Genetic variability in a collection of the tropical forage legume *Centrosema molle Mart. ex Benth.* and its association with the climate where they originate in Venezuela

Genetic variability in a collection of the tropical forage legume *Centrosema molle Mart. ex Benth.* and its association with the climate where they originate in Venezuela

Sandy Molina ,[1] Orlando Guenni,[2] Diógenes Infante [1]*

[1] Unidad de Biotecnología de Plantas, Instituto de Estudios Avanzados (IDEA), Caracas, Venezuela.

[2] Universidad Central de Venezuela. Facultad de Agronomía. Instituto de Botánica Agrícola, Maracay 2101, Venezuela.

Centrosema molle Mart. ex Benth. (*sin. C. pubescens Benth.*) is among the most important tropical legumes used for forage and green cover. With the purpose of studying the genetic variability of a germplasm collection of the species, 34 accessions were evaluated using four AFLP primer combinations. Additionally, the association between the genetic groups derived from the molecular characterization and those produced by eco-geographical origin was analysed based on climate classification. The resulting climatic groups were separated mainly in three altitude ranges: 10-180 m.a.s.l., 90-540 m.a.s.l. and 320-936 m.a.s.l. The four primer combinations produced 164 fragments, of which 57% were polymorphic. No duplicate accessions were present. An analysis of principal components showed the presence of three homogenous groups associated with the eco-geography (in particular the altitude) of the collection site. This may suggest an incipient process of differentiation in the *C. molle* natural populations due to the geographical isolation produced by some particular habitats.

Keywords
AFLP, genetic diversity, genetic resources, natural populations.

Introduction

The Neotropical savannahs have been the source of a number of forage legumes, and Venezuela in particular has a high diversity of native legumes (Flores & Schultze-Kraft, 1994). Among these, there are species of the *Centrosema genus*, which cover 34 species native to the tropical and sub-tropical Americas (Williams & Clements, 1990). In Venezuela, 12 species have been reported (Fariñas, Shultze-Kraft, Calles, Guenni, & Rodríguez, 2006). Commercial cultivars have been released from *C. acutifolium, C. pascuorum and C. molle* (Schultze-Kraft Clement, & Keller-Grein, 1990), but only *C. molle* has had a major agronomic and economic impact outside the region (Humphreys, Ivory, Wong, & Topark-Ngarn, 1990; Teitzel, Cameron, Anning, & Stockweell, 1990).

C. molle is widely distributed from Mexico to southeast Brazil, also extending towards Bolivia and Paraguay (Schultze-Kraft *et al.*, 1990). Currently more than 1,000 accessions have been preserved in various germplasm banks (Schultze-Kraft *et al.*, 1990). *C. molle* enjoys broad phenotypical variability (Schultze-Kraft & Keller, 1985; Grof, Flores, Mendoza & Pizarro, 1990; Keller, Schultze-Kraft, Franco & Ramírez, 2000) and is adapted to environments with low rainfall or extended drought periods (Faría-Marmol, Chirinos, Faría, & González, 1998). The germplasm bank of forage legumes of the National Institute of Agriculture Investigations of Venezuela curates a collection of 74 accessions, including material previously evaluated by the International Tropical Pastures Evaluation Network, RIEPT, (Keller-Grein *et al.*, 2000), and a set of 50 recently collected from a wide range of ecological zones (Guenni *et al.*, 2007).

Molecular markers can be used to identify genes of

agronomic importance and are particularly suited to decision-making in germplasm collections as they simplify the identification of genetic duplicates. In order to assess genetic diversity of a highly valuable forage legume *ex situ* collection, this study sets out to characterize the genetic variability of 34 *C. molle* accessions using AFLP as a platform for DNA fingerprinting. At the same time, this allowed the evaluation for possible relationships between genotype and the eco-geography of the collection site.

