

Full Length Research Paper

Molecular Ecology of genetic diversity of cacao cultivated in the south-east region of Nicaragua

JC Ruiz¹, O Roa Gamboa¹ and I Marin Arguello^{1*}

¹Laboratorio BIOciencia, UNAN-Managua. Recinto Rubén Darío, Managua, Nicaragua

Accepted 25 February, 2011

The use of local germplasm of *Theobroma cacao* L. as a basis for developing a breeding program is limited by the lack of information on genetic variability and relationships among populations. This study covered a sample of 70 cocoa trees in five geographical areas of Nicaragua: El Rama, Nueva Guinea, Muelle de los Bueyes, El Castillo and Los Guatusos. The samples were molecularly analyzed using 10 microsatellite (SSR) DNA. For the 10 locus was achieved amplification of 95 alleles. Populations showed an average observed heterozygosity (H_o) of 0.522, and Nueva Guinea was the population with the value of H_o closest in value of expected heterozygosity (H_e). The latter obtained the lowest inbreeding rate ($f = 0.06$). The value obtained for the average Fixation Index F_{ST} was 0.23, which allows you to catalog the cocoa plantations of southeastern of Nicaragua with a moderate level of genetic diversity. The populations of Muelle de los Bueyes and El Castillo presented the maximum value of genetic distance (0.1784). According to the Principal Coordinate Analysis, the population distribution in a geometric plane can be grouped into two blocks. Finally, the study phylogenetic populations of Castillo and Guatusos are more closely related to each other, than the rest of the population.

Keywords: *Theobroma cacao* L., Microsatellite markers, Germplasm, Consanguinity, Heterocygosis

INTRODUCTION

When the Spanish conquistadors arrived in the New World, indigenous ancestors were using cacao as currency. The units of change were a Hand, five seeds, Quince, fifteen seeds or Tzontly forty cocoa beans (Rocha, 1976, Palma, 1975). Since then, in Mesoamerica, the cultivation of cocoa has a cultural and economic roots (Henderson, 2007). Cocoa, *Theobroma cacao* ($2n = 2x = 20$) whose genome is 390 Mega bases (Mb) distributed in 10 chromosomes, is a neotropical species found in the Mesoamerican region between 20° N and 20° S, considered the area of domestication of cocoa (Henderson, 2007). Commercial cultivars of T. cocoa have a narrow genetic base (Schnell, 2005). The

primary types of cocoa are Forastero, Trinitario (hybrid between Forastero and Criollo) and Criollo, but Motamayor (2008) proposes to classify the species into 10 clusters or groups: Marañón, Curaray, Criollo, Iquitos, Nanay, Contamana, Amelonado, Purus, Nacional and Guiana. De la Cruz, (1995) presents divided *Theobroma* into two sub-species: Cocoa and *sphaerocarpon*. Beside, De la Cruz (1995) suggests that the Criollo and Trinitario types should be considered as traditional crops and not like genetic groups. In addition, the Amazon basin is the region that has greater cacao diversity (Gentry, 1988).

In Nicaragua the cocoa plantations were developed hybrid genetic material derived from crosses between Trinitario and Forastero, and in addition endogenous Criollo cocoa plantations, although it is unknown the status of biodiversity of the latter. The producers, through

*Corresponding author email: ivmarinarguello@gmail.com

social exchange of seeds, are responsible for the spread of cacao trees, which were not derived from clonal gardens or breeding programs. In Nicaragua the annual production of cacao is 1,500 t -2.000 6.500 ha, with an approximate yield of 3.5 quintals mz-1 and priced at U\$ 2-6 per kg (Buchert, 2008). According to the International Cocoa Organization (2007) Nicaragua is ranked 27 among cocoa-exporting countries and their production represents 0.026% of the global total.

One of the main concerns of the world cocoa market is the ability of producing countries to supply the market continuously with fine cocoa quality, value added, as high oil content of cocoa, polyphenol groups (antioxidants) and terpenes (flavor and aroma) (Othman, 2007; Niemenak, 2006). In this regard it is important to understand the genetic diversity of cacao plantations in Nicaragua due to its impact on production sustainability. Therefore, the objective of this study was to establish the degree of genetic variation among populations of cacao grown in Nicaragua and to determine the genetic diversity of cacao and extrapolate cocoa elite populations of the Mesoamerican region.

