and Plant Health, Universidade Estadual do Maranhão, São Luís, Maranhão, Brasil.

Corresponding author:
Maria Eduarda da Silva Guimarães
maria.eduarda.ufv@gmail.com

How to cite: GUIMARÃES, M.E.S., et al. Genetic variability of cocoa from the amazon basin: implications for genetic conservation. *Bioscience Journal*. 2023, **39**, e39020. https://doi.org/10.14393/BJ-v39n0a2023-63126

Abstract

The cocoa and chocolate production chain involves US$60 billion annually and three million farmers around the world, in an area exceeding nine million hectares. The use of wild germplasm will enable to generate new disease- and pest-resistant cultivars and ability to adapt to changing environments. Here we evaluated 145 cocoa accessions, originated from nine Amazonian basins, based on eight fruit traits. Univariate anova showed significant differences ($p<0.05$) for all traits. For seven traits, the variance component within basins was higher (81.5%, on average). Therefore, it is recommended that the collection of wild accessions prioritize a larger number of plants from a few populations of the most divergent basins. The multivariate analyses revealed a greater divergence between the Ji-Paraná-RO and Solimões/Amazonas-PA basins (27.69) and a greater similarity between Alien clones-PER and Solimões/Amazonas-AM (0.66) in relation to their populations. They also revealed that the accessions differentiation occurred according to the river basin system. These results allowed elucidate the genetic structure and distribution of cacao populations. In addition, strengthen the importance of collecting and conserving germplasm to preserve genetic resources.

Keywords: Amazon basin. Cacao. Genebank. Genetic diversity. *Theobroma cacao* L.

1. Introduction

Cocoa (*Theobroma cacao* L.), tree species belonging to the family Malvaceae, is a perennial plant, typical of tropical climate and native from America rain forest. Originated from Amazon basin and initially cultivated in Central America, more than 2000 years ago, cocoa is a crop of great socio-economic importance (Zarrillo et al. 2018). Its production chain involves US$60 billion annually (Rosenblum 2006) and three million farmers in the world, in an area exceeding nine million hectares. In Brazil, the cocoa and chocolate production chain generate US$4 billion annually. There are up to 90,000 cocoa farmers directly involved in a cropped area of about 600,000 hectares, in six States producers (IBGE 2019).

Traditionally, *T. cacao* is divided into three racial groups: Criollo (*T. cacao* var. *cacao*) and Forastero (*T. cacao* var. *sphaerocarpum*). Hybrids between Upper and Lower Amazon Forasteros and South American Criollos arose spontaneously in Trinidad and are Trinitarian racial groups (Dias et al. 2003; Bhattacharjee and Akoroda 2018). However, studies with molecular markers showed a new classification of the genepool of Amazonian cocoa, with ten new groups (Motamayor et al. 2008).

The Amazon region is the diversity center of the *T. cacao* and presents serious risks of genetic erosion by logging and mining, hydroelectric projects and the opening of new agricultural frontiers. The cocoa accessions present in germplasm collections represent only a small portion of existing variability. Since the 20th century, expeditions have been carried out, mostly in the Amazon basin, with the aim of collecting and preserving accessions (Souza et al. 2012).

The use of wild germplasm will enable to generate new cultivars with disease and pest resistance and ability to adapt to changing environments, to ensure the future of the cocoa economy sustainable. The collection of *T. cacao* genetic resources aims to the rescue its populations, which are a repository of genetic variability and offers opportunity to rescue new genes that may provide future genetic gains to the breeding programs (Zhang et al. 2012; Osorio-Guarín et al. 2017).However, the costs of this activity are high and there are difficulties to return to a particular region that have been exploited. In this way, the collector must explore the different alternatives and possibilities presented by collecting (Almeida 2001). To manage cocoa genetic resources efficiently, we must know about the variability, how it is distributed, and which factors influence its distribution. It is believed that the differentiation of *T. cacao* populations in the Amazon region occurs by river basins and this hypothesis has been proven by Dias et al. (2003). This information can be used to improve germplasm collection for genetic conservation and breeding programs.