Materials and Methods

The germplasm

Thirty-four accessions from the most variable eco-geographic points, within the natural range of distribution in the country, were chosen initially because of the availability of the seed from the FLGBV-INIA (Table 1). Depending on the seed availability, these 34 *C. molle* accessions were cultivated during the wet season (May-September) in pots under greenhouse conditions, or in small plots (2 x 2 m2) at the experimental station of the Agronomy Faculty, Universidad Central de Venezuela (UCV), Maracay (10o15´N, 67o39´W; elevation 450 m.a.s.l.). In both cases, plants were grown on a sandy loam soil and were fertilized and irrigated accordingly. Weed control was performed manually.

Geographic and climatic clustering of the collection

Geographic coordinates and the altitude of the collecting site were used to discrimate several accession groups from the analysis of principal components and conglomerates through the “Ward’s” method, using the FloraMAP package (Jones, Guarino, & Harvis, 2002). The different groups were defined according to the altitude and climatic environment of the site of origin (total annual rainfall, rainfall distribution and air temperature changes along the year).

DNA extraction and AFLP fingerprinting

DNA was extracted from *C. molle* leaves following the methodology proposed by Keb-Llanes, González, Chi-Manzanero & Infante, (2002). Three individuals per accession were sampled. The AFLP procedure was a scaled down version of the technique described by Vos *et al.*, (1995). From the DNA extracted from each leaf sample (two per accession), 250ng were double digested to completion with 5U *Msel* and 5U *EcoR1*. Then, 50pmol *EcoR1* adaptor and 5pmol *Msel* adaptor were ligated by adding 1U DNA T4 ligase and incubating at 4°C overnight. This restriction/ligation solution was diluted 1/10 with double distilled sterile water and 2ul used as a pre-amplification template, using pre-selective primers Eco R1 (+A) and Mse 1 (+C). The 20ul pre-amplification PCR contained 2mM MgCl2, 30pmol of each primer, 0.5U Taq polymerase and 0.2mM dNTP, and the cycling regime was 5min at 94°C, followed by 24 cycles of 94°C / 30s, 56°C / 60s and 72°C / 60s, with a final extension step of 72°C / 7min. The pre-amplification reaction was diluted 1/50 and 5ul used as a template for the 20ul selective PCR (2mM MgCl2, 6.7ng MseI+3 primer and 32.8ng EcoRI+3 primer, 0.5U Taq polymerase and 0.2mM dNTP). The cycling regime consisted of 94°C / 5min, followed by 12 cycles of 94°C / 30s, 65°C (reducing by 0.7°C per cycle) / 30s and 72 °C / 60s; followed by 24 cycles of 94°C / 30s, 56°C / 30s, 72°C / 60s, and ending with 72°C / 7min. The samples were concentrated by evaporation under vacuum until reaching a volume of 4ul, and then, 16ul loading buffer was added. The sample was denatured at 96°C for 5min and electrophoresed through 6% denaturing polyacrylamide gels, which were silver stained with Silver Sequence stain kit (Promega Corporation, Madison, WI, USA), following the manufacturer’s instructions. Four AFLP primer combinations were used: E+AAC M+CAA, E+AAC M+CAC, E+ ACG M+CAG, E+AAC M+CAG.

Statistical analysis

Bands were scored as 0 (absent) and 1 (present), and a binary matrix (0 and 1) was built. The data were exported to NTSYSpc (Exeter Software, Seutaken, NY, USA) and Dice coefficients of similarity for two state data (presence/absence) were calculated. Clustering was obtained using UPGMA analysis (Unweighted pair-group method, arithmetic average), and a dendrogram generated using the derived genetic distances. The topology of the cluster was verified by bootstrapping the data matrix 1,000 times with 25% resampling using the program Freetree (Hampl, Pavlicek, & Flegr, 2001).