MATERIALS AND METHODS

Plant material

Between February and June 2010 took place sampling plant material in Nueva Guinea, El Rama, Muelle de los Bueyes, El Castillo and Los Guatusos. From different trees, 70 samples were collected of immature leaves and not young. The tree selection was based on its productive potential and physical condition. The collection of each sample was accompanied with the raising of ecological agriculture tab of the plot where the tree stood. Moreover, each tree was referenced satellite and labeling. Each leaf was wrapped in paper towels and packed in a Ziploc bag and preserved at -20 ° C until processing.

DNA extraction

Prior to DNA extraction leaf was rinsed several times with distilled water and drilled to obtain 200 mg leaf discs were placed in a 1.5 mL Eppendorf tube which was immersed in liquid nitrogen and crushed using a mortar Kontes. Then we used a commercial medium for DNA extraction, DNeasy Plant Mini Kit (Qiagen). We proceeded with the protocol recommended by the commercial company with the variant that AP1 extraction solution was added 10 mg mL⁻¹ of Polyvinylpyrrolidone (Sigma). The DNA was quantified using Biospectofotómetro (Shimadzu UV-1200) and all DNA samples were diluted to a concentration of about 5 ng mL⁻¹ (Saunders et al., 2004; Zhang et al., 2006).

Microsatellite markers

For molecular characterization was implemented 10 microsatellite (SSR) reported by Saunders (2004) and considered the comments of Zhang (2006). The name of microsatellite markers, the sequence of the primers, temperature, alignment and size are reported by Saunders (2004). The PCR reaction was performed in a volume of 50 µL, containing 5 ng of genomic DNA. All the PCR reactions contained: 1.0 U Platinum Taq DNA polymerase (Invitrogen), 0.2 mM dNTPs (Invitrogen), 0.2µM of each primer (Invitrogen), 1.5 mM MgCl₂, 1X buffer (20 mM Tris-HCl, pH 8.0, 40 mM NaCl, 2 mM sodium phosphate, 0.1 mM EDTA, 1 mM DTT, 50% glycerol). The conditions of the thermocycler (Eppendorf) were: 4 min denaturation at 94 °C, followed by 33 cycles of denaturation at 94 °C for 30 s, 1 min at a temperature of alignment of 46 °C or 51 °C and 1 min extension at 72 °C, with a final extension of 72 °C for 7 min.

Electrophoresis

Amplified products were separated on a vertical electrophoresis apparatus (T-Rex Aluminum Backed Sequencer, Model FB-SEQ-3545-FISHER) in a denaturing gel (6 M urea) 6% polyacrylamide (Fisher) running at 55W, 1700V by 2 or 3 h in 1X TBE buffer (Qiagen) and detected by silver staining. The size of the amplified bands was determined with a commercial DNA marker, 125bp DNA Ladder (Invitrogen).

Statistical data processing

The statistical program Genetic Analysis in Excel v6.1 (GenAIEx) (Peakall, 2006) was used to determine the average number of alleles per locus, allele frequency, the percentage of polymorphism loci, the value of observed and expected heterozygosity, the pattern of alleles in the population, and the Principal Coordinates Analysis (PCA). The statistical program Genetix 4.02 (Belkhir, 2001) was implemented for statistical determination of Nei genetic diversity and the calculation of the F-statistics (F-Statistic of Wright). The phylogenetic tree was constructed using Phylogeny Inference Package (PhyIip v3.6) program and based on molecular data. Other support programs used were Arlequin v3.1 and Convert (Glaubitz 2004).

RESULTS

With the implementation of 10 microsatellite (SSR) was determined the variation at the level of allele. This was amplified 95 alleles with an average of 9.5 alleles per locus. In our study, Y16988 locus has the highest earning