The objectives of this study were: i) to assess the variability and its distribution pattern between and within populations of different cocoa Amazon river basins; ii) to evaluate the hypothesis of differentiation of cocoa populations by river basins.

2. Material and Methods

Cocoa germplasm collection and local

For this study we used the cocoa germplasm collection of the Experimental Station at Ouro Preto (Estex-OP) (latitude 10° 44' 30" S, longitude 62° 13' 30" W, and altitude 280 m above sea level), located in the county of Ouro Preto do Oeste, 340 km from Porto Velho, Rondônia, Brazil. Rondônia presents climate of type Aw-rainy tropical, with the coldest month temperature of 18 °C and a well-defined dry season.

The Estex-OP cocoa germplasm collection contains 615 clonal accessions from multiple provenances, predominantly composed by wild populations of Rondônia. For the present study, we selected 145 clonal accessions based on: physiological maturity (plants over ten years of field); presence of at least three plants per accessions at the same age of field; plants in good agronomic conditions for evaluation, considering size and canopy architecture; better control of production and better condition of competitiveness between plants of the plot, assessed as less number of dead plants in the neighborhood. The planting of the clones studied was held from 1977 to 1984, in plots of varying sizes from 5 to 36 plants, arranged in single, dual or contiguous rows. The 145 clones evaluated belong to nine river basins (Table 1). The number of plants evaluated by clonal accession ranged from 4 to 14 plants.

The evaluations were carried out during the 37 harvests from four crop consecutive years. We evaluated eight traits per plant: total number of healthy fruits (TNHF); seed wet weight of healthy fruits (SWWHF), in g; seed wet weight mean of healthy fruits (SWWMHF), in g; percentage of fruits with witches' broom (PFWB); percentage of bored fruits (PBF); percentage of fruits with germinated seeds (PFGS); number of branches with witches' broom (NBWB) and number of inflorescence with witches' broom (NIWB).

River basins

Aimed at collecting cocoa germplasm, Almeida et al. (1987) divided the Brazilian Amazonian region in 186 river basins, classified as follows: four of first order, 60 of second order and 122 of third-order. In the present study, we sampled nine basins (Table 1), being three of first-order (1. Delta/Estuário-PA, 3.

Solimões/Amazonas-PA and 7. Solimões/Amazonas-AM basins); three of the third order (2. Ji-Paraná-RO, 4. Jamari-RO and 8. Acre-AC) (Figure 1); and alien clones from Trinidad and Tobago (5. Alien clones_TTO), and from Peru_PER (6. Alien clones_PER) and clones from plantations of Rondônia-RO (9. Clones-RO), as outgroups.

Table 1. Identification of 145 cocoa clonal accessions from nine river basins, evaluated at Ouro Preto do Oeste, RO, Brazil.

Basins	Clones	Collection region	N° of evaluated clones
1	BE 9 and 10	Belém-PA	2
2	CAB 008 to 013	Presidente Médici-RO	44
	CAB 417 to 474	Ouro Preto do Oeste, Jaru, Mirante da Serra e Ji-Paraná-RO	
3	CAB 035 to 042	Santarém-PA	6
4	CAB 218 to 414	Ariquemes-RO	63
5	ICS 9 and 100	Farms from Trinidad and Tobago	2
6	IMC 67	Iquitos, Peru	4
	PA 150	Parinari, Peru	
	POUND 7 and 12	Rio Nanay, Loreto, Peru	
7	MA 11 to 15	Careiro islands, Manaus-AM	4
8	RB 39	Rio Branco-AC	1
9	SA 002 to 020	Plantations-RO	19
	SA 046 and 047	Plantations-RO	

Basins: 1. Delta/Estuário-PA; 2. Ji-Paraná-RO; 3. Solimões/Amazonas-PA; 4. Jamari-RO; 5. Alien clones_TTO; 6. Alien clones_PER; 7. Solimões/Amazonas-AM; 8. Acre-AC; 9. Clones-RO.