Accession No. (FLGBV/CIAT)*	State	Latitude (N)	Longitude (W)	Elevation (m.a.s.l)	Climatic group	Genetic group
431 (CIAT 15133)	Trujillo	9.37	-70.70	900	1	1
135	Monagas	9.95	-63.97	633	1	2
79	Aragua	10.48	-67.72	351	1	2
434 (CIAT 15150)	Barinas	7.67	-71.37	260	1	2
428 (CIAT 5634)	Portuguesa	9.37	-69.98	700	1	2
427 (CIAT 5631)	Trujillo	9.40	-70.48	490	1	2
131	Monagas	9.72	-63.82	162	1	3
51	Aragua	9.88	-66.95	480	1	2
422 (CIAT 5169)	Aragua	10.07	-66.93	420	1	1
108	Yaracuy	10.17	-68.72	408	2	2
110	Carabobo	10.18	-68.18	663	1	3
83	Aragua	10.27	-67.48	522	2	3
101	Aragua	10.42	-67.77	45	3	3
432 (CIAT 15144)	Zulia	8.72	-72.53	10	3	3
439 (CIAT 15872)	Miranda	10.18	-66.45	320	1	3
440 (CIAT 15875)	Miranda	10.22	-66.25	60	3	3
69	Bolívar	7.22	-65.52	84	3	3
63	Guárico	7,72	-66.25	57	3	2
59	Guárico	8.08	-67.05	30	1	3
117	Anzoátegui	8.27	-64.32	66	3	2
436 (CIAT 15160)	Barinas	8.33	-69.55	180	3	3
28	Cojedes	8.98	-68.18	90	2	2
38	Cojedes	9.05	-68.50	90	2	3
52	Aragua	9.80	-66.83	267	2	3
106	Carabobo	10.48	-68.2	15	1	3
87	Cojedes	9.82	-68.52	459	2	3
42	Cojedes	9.72	-68.55	189	1	3

Table I. State, latitude, longitude, and altitude of the collection sites. The accessions are grouped according to the eco-geographic and AFLP molecular characterization.

* Accessions originally from the Forage Legumes Germplasm Bank of Venezuela (FLGBV) and from the International Center for Tropical Agriculture (CIAT)

Results and Discussion

Climate of origin

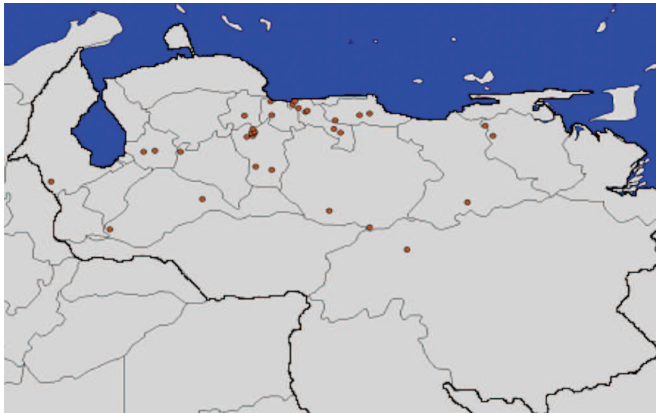

FIGURA 1. Geographical distribution in northern Venezuela of the 34 *C. molle* accessions used in this study. The map was generated by Flora-MAP. Dots represent the locations of the collection sites.

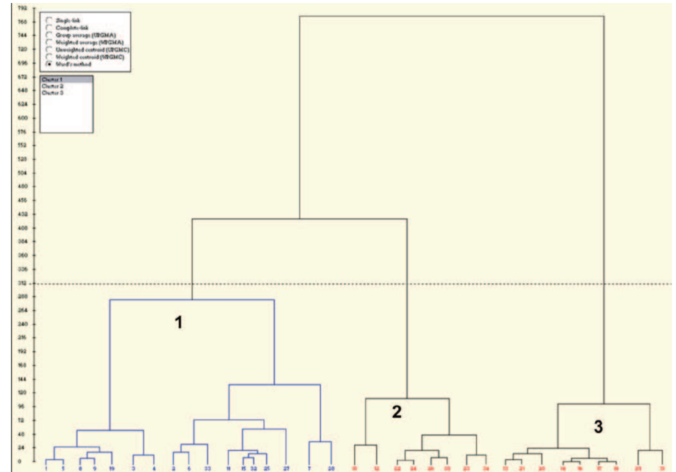

FIGURA 2. Dendrogram resulting after the application of the "Ward's" method to generate the different climatic groups (1, 2 and 3) represented in Table II. The climatic clusters are originated on the basis of latitude and longitude of the collecting site.

