Molecular genetic evaluation of the *Stylosanthes capitata* Vog. and *Stylosanthes macrocephala* Ferr. et Costa germoplasms

Guilherme de Toledo-Silva[1], Melissa Oliveira Santos-Garcia[2], Rosângela Maria Simeão Resende[3], Lucimara Chiari[4], Cláudio Takao Karia[5], Marcelo Ayres Carvalho[6], Fábio Gelape Faleiro[7], Maria Imaculada Zucchi[8], Anete Pereira de Souza[9]

Introduction

The genus *Stylosanthes* Sw. is a member of the family Fabaceae, tribe Aeschynomeneae, subtribe Stylosanthinae and includes 48 herbaceous-subshrub species of neotropical origin (COSTA, 2006). The major diversity center is central Brazil (STACE and CAMERON, 1984), which contains 45% of all species of the genus and is the site where the greatest phenotypic variation and endemism are observed (COSTA, 2006).

Several *Stylosanthes* species are important sources of pasture legumes for tropical and subtropical environments, and they are also used for soil improvement through nitrogen fixation and for the reclaiming of degraded wastelands. *Stylosanthes macrocephala* is a diploid species (2n=20) that belongs to the Styposanthes Section (MANETJE, 1984). *S. capitata* Vog. (2n=40) likely originated from hybridization of the diploid species *S. pilosa* Ferr. et Costa and *S. macrocephala* Ferr. et Costa, which are both from the Styposanthes section (LIU et al., 1999).

Here, we used polymorphic microsatellite markers to evaluate the genetic diversity and population structure present in the germplasm collections of *S. macrocephala* and *S. capitata*. We also used the microsatellite data to assemble core collections for both species in order to facilitate the use and maintenance of these resources.

Material and methods

A total of 134 accessions of *S. macrocephala* and 192 accessions of *S. capitata* from Embrapa-Cerrados were evaluated using microsatellites developed by Santos et al. (2009a, 13SSR loci) and Santos et al. (2009b, 15 SSR loci), respectively.

DNA samples were extracted from fresh leaves using the cetyltrimethyl ammonium bromide (CTAB) extraction method as described by Faleiro et al. (2003). The PCR-amplified DNA fragments were separated by electrophoresis on 6% denaturing polyacrylamide gels and were then silver stained (Creste et al., 2001).

Genetic Data Analysis (GDA) (LEWIS and ZAYKIN, 2000) was used to estimate the observed and expected heterozygosities. Allele frequencies and Roger's distance modified by Wright (1978) were calculated between all pairs of accessions using Tools for Population Genetic Analysis (TFPGA) (MILLER, 1997).

The software STRUCTURE 2.0 (PRITCHARD et al., 2000) was utilized to infer population structure and assign accessions to populations based on the SSR genotypes. Identification of the number of distinct clusters. DARwin 5.0 software (PERRIER and JACQUEMOUD-COLLET 2006) was used to define the genetic relationships among accessions based on Roger's genetic distance and the Neighbor-Joining clustering method. FSTAT (GOUDET, 2001) was used to calculate Nei's G_{ST} among the groups formed by the STRUCTURE analysis.

Finally, we assembled a core collection that should represent the entire genetic diversity explored in this study using the software CoreFinder, available at http://www.appliedgenomics.org (CIPRIANI et al., 2010).

Results

Embrapa germplasm collections of *S. macrocephala* and *S. capitata* were genotyped by use of microsatellite markers. The number of alleles per locus ranged from two to 11 and averaged 4.7 alleles per locus in *S. macrocephala*. Individual markers detected from 0.02 to 0.85 of expected heterozygosity, with an average of 0.36. The observed heterozygosity was low, ranging from 0.01 to 0.17 and averaging 0.08 across all loci. For *S. capitata*, the allele number per locus ranged from two to nine and averaged 3.4 alleles per locus. The expected heterozygosity