Table 1. General statistics of the 10 microsatellites used in cocoa samples

Locus	SSR [†]	N _a ^{††}	Alleles (bp) [‡]	H _o [§]	H _e [¶]
Y16980	mTcCIR6	10	230, 235, 240, 250, 275, 280, 285, 290, 295, 300	0.504	0.609
Y16981	mTcCIR7	9	155, 160, 165, 170, 173, 175, 180, 185, 190	0.684	0.653
Y16982	mTcCIR8	8	275, 280, 285, 290, 295, 300, 305, 310	0.611	0.507
Y16985	mTcCIR11	11	210, 280, 285, 290, 295, 300, 303, 305, 310, 320, 470	0.399	0.620
Y16988	mTcCIR15	15	220, 225, 230, 235, 240, 245, 250, 255, 260, 270, 275, 280, 290, 300, 310	0.514	0.702
Y16991	mTcCIR18	9	305, 310, 315, 320, 325, 330, 340, 345, 350	0.512	0.624
Y16991	mTcCIR22	7	270, 275, 280, 285, 290, 295, 300	0.395	0.419
Y16998	mTcCIR26	7	280, 285, 290, 295, 300, 305, 310	0.492	0.498
AJ271943	mTcCIR40	9	250, 255, 260, 265, 270, 275, 280, 285, 290	0.631	0.670
AJ271958	mTcCIR60	10	180, 185, 190, 195, 200, 205, 210, 215, 220, 225	0.476	0.670
Total / Average		95		0.522	0.597

† microsatellite representation; † † alleles number; ‡ base pair; § observed heterozygosity; ¶ expected heterozygosity.

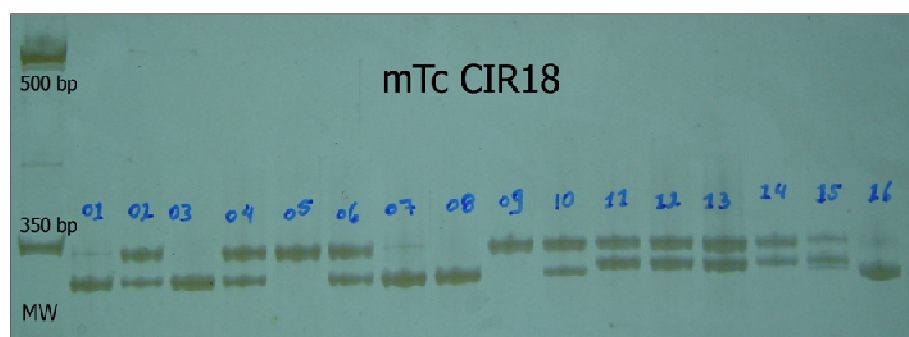


Figure 1. Silver staining 6% polyacrylamide gel of amplified molecular marker (mTcCIR18). MW: molecular weight of DNA marker (Invitrogen, 125 bp DNA Ladder). The numbers represent the samples of cacao analyzed

a total of 15 alleles (Table 1). This locus has a better chance of change in the frequency of any of its alleles.

The Figure 1 shows the homozygous or heterozygous condition for a particular locus identified using molecular markers. Viewing a single band determines the

homozygous condition; otherwise two bands mark the state of heterozygosity.

An important parameter in genetic diversity of a population is the relationship between values for observed heterozygosity (H_o) and expected

Table 2. Parameters of genetic diversity of the populations under study

Population	Parameters genetic diversity									
	N [¶]	K ^{¶¶}	H _o	H _e	f [§]	H _s [†]	H _T ^{††}	D _m [‡]	D _{ST} [♯]	G _{ST} ^{♯♯}
El Castillo	23	6.3	0.609	0.733	0.19240					
Los Guatusos	10	4.3	0.548	0.619	0.16758					
Nueva Guinea	11	4.3	0.555	0.562	0.06082					
El Rama	17	4.1	0.448	0.553	0.21885					
Muelle de los Bueyes	9	3.6	0.450	0.520	0.19245					
Average 4.5	0.522	0.600	0.16642	0.5974	0.7781	0.2258	0.1807	0.2322		

¶ alleles number; ¶¶ average of alleles by locus; § Coefficient of Consanguinity; † Genetic Diversity inside of population; †† Total Genetic Diversity ; ‡ Absolute Genetic Differentiation; ♯ Genetic Diversity Between Populations ; ♯♯ Relative Magnitude of Gene Differentiation among Subgroups.