Figure 1 represents the geographic distribution of the collecting sites in Venezuela. The sites of origin covered a total of 12 states and included a large variation in habitats, ranging from savannahs to forests, savannah-forest ecotones and disturbed areas such as grazed paddocks and roadsides. Overall, the altitude of the collecting sites varied between 10 m.a.s.l. (FLGBV 432, Zulia state) and 936 m.a.s.l. (FLGBV 73, Aragua state) (Table 1). In Figure 2, a dendrogram generated by the "Ward's" method is presented, with the three resulting groups. Table 1 includes the accessions within each climatic group, with differences among them based mainly on altitude. The resulting climatic groups are described as follows:

Group 1: comprises 17 accessions (Table 1), of which ten originate from altitudes between 320 m.a.s.l and 936 m.a.s.l., five between 160 m.a.s.l and 288 m.a.s.l., and two (FLGBV 106 and 59) from 15 m.a.s.l and 30 m.a.s.l.. The climate has a unimodal rainfall pattern, characterized by a drought period of three months with a rainfall range of 24-40mm/month, for a total of 96mm, and a rainy period with a monthly rainfall range of 56-212mm, resulting on a total annual rainfall of 1459mm. The minimum and maximum annual range of temperature is 20.5-22.5°C and 31-33°C, respectively. The annual mean temperature range is 26.1-27.8°C.

Group 2: consists of eight accessions (Table 1), occupying a lower altitudinal range (90-540 m.a.s.l.) than in Group 1. Only two accessions had a collection site under 100 m.a.s.l. (FLGBV 28 and 38), with the remainder collected above 260 m.a.s.l. The rainfall pattern is unimodal and the climate is characterized by two dry months with a rainfall of 29-34mm/month. The rainy period has a 71-76mm/month, with a total of 1198mm per year. The minimum and maximum annual temperature range is 17-18.5°C and 28-30°C, respectively, with an annual mean of 22.6-24.5°C.

Group 3: clusters nine accessions (Table 1), seven of which originated from low altitude sites (10-180 m.a.s.l.). FLGBV 90 and 94 were collected at 318 m.a.s.l. and 399 m.a.s.l., respectively. The rainfall pattern is unimodal, it has a dry period of four months with a rainfall range of 4-38mm/month. The rainy period averages 75-223mm/month, accounting a total of 1281mm during the year. The minimum and maximum annual range of temperature is 15.5°C-19°C and 29°C-34°C, respectively, for an annual mean of 23.9-26.0°C.

AFLP profiles

The four primer combinations produced a total of 164 fragments sized between 80 and 800 base pairs (bp). The most informative primer combination was E+AAC M+CAA (Figure 3), generating 43 out of the 93 polymorphic products. The number of fragments per individual and primer combination varied between 25 and 54, thus confirming the capacity of AFLP to produce genetic information for *C. molle*. The least similarity coefficient (0.83) was observed between accessions FLGBV 422 and 101, whereas the highest (0.96) was between FLGBV 87 and 21. The clustering assessment (Figure 4) produced three homogeneous groups. The first group contained FLGBV 431 and 422; the second FLGBV 135, 79, 51, 434, 428, 427, 28, 63, 108 and 117; and the third FLGBV 131, 101, 439, 87, 21, 42, 94, 45, 90, 85, 80, 59, 38, 440, 69, 436,