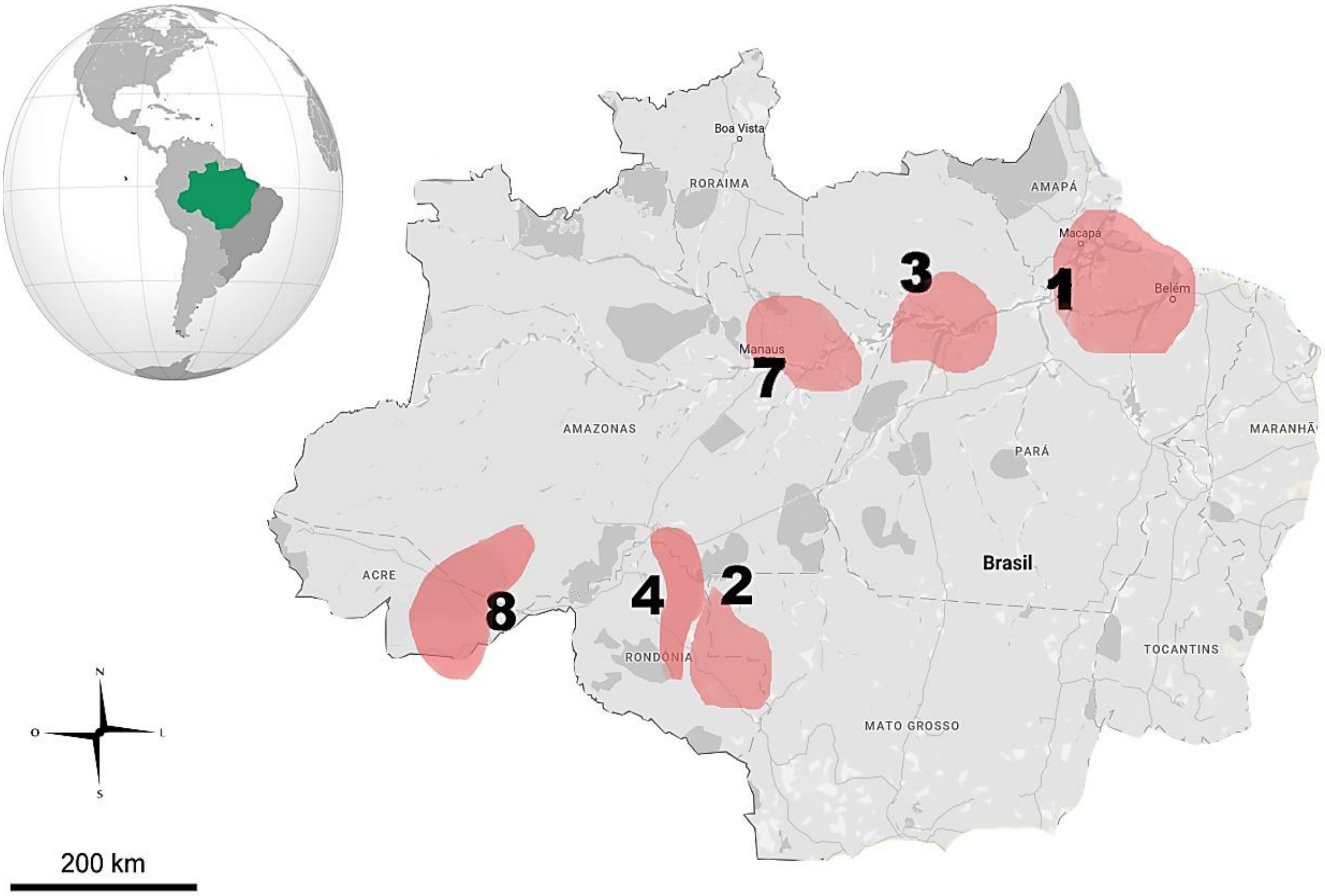

Figure 1. Map of the Brazilian Amazon region, with six of the nine river basins evaluated. 1. Delta/Estuário-PA, 2. Ji-Paraná-RO, 3. Solimões/Amazonas-PA, 4. Jamari-RO, 7. Solimões/ Amazonas-AM, 8. Acre-AC.