1.Genetics and Molecular Biology PhD Student, Fapesp scholar, Universidade Estadual de Campinas, Genetic and Molecular Analysis Laboratory, Campinas, SP, CEP 13.083-875. E-mail: toledoesilva@gmail.com
2.Postdoctoral position, Universidade Estadual de Campinas, Fapesp scholar, Genetic and Molecular Analysis Laboratory, Campinas, SP, CEP 13.083-875. E-mail: melbio5@gmail.com
3.Researcher of Embrapa Gado de Corte, Campo Grande, MS, CEP 79.002-970. Email: rosangela@cnpgc.embrapa.br
4.Researcher of Embrapa Gado de Corte, Campo Grande, MS, CEP 79.002-970. Email: lchiari@cnpgc.embrapa.br
5.Researcher of Embrapa Cerrados, Planaltina, DF, CEP 73.310-970. Email: karia@cpac.embrapa.br
6 Researcher of Embrapa Cerrados, Planaltina, DF, CEP 73.310-970. Email: marcelo@cpac.embrapa.br
7.Researcher of Embrapa Cerrados, Planaltina, DF, CEP 73.310-970. Email: ffaleiro@cpac.embrapa.br
8.Researcher of São Paulo Agency of Technology and Agro-Business, Pólo Apta Centro Sul, Piracicaba, SP, CEP 13400-970. E-mail: mizucchi@apta.sp.gov.br
9.Teacher of Genetics and Evolution Department, Biology Institute and researcher of Molecular Biology and Genetic Engineering Center of Universidade Estadual de Campinas - Campinas, SP, CEP 13083-970. E-mail: anete@unicamp.br
Financial support: FAPESP (Grant # 2005/51010-0)

ranged from 0.27 to 0.74, with an average of 0.51, whereas the observed heterozygosity ranged from 0.04 to 0.87, with an average of 0.48. Genetic distances among *S. macrocephala* accessions ranged from 0 to 0.83 and averaged 0.54. In *S. capitata*, genetic distances ranged from 0 to 0.85, averaging 0.50.

STRUCTURE analysis identified five groups in the germplasm collection of *S. macrocephala* (Fig. 1). Using K=5, 57% of the accessions were assigned to one group with more than 80% probability. When those groups were considered to be different populations, the observed heterozygosity varied from 0.03 in group D to 0.14 in group C, while the expected heterozygosity ranged from 0.14 in group D to 0.38 in group C. Considering *S. capitata*, Bayesian clustering identified 4 groups in the germplasm collection of Embrapa-Cerrados (Fig. 1). With a K value of 4, 68% of the accessions were identified in a single group with more than 80% probability. Considering the groups obtained with STRUCTURE, the observed heterozygosity varied from 0.40 in group A to 0.56 in group C, while the expected heterozygosity ranged from 0.40 in group A to 0.49 in groups C and D. The proportion of genetic diversity due to differences among the *S. macrocephala* and *S.capitata* groups clustered by STRUCTURE was assessed by the estimate of Nei's gene diversity among groups (G_{ST}). It indicated that approximately 27% of the diversity was due to differences among groups in *S. macrocephala*. For *S. capitata* the same parameter value was approximately 11%. A tree was constructed based on the Roger's genetic distances and Neighbor-Joining (NJ) clustering methods (Fig. 2). In general, the groups formed by STRUCTURE were subdivided into more than one group in the distance method.

We also assembled core collections to represent 100% of the diversity found in this study for both species. For *S. macrocephala*, the 61 alleles identified at 13 microsatellite loci were fully represented by 23 accessions. For *S. capitata*, the 51 alleles identified at 15 loci were fully represented by 13 accessions.

Discussion

The microsatellite loci used in the present work were suitable for evaluating all the accessions of *S. macrocephala* and *S. capitata*. *S. capitata* accessions showed a high level of heterozygosity. The allopolyploid origin of *S. capitata* may influence the observed heterozygosities, as the *Stylosanthes* allotetraploid species shows a high level of fixed heterozygosity (VANDER STAPPEN et al., 2002)..