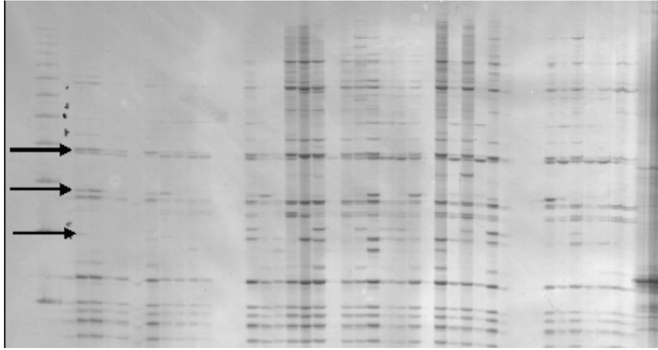

FIGURA 3. Partial E+ACC/M+CAA AFLP profile of the germplasm set on 6% polyacrilamide gel. Arrows indicate different polymorphic bands.

52, 106, 73, 83, 110 and 432. The topology of the above three groups was verified by bootstrapping, obtaining a dendrogram similar to that of Figure 4 (not shown), thus validating the existence of the different molecular groups.

A comparison between genotype and the eco-geography of the collection site is presented in Table 1. The molecular groups partially resemble those resulting from the eco-geographic classification. For example, the whole genetic group 1 and 60% of the accessions of group 2 are included in the climatic group 1. A large number of individuals of this group is comprised in accessions mostly from high altitudes. On the other hand, in the genetic group 3, 59% of the accessions are included in the climatic groups 2 and 3, which comprise materials from intermediate and low altitudes sites, respectively.

The genetic characterization demonstrated that the current *C. molle* collection may be clustered in three groups with a minimum similarity coefficient of 0.77. To confirm this actual level of genetic variability within the species, further assessment on other germplasm collections, as well as on native populations, should be made with these AFLP markers. The results of this work also showed that there was no evidence of duplicates in the collection, between accessions and even within accessions, because the AFLP profiles were identical between the two samples from the same plant used to calculate the dendrograms, but were not identical among the three plants sampled per accession to measure the similarity coefficient within accessions. In this case the coefficient was bigger and varied between 0.94 and 0.96. This is consistent with the fact that genetic identity does not exist in sexual or asexual reproduction (Infante, Molina, Demey, & Gámez, 2003;2006) and even throughout *in vitro* micropropagation (González, Alemán, & Infante, 2003; Sanchez-Teyer, Quiroz, Loyola, & Infante, 2003), when it is measured with high output molecular marker

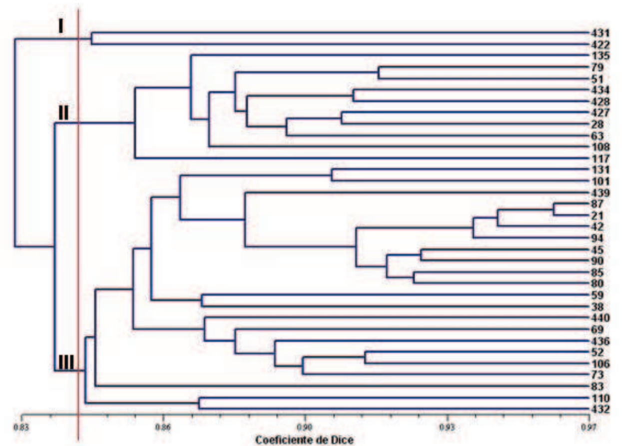

FIGURA 4. UPGMA dendrogram based on 93 variables bootstrapped 1,000 times with 25% resamplings. Numbers I, II and III refer to the genetic clusters shown in Table 1.

like AFLP or Inverse sequenced-tagged repeat (ISTR). So the AFLP technique represented a useful means of individual identification in the collection.