Statistical procedures

The data were summed by year and traits and then calculated the corresponding average values of the four years. Exploratory data analyses for normality test were performed. Six traits evaluated presented slight deviations from normality, but the analyses of variance (Anova) were processed from the original data. The data concerning the traits NBWB and NIWB, expressed as a percentage and represented by the variable x, presented large deviations from normality, and thus, were transformed using the following function Y:

$$Y = \sqrt{x+0.5}$$

Data were processed by univariate Anova, in hierarchical scheme (Oikawa and Iemma 1999; Dias and Barros 2009), for estimation of variance components in each of the three levels of population subdivision: among basin (B_i), among accessions within basins ($A_{j(i)}$) and between trees within accessions and basins ($T_{k(ij)}$). The random effects of B_i, $A_{j(i)}$, $T_{k(ij)}$ were assumed independent with averages 0 and variances equal to $\sigma_b^2, \sigma_a^2, \sigma_t^2$, respectively. The estimated total variance σ_T^2 was divided as follows:

$$\sigma_T^2 = \sigma_b^2 + \sigma_a^2 + \sigma_t^2$$

The following model was used

$$Y_{ijk} = \mu + B_i + A_{j(i)} + T_{k(ij)}$$

Where:
Y_{ijk}= average value of four years of trait Y measure on kth tree of jth accession of ith basin;
μ = general mean;
B_i = effect of ith basin, with i = 1, 2, ..., 9;
$A_{j(i)}$= effect of jth accession within ith basin, with j = 1, 2, ..., 145;
$T_{k(ij)}$= random sampling error defined by the effect of kth tree within jth accession of ith basin.

The Mahalanobis (1936)'s distance was also performed by means of the D^2 statistic to quantify the genetic diversity between cocoa trees of basins. The distance D^2 between two basins i and i' in *p* traits is defined as:

$$D^2 = d'W^{-1}d$$

Where d is a vector of difference between the means of basins i and i' for all the traits *p* and d' is its transposed, and W is the error covariance matrix *p* x *p* between trees within accessions and basins, obtained from Anova. The F test was applied to study the levels of significance of D^2 values. The univariate Anova and the Mahalanobis D^2 distance analyses were performed using the Nested and Candisc procedures, respectively, of the SAS (SAS Institute Inc. 1989).

3. Results

Genetic variability within and between cocoa populations from river basins

There were significant differences ($p<0.05$) for all traits evaluated, within and between cocoa populations from river basins. For seven of the eight traits, the variance component within basins [among accessions within basins ($A_{j(i)}$) and between trees within accessions and basins ($T_{k(ij)}$)] was higher, with exception for SWWMHF (Table 2).

Genetic variability between cocoa populations from different river basins

The values of Mahalanobis distances, performed for the differentiation between cocoa populations from river basins, ranged from 0.66 to 27.69 (Table 3). The cocoa populations of the Ji-Paraná-RO (2) and

Solimões/Amazonas-PA (3) basins were more divergent (27.69), whereas those of the Alien clones_PER (6) and Solimões/Amazonas-AM (7) basins were the most similar. When submitted to the F statistics, most of the D^2 distance values were significant (Table 3), in agreement with the results of the univariate Anova (Table 2).

Table 2. Summary of the hierarchical univariate analysis of variance and its components of variance (CV, in %) for each source of variation.