The genetic distances here described for *S. macrocephala* and *S. capitata* were higher than those reported for other *Stylosanthes* species. In addition to the more polymorphic nature of microsatellites, the higher genetic distances estimated here may also be due to the higher number of accessions studied.

STRUCTURE analysis showed five distinct groups on *S. macrocephala* germplasm collection, whereas between *S. capitata* accessions were formed four clusters. Among *S. macrocephala* clusters, group D was the most homogeneous, with the lowest genetic diversity, whereas group C showed the highest genetic diversity. *S. capitata* results show homogeneity among the groups, which may be a consequence of the high number of accessions from only two sampling locales, while the other geographical areas were represented by fewer accessions.

G_{ST} value found in *S. macrocephala* groups is similar to that observed for populations of species belonging to the Fabaceae family using isozymes (Hamrick and Godt, 1996). For *S. capitata* groups, the obtained value is lower than that found for other *Stylosanthes* species.

S. macrocephala accessions represent the Brazilian States of Minas Gerais, Goiás, Bahia, and Distrito Federal, with one accession sampled in the State of Piauí and with several other being of unknown origin. Based on STRUCTURE cluster analysis and expected heterozygosity values, we hypothesized that the *S. macrocephala* populations were derived from the population from the state of Bahia and are still in the process of differentiation. Accessions from Bahia were distributed in all five groups, indicating that those accessions are ancestors of the accessions from other states. In addition, group C, which is mostly made up of accessions from Bahia, presented the higher expected heterozygosity value, which also indicates a center of origin in this state. This hypothesis deserves further investigation, including natural population analysis.

The *S. capitata* accessions analyzed in the present work were collected in the diverse Brazilian states of Distrito Federal, Bolivia, Colombia, and Venezuela. The majority of accessions were collected in the states of Bahia and Goiás, with 54 and 40 accessions respectively. As most Brazilian states were represented by just a few accessions, it is difficult to reconcile the genetic clustering results with the different geographic regions. Studies with natural populations from diverse Brazilian regions are necessary to better understand the distribution of the genetic diversity of *S. capitata* in Brazil.

We constructed a tree based on the Roger's genetic distances NJ clustering methods for both species. In *S. macrocephala*, group C, which was the most diverse of the groups, was subdivided into four groups, while groups A and D remained mostly intact. Groups B and E were mostly divided into two distinct groups. In the *S. capitata* NJ tree, group A was subdivided into three groups. Accessions from STRUCTURE groups B and C remained together in the NJ tree, while group D was divided across all NJ groups. The Neighbor-Joining tree and STRUCTURE methods are approaches with different assumptions, and differences can be observed when comparing their results (WANG et al., 2009). Used together, both analyses can lead to a deeper understanding of the genetic diversity in the *S. macrocephala* and *S. capitata* germplasm collections.

The suggested core collections for *S. macrocephala* and *S. capitata* are represented by 23 and 13 accessions, respectively. Both results show that few accessions are needed to represent the genetic diversity of the entire collection. Due to the sampling present in the collection, which represents some regions with a large number of accessions and others with fewer, the overall genetic diversity was relatively low, and the number of accessions

needed for the core collection was low as well. In *S. capitata*, the high level of heterozygosity may be another reason for the size of the core collection (CIPRIANI et al., 2010). These results achieved the reduction in the number of accessions, helping in identification, evaluation, and utilization of the germplasm collections located in Embrapa Cerrados.

The present microsatellite analysis of both species of the genus *Stylosanthes* identified population structure among the analyzed accessions. The study will facilitate the management and utilization of the accessions and suggests conclusions about the geographical relationships of *S. macrocephala*. However, more samples from natural populations are needed from *S. capitata* for this type of analysis. Furthermore, the data indicated that a small number of microsatellite markers is effective for the assessment of genetic diversity in the *S. macrocephala* and *S. capitata* species, leading to a quick and low-cost methodology of screening *Stylosanthes* germplasm collections.