Table 3. Estimation of Fixation Index (F-Wright) ^P for the 10 loci studied

Locus	SSR	F _{IT} [†]	F _{ST} [‡]	F _{IS} [§]
Y16980	mTcCIR6	0.36338	0.23398	0.16892
Y16981	mTcCIR7	0.39264	0.24157	0.19919
Y16982	mTcCIR8	0.39278	0.23493	0.20633
Y16985	mTcCIR11	0.35616	0.23794	0.15514
Y16988	mTcCIR15	0.36169	0.24308	0.15671
Y16991	mTcCIR18	0.35986	0.23041	0.16821
Y16991	mTcCIR22	0.36392	0.22175	0.18268
Y16998	mTcCIR26	0.37466	0.23081	0.18702
AJ271943	mTcCIR40	0.38277	0.24788	0.17934
AJ271958	mTcCIR60	0.37068	0.24939	0.16159
Average		0.37234	0.23708	0.17770

♯ calculated with a confidence interval 95% Bootstraps based on 1000 replications per locus; † Total Fixation Index; ‡ Fixation Index Among Populations; § Fixation Index Within Populations

heterozygosity (H_e). Heterozygosity is determined by the number of times (frequency) repeated the different alleles for a particular locus. Table 1 show that in 8 of the 10 loci analyzed, the value of H_o is less than the value recorded for H_e.

The values for the Fixation Index also known as the Coefficient of Inbreeding (*f*) is showed in table 2. This parameter determines the reduction of heterozygosity or outcrossing level. Values can range from -1 to +1. The study did not yield any negative value for *f* indicating that the populations under study there is no excess of heterozygosity.

The distribution of variability within and between the populations studied is determined by the Total Genetic Diversity, Total Heterozygosity (H_T), Heterozygosity for each Population (H_s) and Genetic Diversity among

populations (D_{ST}). In this study the value of the H_T is the result of the sum of (H_s) and D_{ST}. The value reported for D_{ST} (0.1807, Table 2) indicates that loci for the five populations studied has few heterozygotes and several types of homozygotes.

The Coefficient of Genetic differentiation among populations (G_{ST}) is calculated by the relationship D_{ST} / H_T, in the present study, the G_{ST} value of 0.2322 (Table 2) and F_{ST} of 0.23708 (Table 3), very close together, indicating that diversity among populations is supported by the presence of two alleles at a locus.

The parameter that determines the absolute magnitude of Genetic Divergence (D_m) expresses the variability between interpopulations with respect to the intrapopulations. This ratio is calculated by ND_{ST} / n-1, where n represents the number of populations.

Table 4. Genetic Distance (Nei, 1972) among the five populations[§] under study south-eastern of Nicaragua

	CA	GU	NG	RA	MB
CA (23)	0.000				
GU (10)	0.507	0.000			
NG (11)	1.285	1.864	0.000		
RA (17)	1.674	1.633	0.180	0.000	
MB (9)	1.784	1.386	0.360	0.110	0.000

§ El Castillo, CA ; Los Guatusos, GU; Nueva Guinea, NG ;El Rama, RA; Muelle de los Bueyes, MB

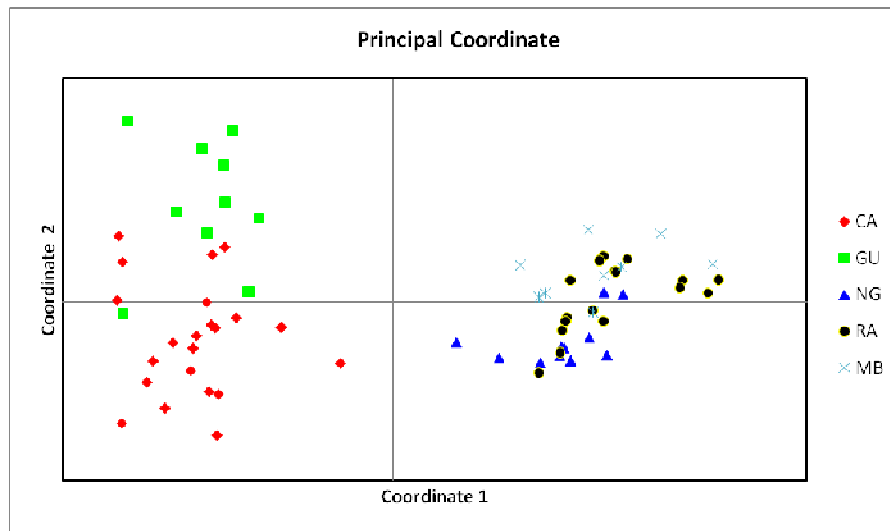


Figure 2. Dimensional array of association (distance) between individuals from different populations (El Castillo, CA; Los Guatusos, GU; Nueva Guinea, NG; El Rama, RA; Muelle de los Bueyes, MB).