A genotypic characterization of a *Desmodium ovalifolium* collection also showed little variation (Fisher, Berndl, & Schultze-Kraft, 2004) compared to the relative high variation found among just eight accessions of an *Arachis pintoi* collection (Maas, Torres, & Ocampo, 1993). The analysis of the geographic distribution of the accessions showed the existence of three climatic patterns that form a continuous sequence along the whole gradient of altitudes covered by the species natural ditribution. An increase in total annual rainfall and a progressive decrease in both, temperature and the extension of the dry period with increasing altitude, were expected from climatic group 3 to 1. However, the inclusion of some particular accessions coming from altitudes well beyond the range found for the majority of those included in each group, certainly modified the mean value of the climatic variables in each climatic group generated by FloraMAP.

The existence of climatic groups differentiated by the altitude together with the results of the genetic analysis may indicate an incipient process of differentiation in the *C. molle* natural populations adapted to higher altitudes, probably due to the geographical isolation produced by those particular habitats. This is in agreement with the results obtained in other legume species with forage value (Maass, Jamnadass, Hanson, & Pengelly, 2005), whereas in other cases there has been a low association with the place of origin (Heide, Fischer, Berndl, & Schultze-Kraft, 2007). Alternatively, the genetic groups found in this study could be the result of the influence of other equally important environmental factors such

as the physical and chemical characteristics of the local soil. Further studies are required to confirm this hypothesis.

Most *Centrosema* species (including *C. molle*) have been assumed to be allogamic, so for the purpose of seed bank management, they have been handled as self-pollinated species. However, it is clear that the rate of outcrossing varies from 15% to 56% among some of the agronomically important *Centrosema* species, with a 30% - 56% range for *C. molle* (Miles, Clements, Grof, & Serpa, 1990; Escobar, 1991; Pentead, Garcia, & De la Vega,, 1996). Based on the level of heterozygosity and genetic variation within accessions, Penteado *et al.*, (1996) postulated that *C. molle*, as well as *C. acutifolium* and *C. brasilianum*, should be considered species in which outcrossing occurs at a high rate. These authors concluded that the presence of both fertilization systems should be taken into account for collecting, multiplying and conservation purposes. Consequently, given the relatively high level of outcrossing reported in these species, a significant genetic flow within accessions in a given *ex situ* collection may be expected, leading to possible losses of heterozygosity and allelic richness among them. However, in our study, there was not enough time for this to take place since the plants used for the genetic characterization came from the first batch of seed multiplied in the field and glasshouse after collecting from the natural populations. Nonetheless, the lower level of AFLP variation between samples within accessions suggested low heterozygosity/allele richness. We have to conclude that either there is very little outcrossing in the wild or the populations are very isolated from one another and represent homogeneous populations descended from a small number of founders. Another possibility is that the markers are not neutral and there have been extensive natural selection in the wild. This last possibility seems less probable due to the random nature of the used markers. However, this possibility can be discarded using a different marker system, as ISTR, a random generated marker system based in retrotransposons sequences (Demey, Gamez, Molina, & Infante, 2004).

Likewise, from the context of genetic variability conservation within the species, it seems important to invest more effort in preserving germplasm coming from more remote zones and in maintaining these materials relatively isolated in the future *ex situ* collections. With the morpho-agronomic characterization of this collection to be carried on in the near future we may be able to produce more conclusive information about the degree of genetic variation among accessions and the significance that this variability has in terms of their agronomic potential.

We conclude that the molecular analysis has demonstrated the existence of genetic variability in the collection of *C. molle*, which may have a potential for the selection of useful biological characters for a better adaptation to savannah ecosystems. Also, the potential correlation established between genetic and eco-geographic information represents a useful tool for handling collections and developing programs oriented to the conservation, propagation and agronomic improvement of *Centrosema* species.

Acknowledgements

This study was partially supported by FUNDACITE-Aragua, through funds from the Fondo Nacional de Ciencia y Tecnología (FONACIT) of Venezuela, under the scope of the research project No. 2004-FCT-01-15-9-2 (1749). Our gratitude to Teodardo Calles and Julio Tam, who kindly collaborated in processing the information generated by FloraMAP.