Traits	Sources of variation					
	Basins (B)		Accessions (A)/B		Trees (T)/A/B	
	MS	CV (%)	MS	CV (%)	MS	CV (%)
TNHF[1]	3878.22*	16.28	797.16*	34.68	164.10	49.03
SWWHF[2]	2820111.8*	1.62	3583698.8*	34.54	904634.0	63.83
SWWMHF[3]	57143.50*	77.00	966.60*	17.58	55.24	5.40
PFWB[4]	958.78*	13.56	275.21*	45.26	37.89	41.17
PBF[5]	356.30*	5.44	165.67*	37.12	37.30	57.43
PFGS[6]	44.54*	17.02	4.82*	14.27	2.22	68.69
NBWB[7]	7.43*	11.22	2.43*	50.48	0.28	38.29
NIWB[8]	3.12*	5.91	1.06*	48.24	0.16	45.84
Degrees of freedom	8		136		618	
Component of Variance (CV %)		18.50		35.30		46.20
Component of Variance (CV, %) within and between basins		18.50				81.50

[1] Total number of healthy fruits (TNHF); [2] Seed wet weight of healthy fruits (SWWHF), in g; [3] Seed wet weight mean of healthy fruits (SWWMHF), in g; [4] Percentage of fruits with witches' broom (PFWB); [5] Percentage of bored fruits (PBF); [6] Percentage of fruits with germinated seeds (PFGS); [7] Number of branches with witches' broom (NBWB) and [8] Number of inflorescences with witches' broom (NIWB). * Significant at 5% of probability by F test.

Table 3. Mahalanobis distance matrix (D^2, above of diagonal) between cocoa populations originated from nine river basins and significance from F test (p values, below of diagonal).

Basins	1	2	3	4	5	6	7	8	9
1	0	17.65	4.47	12.01	5.19	0.90	0.83	11.53	2.62
2	0.0001	0	27.69	1.38	26.12	18.25	17.06	7.50	15. 46
3	0.0001	0.0001	0	22.52	1.61	5.72	3.76	17.00	5.99
4	0.0001	0.0001	0.0001	0	21.91	12.71	11.72	4.83	10.24
5	0.0001	0.0001	0.2113	0.0001	0	6.93	5.02	18.65	5.42
6	0.5955	0.0001	0.0001	0.0001	0.0001	0	0.66	13.03	2.07
7	0.7502	0.0001	0.0001	0.0001	0.0001	0.4681	0	10.04	1.54
8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0	10.54
9	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0

1. Delta/Estuário-PA; 2. Ji-Paraná-RO; 3. Solimões/Amazonas-PA; 4. Jamari-RO; 5. Alien clones_TTO; 6. Alien clones_PER; 7. Solimões/Amazonas-AM; 8. Acre-AC; 9. Alien clones_RO. Underlined D^2 values were not significant by F test.

4. Discussion

In general, partitioning of variation within and between populations revealed that most (on average, 81.50%) of the variability was detected within basins (populations). These results corroborate with those obtained in other studies involving morphological traits (Dias et al. 2003) and DNA molecular markers (Thomas et al. 2012). Genetic differentiation between and within cocoa populations of river basins, on a given spatial scale, depends on natural selection, genetic drift, and the distance and magnitude of the gene flow (Dias et al. 2003). *T. cacao* presents an exclusive mechanism of incompatibility that favors cross fertilization. During microsporogenesis, the trees present gametophytic and sporophytic factors that control the phenotype of pollen incompatibility (Lanaud et al. 2017). This incompatibility system may limit self-fertilization, restrict the gene flow and yield of self-incompatible trees. Within a single cocoa population, some trees may exhibit self-compatibility for varying situations (Lanaud et al. 2017).

The cocoa populations from the Ji-Paraná-RO (2) and Solimões/Amazonas-PA (3) basins were the most divergent in relation to the others. It is believed that wild cocoa populations from the Jamari-RO (4) and Ji-Paraná-RO basins are genetically related and belong to the same genepool (Almeida 2001). However, Mota (2003) found divergence between the populations of that two basins due to recent processes of human intervention, through cropping and selection.

The distinction between the populations of the studied river basins seems to confirm that the differentiation of wild Amazonian cocoa populations evolved according to each system of river basins, as verified by Dias et al. (2003). Such differentiation may be related to the geographic distance between populations (Dias et al., 2003). The basins of Ji-Paraná-RO and Solimões/Amazonas-PA, for example, are distant of more than 800 km. Geographic isolation prevented gene flow and drove the greatest genetic distance between the cocoa populations of these basins.