The probability that two alleles of the same locus in an individual, are identical from the same offspring is denoted as the Fixation Index or Coefficient of Differentiation (F) and in the context of the present study, this index corresponds to the values F (Table 3), where the subscripts representing the level of hierarchy, which are being compared.

It is instructive to determine the proximity or the Genetic Distance between populations, to achieve this we determined the Coefficient of Genetic Distance (Table 4). This parameter allows compared pairs of populations of the five populations under study.

The Principal Coordinates Analysis (PCA) can transform a number of variables correlated with a small number of uncorrelated variables, facilitating visual interpretation (graph) of the similarity or dissimilarity that

exists between individuals or populations. This relationship is embodied in a plane where the geometric distance between points reflects the genetic distance. The representation does not take into account the evolutionary processes, there is a phylogenetic tree and the relationship between two points is determined by calculating the frequency of alleles. In the Figure 2 shows the distribution in a geometric plane of each sampled tree, represented by a point.

If a pair of genes from different species or populations is developed in parallel and whether the degree of divergence between the two genes / populations implies that they have originated from the generation t , we can infer that these genes diverged from an ancestor common at time period $t / 2$. This reasoning allows us to construct the phylogenetic tree based on genetic

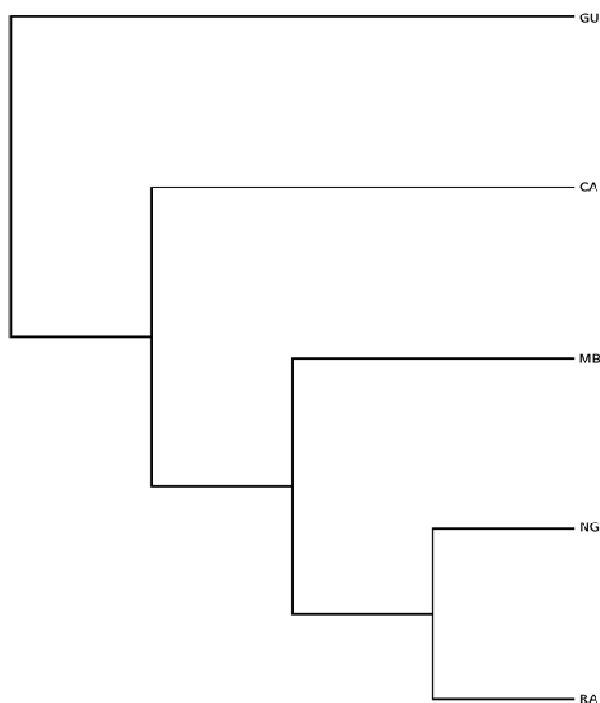


Figure 3. Dendrogram for the phylogenetic tree of the five geographical cacao populations included in this study. Molecular phylogenetics tree is aimed at relationships for populations based on molecular data (allele and haplotype frequencies)

distance. Figure 3 shows the phylogenetic tree representing the relationships of five populations under study, based on the degree of similarity or number of shared characteristics.

DISCUSSION

The genetic variation between individuals or populations may manifest at different levels. The first level comprises a set of indicators that measures the change internally in the population (frequency of alleles, polymorphism, observed and expected heterozygosity, mean number of alleles per locus, effective number of alleles per locus, fixation index, etc.) and the second level includes another set of indicators to measure the variation of a population relative to other populations (genetic distance, inbreeding coefficient, geometric correlation between genetic distance, cluster analysis, phylogeny, etc.). These indicators were identified in this study

In Brazil's cacao plantation Sereno (2006) used 11 microsatellite and report the amplification of 49 alleles

with an average of 4.45. A higher average (5.6) reported by Zhang (2009) a study developed in the international collection of cocoa held in CATIE, Botanical Garden (Costa Rica). This study reports an average of 9.5 alleles per locus. Value obtained due to the size of the area sampled. Populations of trees sampled in this study are scattered and are not under the care of a germplasm collection.

This study reported a deficiency of heterozygosity. It is very likely that the sampling would have developed among related trees (consanguinity). This means that a percentage of the sampled trees, have arisen from a facultative selfing or almost total, what makes the union of gametes not random. A second explanation could be the result of the Wahlund effect, which occurs when the samples have different allele frequencies from two mixed populations.