References

1. Demey, JR.; Gamez, E.; Molina, S. & Infante, D. (2004). Comparative study of the discriminating capacity of AFLP and ISTR markers for genetic analysis of *Agave fourcroydes*. Plant Molecular Biology Reporter. 22:29-35

2. Escobar, G. (1991). Aspectos coevolutivos de la polinización de plantas por insectos: polinización de *Centrocema macrocarpum,* un caso práctico. PhD Thesis. Universidad del Valle.

3. Faría-Marmol, J.; Chirinos, Z.; Faría, J. & González, I. (1998). Evaluación de accesiones de *Centrosema pubecens* en regiones secas del Estado Zulia. Archivo Latinoamericano de Producción Animal. 6:79-92.

4. Fariñas, J.; Shultze-Kraft, R.; Calles, T.; Guenni, O. & Rodríguez, I. (2006). Revisión de *Centrosema* (DC) Benth. (*Fabaceae*) para Venezuela. In: Mejía, M (ed.) IX Congreso Latinoamericano de Botánica. Santo Domingo, República Dominicana: p. 691.

5. Fisher, E.; Berndl, T. & Schultze-Kraft, R. (2004). Identification of core collection in *Desmondium ovafolium* based on markers data. In Kurt, J.; Peters; Kirschke, D.; Manig, W.; Bürkert, A.; Schultze-Kraft, R.; Bharati, L.; Bonte-Friedheim, C.; Deininger, A.; Bhandari, N. & Weitkamp, H. (eds.) Rural Poverty Reduction through Research for Development and Transformation: international research on food security, natural resource management and rural development. Druckhaus Köthen GmbH, Köthen. p. 293.

6. Flores, A. & Schultze-Kraft, R. (1994). Recolección de recursos genéticos de leguminosas tropicales en Venezuela. Agronomía Tropical. 44:357-371.

7. González, G.; Alemán, S. & Infante, D. (2003). Asexual Genetic Variability in *Agave fourcroydes* II: Selection among individuals in a clonally propagated population. Plant Science. 165:595-601.

8. Grof, B.; Flores, A.; Mendoza, P. & Pizarro, E. (1990). Regional experience with *Centrosema*: Northern South America. In: Schultze-Kraft, R. & Clements, RJ. (eds.) *Centrosema:* Biology, Agronomy, and Utilization. Colombia: Centro Internacional de Agricultura Tropical (CIAT). p. 391-420.

9. Guenni, O.; Calles, T.; Gil, J.; Fariñas, J.; Espinoza, F.; Sanabria, D. & Schultze-Kraft, R. (2007). Surveying and collecting native *Centrosema*, *Stylosanthes* and *Desmodium* germplasm in Venezuela. Plant Genetic Resources Newsletter. 148: 38-43.

10. Villiers, E.M.; Fauquet, C.; Broker, T.R.; Bernard, H.U. & Hausen H.Z. (2004). Classification of papillomaviruses. Virology, 324:17-27.

11. Woese, C.R. (1998). The Universal Ancestor, Proceedings of the National Academy of Sciences of the United States of America. 95:6854-6859.

12. Zmasek, C.M. & Eddy S.R. (2001). ATV: display and manipulation of annotated phylogenetic trees, Bioinformatics. 17: 383-384. **13**. Cox, J.L., Farrell, R.A., Hart, R.W., and Langham, M.E. The transparency of the mammalian cornea. J. Physiol. 210, 601, 1970.

(*) Author for correspondence:
Diógenes Infante,
Centro Nacional de Biotecnología Agrícola,
Instituto de Estudios Avanzados IDEA,
Baruta, Miranda,
República Bolivariana de Venezuela
Fax: 58-212-903-5093
e-mail: dinfante@idea.gob.ve