Another explanation for the differentiation between cocoa populations from different basins may be the action of past evolutionary processes, as reported in studies with molecular markers (Mota 2003; Motamayor et al. 2008; Thomas et al. 2012). In the present work we considered the combination of three evolutionary hypotheses: river barriers (Patton et al. 1994), refuges (Haffer 1969) and Amazonian lake (Klammer 1984). This combination was also used by Dias et al. (2003) to discuss and explain the variability among the cocoa populations of four basins of the Brazilian Amazon. However, these authors evaluated data from 64 progenies, being 16 in each basin in just one year. In the present study, a greater number of accessions of nine basins were evaluated during four crop consecutive years. However, the number of accessions per basin ranged from 1 to 63.

According to the hypothesis of refuges, climatic changes imposed by glaciation during the Pleistocene and Post-Pleistocene, resulted in the fragmentation of the Amazon rainforest in dry seasons, forming forest refuges in more humid regions of the basin. During the wet seasons, these fragments were reconnected. The process of retreat and expansion was repeated several times during the Quaternary, leading to a rapid differentiation of Amazonian forest populations. Isolation probably limited gene flow among cocoa populations, which evolved allopathically under natural selection and genetic drift, as suggested by other authors (Dias et al. 2003; Mota 2003). In some extensions, the basins evaluated correspond to areas of refuge, initially proposed by Haffer (1969) and later by Prance (1973). The basins of Delta/Estuário-PA (1), Ji-Paraná-RO (2), Solimões/Amazonas-AM (7) and the Peruvian (6) basins corresponding to the refuges of Xingu, Aripuanã, Manaus and East Peru, respectively.

The hypothesis of river barriers was suggested by Wallace (1854), in his study with primates from the Amazon region. According to Wallace, the barriers formed by the great rivers of the Amazon forest influenced the development of local biodiversity, later corroborated by Patton et al. (1994). The populations of the Amazonian forest, which were continuously distributed, were isolated geographically when the network of the great Amazonian rivers formed in the late Tertiary and early Quaternary. These barriers may have blocked the gene flow among populations and led to speciation of some plants and animals on opposite banks of rivers (Nazareno et al. 2017; Pirani et al. 2019). Consequently, the species evolved by natural selection and genetic drift. The discontinuous distribution of small populations along a watercourse seems to reinforce the role of rivers in the allopatric divergence of natural Amazonian populations (Almeida et al. 1987; Dias et al. 2003; Mota 2003). This pattern of diversification has often been associated with refuges theory (Haffer 1969; Dias et al. 2003; Mota 2003). It is believed that gene flow is more restricted in large rivers and their headwaters than in small rivers and their headwaters (Dias et al., 2003; Nazareno et al. 2017). Thus, the divergence between basins is lesser than among populations into basins. The fact that in most cases the cocoa populations of the basins were genetically distant, as revealed by the D^2 statistic, shows that the gene flow between them may have been reduced.

According to the Amazonian lagoon hypothesis, in the last four million years, in the Plio-Pleistocene, sea level rise has become a barrier to the discharge of the water system of the Amazon River. This process resulted in the formation of the Amazonian lake that covered most of the Amazon, at the end of the Tertiary (Pliocene). It is presumed that smaller portions of the forest were covered during a series of sea level alternations during the Quaternary. The flood has pushed the rainforests to peripheral regions of the Amazon Basin, the present boundaries of the uplands. The differentiation that occurred at the edges of the basin increased the isolation effect of the broader rivers, separating animal and plant populations into “forest islands”, as evidenced by the corresponding terrain levels in the Lower Amazonas valley (Klammer 1984; Marroig and Cerqueira 1997).