In relation to the divergence between H_o and H_e values (Table 2) for population El Castillo and El Rama, they are far apart, it is likely that one or more considerations for the Hardy-Weinberg (HW) are not being met (being diploid, sexually reproducing, not this cross generational,

gene with two alleles, the allele frequency is identical male and female gametes, random crossover, mutation and migration is ignored and, finally, that natural selection does not affect the alleles). The relationship between H_o and H_e makes the link between alleles and genotypes frequency, also determines the proportion of genotypes in a population remain constant from generation to generation.

Obtain positive values for Coefficient of Consanguinity (f) 0.06 to 0.21 denote different levels of heterozygote deficiency, probably due to the effect of consanguinity. This coefficient (f) can also be considered as decreasing the proportion of heterozygosity at the crossing between related species. Inbreeding tends to expose harmful and recessive alleles, a situation that may be affecting the productivity of cocoa farms in the study area. This phenomenon is referred as biparental inbreeding that occurs when genetic information of the progeny comes from two different parents but related to each other, this could be explain the level of abortions in cacao tree, inducing by incompatibility between gametes (Hartl, 2007).

The G_{ST} value 0.2322 indicates that most of the total variation is distributed among populations and the variation or diversity is not exclusive to any of the populations studied. Beside a value of D_m 0.225 suggest that variability present in the populations is supported by the diversity of the individuals in the population and not by the variations among populations.

Thus, F_{ST} is the probability that any particular locus in an inbred population have alleles identical by descent. The reported value for F_{ST} equal to 0.23 indicates that the populations studied have a moderate genetic diversity. F_{IS} index is a critical parameter in measuring genetic variation. This parameter has a minimum value of 0 (indicating absence of genetic divergence) and the maximum value of 1 (indicating fixation for alternate alleles in different subpopulations). Also set heterozygosity deficiency in individuals within the population.

Based on data obtained from Table 2 and 3, it shows that the cacao trees are showing an overlap of generations and fertilization preferably between related plants that grow close together and the place occupied by the mother plant. This situation reduces the genetic diversity of cacao plantations. However, the trend of loss of heterozygosity can be reversed by providing, in just a one generation the random crossing between no related trees.

Thus Muelle de los Bueyes (MB) and El Castillo (CA) are the populations with the highest distance value 0.1784, being reported as the most diverse populations analyzed. Conversely, the closest are the Rama (RA) and Muelle de los Bueyes with a distance value 0.110, providing a lower limit differentiation. This value determines the proportion of the genetic elements (alleles, genes, gametes, genotypes) that share the two

populations compared. A value equal to zero indicates that population are absolutely identical.

Regarding to the PCA analysis, the points that cross from one quadrant to another indicate genetic similarity between individuals (Figure 1). Also seen in certain quadrants as there are points that fit perfectly one on another as if they were a single point. Observe how population tend fall into two main sections. One block represented by the populations of Los Guatusos and El Castillo, and a second block consisting of El Rama, Muelle de los Bueyes and Nueva Guinea. Within this second group, the population of El Rama, has a heterogeneous profile of individuals closely linked to populations of Muelle de los Bueyes and Nueva Guinea.

The phylogenetic tree permit grouped populations in relation who submit the smallest genetic distance. Each branch represents the most recent common ancestor. In our study, less connected populations or "neighbors" from the point of view common ancestor belong to Los Guatusos and El Rama.

CONCLUSION

In conclusions, the present study generated information on the genetic diversity of cacao grown in the five villages in Nicaragua and is able to generate information regarding the degree of relatedness between populations, genetic distance between populations, establish the level of genetic diversity and heterozygosity. Based on determined parameters we can conclude that the genetic diversity of cacao in the study area is moderate. All this information would provide the baseline for establishing a breeding and conservation program of local biodiversity of cocoa. Breeding programs focused at the selection of cocoa tree with a particular property (flavor and aroma).

ACKNOWLEDGEMENTS

Our appreciation go to the producers and field technicians who actively participated in the taking of samples. Thank categorically to IICA-Nicaragua and the Austrian Development Cooperation-ProDeSoc by the financial support provided in the study. To all of them our gratitude.