According to Mota (2003), the occurrence of cocoa populations private alleles in the Solimões/Amazonas and Acre basins shows greater diversity in the more peripheral regions of the Amazon. This result reinforces the hypothesis that the differentiation of the cocoa populations results from the isolation imposed by the river barriers or by the supposed Amazonian lake. These hypotheses could explain the greater diversity in the periphery (highlands) than in the central Amazon region (Bush 1994). This suggests greater gene flow converging from the edges to the center of the Amazon Basin (Marroig and Cerqueira 1997).

5. Conclusions

This study provides new insights for the genetic structure and distribution of cocoa populations aiming collection, management, conservation and breeding. Germplasm collection should consider the spatial distribution of genetic variation. Since most of the variation was concentrated between trees within accessions and basins, to improve the collection process, a greater number of trees should be prioritized in a few populations of the most divergent basins. River basins appear to be a factor of variation in the structuring of wild cocoa populations.

Authors' Contributions: GUIMARÃES, M.E.S.: Conception and design, acquisition of data, analysis and interpretation of data, drafting the manuscript, final approval; DIAS, L.A.S.: Analysis and interpretation of data, drafting the manuscript, final approval; ALMEIDA, C.M.V.C.: Acquisition of data, analysis and interpretation of data, drafting the manuscript, final approval; SOUZA, C.A.S.: Acquisition of data, analysis and interpretation of data; ALMEIDA, O.F.: Acquisition of data, analysis and interpretation of data; CORRÊA, T.R.: Acquisition of data, analysis and interpretation of data. All authors have read and approved the final version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Ethics Approval: Not applicable.

Acknowledgments: The authors would like to thank the funding for the realization of this study provided by the Brazilian agencies CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -Brasil), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico -Brasil) and FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais).

References

ALMEIDA, C.M.V.C., et al. Evolução do programa de conservação dos recursos genéticos de cacau na Amazônia brasileira. CEPLAC/DEPEA, Belém, Pará: Boletim Técnico 5, 1987.

ALMEIDA, C.M.V.C. Ecology of natural populations. 2001. In: L.A.S. DIAS eds. Genetic improvement of cacao, Rome: FAO Ecoport version.

BHATTACHARJEE, R. and AKORODA, M. Taxonomy and classification of cacao. 2018. In: P. UMAHARAN eds. *Achieving sustainable cultivation of cocoa*. Cambridge: Burleigh Dodds Science Publishing, pp. 3-18.

BUSH, M.B. Amazonian speciation: a necessarily complex model. *Journal of Biogeography*. 1994, **21**(1), 5-17. https://doi.org/10.2307/2845600

DIAS, L.A.S., et al. Variation and its distribution in wild cacao populations from the Brazilian Amazon. *Brazilian Archives of Biology and Technology.* 2003, **46**(4), 507-514. https://doi.org/10.1590/S1516-89132003000400003

DIAS, L.A.S. and BARROS, W.S. Biometria experimental. Suprema, Viçosa, Minas Gerais, 2009.

HAFFER, J. Speciation in Amazonian forest birds. *Science*. 1969, **165**(3889), 131-137. https://doi.org/10.1126/science.165.3889.131

IBGE. Produção agrícola municipal: culturas temporárias e permanentes. 2019. Available from:https://biblioteca.ibge.gov.br/visualizacao/periodicos/66/pam_2019_v46_br_informativo.pdf

KLAMMER, G. The relief of the extra-Andean Amazon basin. 1984. In: H. SIOLI eds. The Amazon: Limnology and landscape ecology of a mighty tropical river and its basin, Dordrecht: Dr. W. Junk Publishers, pp. 47-83.

LANAUD, C., et al. Deciphering the Theobroma cacao self-incompatibility system: from genomics to diagnostic markers for self-compatibility. *Journal of Experimental Botany*. 2017, **68**(17), 4775–4790. https://doi.org/10.1093/jxb/erx293

MAHALANOBIS, P.C. On the generalized distance in statistics. *Proceedings from National Institute of Science of India*. 1936, **12**(1), 49-55.