REFERENCE

- Azofeifa-Delgado A (2006). Uso de marcadores moleculares en plantas; Aplicaciones en frutales del trópico. *Rev. Agron. Mesoamericana* 17(2): 221-242.
- Bartley BG (2005). *The Genetic Diversity of Cacao and its Utilization*. CABI Publishing, Massachusetts. 341 p.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001). GENETIX: logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France. CD.

- Buchet JC (2009). Tendencias mundiales del cacao fino. II Foro El Cacao Fino o de Aroma en la Estrategia Cacaotera de Nicaragua. Managua. 30 p.
- Comparative study of different cocoa (*Theobroma cacao* L.) clones in terms of their phenolics and anthocyanins contents. *J. Food Comp. Analysis* 19: 612-619.
- De la Cruz M, Whitkus R, Gomez-Pompa A, Mota-Bravo L (1995). Origins of cacao cultivation. *Nature* 375:542-543.
- Gentry A (1988). Tree species richness of upper Amazonian forest. *PNAS* 85:156-159.
- Gillespie, J. 2004. Population Genetics. A Concise Guide. 2nd Edition. The Johns Hopkins Press. London. 214 p.
- Glaubitz JC (2004). CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol. Ecol. Notes* 4: 309-310.
- Hartl D, Clark A (2007). Principles of Population Genetics. 4th Edition. Sinauer Associates, Inc. Publishers. Massachusetts. 652 p.
- Henderson J, Joyce R, Hall G, Hurst W, McGovern P (2007). Chemical and archaeological evidence for the earliest cacao beverages. *PNAS* 104 (48): 18937-18940.
- Motamayor JC, Risterucci A, López P, Ortiz C, Moreno A, Lanaud C (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89:380-386.
- Motamayor JC, Lachenaud P, Silva e Mota J, Loor R, Kuhn D, Brown S, Schenell R (2008). Geographic and genetic population differentiation of the amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE* 3(10): e3311.
- Nei M (1973). Analysis of gen diversity in subdivided populations. *PNAS* 70: 3321-3323.
- Niemenak N, Rohsius C, Elwers S, Omokolo-Ndoumou D, Lieberei R (2006).
- Nuez F, Carillo JM (2000). Los Marcadores Genéticos en la Mejora Vegetal. Editorial de la Universidad Politécnica de Valencia. España. 579 p.
- Organización Internacional del Cacao (2007). ICCO Informe anual. Londres. 52 p.
- Othman A, Ismail A, Ghani N, Adenan I (2007). Antioxidant capacity and phenolic content of cocoa beans. *Food Chemistry*, 100: 1523-1530.
- Palma I (1975). Moneda y bancos en Nicaragua. Imprenta Nacional, Managua. 84pp.
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Eco. Notes* 6: 288-295.
- Pritchard JK, Stephens M, Gonnely P (2000). Inference of population structure using multilocus genotypes data. *Genetics* 155:945-959.
- Rocha G (1976). Breve ensayo sobre la numismática de Nicaragua. En Cuadernos Universitarios, Segunda Serie, No. 18. p. 93-117.
- Saunders J, Mischke S, Leamy E, Hemeida A (2004). Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. *Theor. Appl. Genet.* 110:41-47.
- Schell RJ, Olano C, Brown J, Meerow A, Cervantes-Martínez C (2005). Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedling and association of microsatellite alleles with productivity. *J. Amer. Soc. Hort. Sci.* 130(2):181-190.
- Sereno M, Albuquerque P, Vencovsky R, Figueira A (2006). Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conser. Genet.* 7:13-24.
- Zaid A, Hughes H, Porceddu E, Nicholas F (2001). Glossary of Biotechnology for Food and Agriculture. FAO Research and Technology Paper 9. Rome. 305 p.
- Zhang D, Boccara M, Motilal L, Butler D, Umaharan P, Mische S, Meinhardt L (2008). Microsatellite variation and population structure in the Refractario cacao of Ecuador. *Conser. Genetic.* 9 (2): 327-337.
- Zhang D, Mischke S, Goenaga R, Hemeida A, Saunders J (2006). Accuracy and reliability of high-throughput microsatellite genotyping for cacao clone identification. *Crop. Sci.* 46:2084-2092.
- Zhang D, Mischke S, Johnson E, Phillips-Mora W, Meinhardt L (2009). Molecular characterization of an international cacao collection using microsatellite markers. *Tree Genet. Genomes* 5:1-10.