MARROIG, G., CERQUEIRA, R. Plio-Pleistocene South America history and the Amazonas Lagoon hypothesis: a piece in the puzzle of Amazonian diversification. *Journal of Computational Biology*. 1997, **2**(2), 103-119.

MOTA, J.W.S. 2003. Análise da diversidade genética de germoplasma de Theobroma cacao L. da Amazônia brasileira por microssatélites**.** Tese de Doutorado, Universidade Federal de Viçosa, Viçosa, Minas Gerais, 97p.

MOTAMAYOR, J.C., et al. Geographic and genetic population differentiation of the Amazonian chocolate tree. *PLoS One*. 2008, **3**(10), 1-8. https://doi.org/10.1371/journal.pone.0003311

NAZARENO, A. G., DICK, C. W., LOHMANN, L. G. Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. *Molecular Ecology*. 2017, **26**(14), 3636–3648. https://doi.org/10.1111/mec.14142

OIKAWA, S.M., IEMMA, A.F. Hipóteses estatísticas com dados desbalanceados nos modelos com três fatores de efeitos fixos hierarquizados. *Scientia Agricola*. 1999, **56**(4), 1123-1134. https://doi.org/10.1590/S0103-90161999000500014

OSORIO-GUARÍN, J. A., et al. Colombia a source of cacao genetic diversity as revealed by the population structure analysis of germplasm bank of Theobroma cacao L. *Frontiers in Plant Science.* 2017, **8**(1), 1-13. https://doi.org/10.3389/fpls.2017.01994

PATTON, J.L., SILVA, M.N.F. and MALCOLM, J.R. Gene genealogy and differentiation among arboreal spiny rats (*Rodontia: Echimydae*) of the Amazon basin: a test of the riverine barrier hypothesis. *Evolution*, 1994, **48**(4), 1314-1323. https://doi.org/10.1111/j.1558-5646.1994.tb05315.x

PIRANI, R. M., et al. Testing main Amazonian rivers as barriers across time and space within widespread taxa. *Journal of Biogeography*. 2019, **46**(11), 2444–2456. https://doi.org/10.1111/jbi.13676

PRANCE, G.T. Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon basin, based on evidence from distribution patterns in Caryocaraceae, Chrysobalanaceae, Dichapetalaceaeand Lecythidaceae. *Acta Amazonica*. 1973, **3**(3), 5-26. https://doi.org/10.1590/1809-43921973033005

ROSENBLUM, M. Chocolate: uma saga agridoce preta e branca. Rio de Janeiro: Rocco, 2006.

SAS Institute Inc. SAS/STAT user's guide, version 6. 4th edition, volume 2, SAS Institute, Cary, USA, 1989.

SOUZA, A.G.C., et al. Contribution of the institutions in the Northern region of Brazil to the development of plant cultivars and their impact on agriculture. *Crop Breeding and Applied Biotechnology*. 2012, **12**(1), 47–56. https://doi.org/10.1590/S1984-70332012000500006

THOMAS, E., et al. Present spatial diversity patterns of *Theobroma cacao* L. in the neotropics reflect genetic differentiation in Pleistocene refugia followed by human-influenced dispersal. *PLoS One*. 2012, **7**(10), 47-67. https://doi.org/10.1371/journal.pone.0047676

WALLACE, A.R. On the monkeys of the Amazon. *Annals and Magazine of Natural History.* 1854, **14**(84), 451-454. https://doi.org/10.1080/037454809494374

ZARRILLO, S., et al. The use and domestication of Theobroma cacao during the mid-Holocene in the upper Amazon. *Nature Ecology & Evolution.* 2018, **2**(12), 1879-1888. https://doi.org/10.1038/s41559-018-0697-x

ZHANG, D., et al. Genetic diversity and spatial structure in a new distinct Theobroma cacao L. population in Bolivia. *Genetic Resources and Crop Evolution*. 2012, **59**(2), 239–252. https://doi.org/10.1007/s10722-011-9680-y

Received: 6 September 2021 | **Accepted:** 12 April 2022 | **Published:** 03 February 2023