Ackowledgements

The authors would like to thank Alexandre Franco Garcia for his assistance with the graphical representations; Both M.O. Santos-Garcia and G. Toledo-Silva received scholarships from FAPESP (2005/52211-9, 2010/50301-9). A.P. Souza is recipient of a research fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

CIPRIANI, G.; SPADOTTO, A.; JURMAN, I. et al. The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L .) accessions uncovers new synonymy and parentages and reveals a large admixture amongst varieties of different geographic origin. **Theoretical and Applied Genetics**, vol. 121, p.1569-1585, 2010.

COSTA, N. **Revisão do género *Stylosanthes***. PhD thesis, Universidade Técnica de Lisboa. 2006.

CRESTE, S.; TULMANN NETO, A.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, vol.19, p.299-306, 2001.

FALEIRO, F.G.; FALEIRO, A.S.; CORDEIRO, M.R. et al. Metodologia para operacionalizar a extração de DNA de espécies nativas do cerrado visando a análises moleculares. **Comunicado Técnico Embrapa Cerrados** vol.92, n.5, 2003.

GOUDET, J. **FSTAT (Version 1.2): a computer program to calculate F-statistics**. Avaiable in: http://www2.unil.ch/popgen/softwares/fstat.htm. 2001.

HAMRICK, J.; GODT, M. Effects of life history traits on genetic diversity in plant species. **Philosophical Transaction of the Royal Society B** vol. 351, p.1291-1298, 1996.

LEWIS, P.O.; ZAYKIN, D. **Genetic Data Analysis: Computer program for the analysis of allelic data**. Avaiable in: http://en.bio-soft.net/dna/gda.html. 2000.

LIU, C.; MUSIAL, J.; THOMAS, B. Genetic relationship among *Stylosanthes* species revealed by RFLP and STS analyses. **Theoretical and Applied Genetics** , vol.99, p.1179-1186, 1999.

MANETJE, L. Considerations on the taxonomy of the genus Stylosanthes. In: STACE, H.M.; EDYE, L. (eds.) **The biology and agronomy of *Stylosanthes*.** Sydney: Academic Press, 1984, pp. 1-22.

MILLER, M.P. **Tools for population genetic analysis (TFPGA): A Windows program for the analysis of allozyme and molecular population genetic data.** Avaiable in: http://www.marksgeneticsoftware.net/tfpga. 1997.

PERRIER, X.; JACQUEMOUD-COLLET, J. **DARwin software**. Avaiable in: http://darwin.cirad.fr/darwin. 2006.

PRITCHARD, J.K.; STEPHENS, M.; DONNELLY, P. Inference of Population Structure Using Multilocus Genotype Data. **Genetics** ,vol.155, p. 945-959, 2000.

SANTOS, M.O.; SASSAKI, R.P.; FERREIRA, T.H.S. et al. Polymorphic microsatellite loci for *Stylosanthes macrocephala* Ferr. et Costa, a tropical forage legume. **Conservation Genetics Resources**, vol.1, p.481-485, 2009a.

SANTOS, M.O.; SASSAKI, R.P.; CHIARI, L.; et al. Isolation and characterization of microsatellite loci in tropical forage *Stylosanthes capitata* Vogel. **Molecular Ecology Resources**, vol. 9, p.192-194, 2009b.

STACE, H.M.; CAMERON, D. Cytogenetics and the evolution of *Stylosanthes*. In: STACE, H.M., EDYE, L. (eds.) **The biology and agronomy of *Stylosanthes*.** Sydney: Academic Press, 1984, pp. 49-72.

VANDER STAPPEN, J.; GAMA LOPEZ, S.; DAVILA, P. et al. Molecular evidence for the hybrid origin of a new endemic species of *Stylosanthes* Sw.(Fabaceae) from the Mexican Yucatán Peninsula. **Botanical journal of the linnean society**, vol. 140, n.1, p. 1-13, 2002.

WANG, M.L.; ZHU, C.; BARKLEY, N.A.; et al. Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. **Theoretical and applied genetics** vol. 120, p.13-23, 2009.

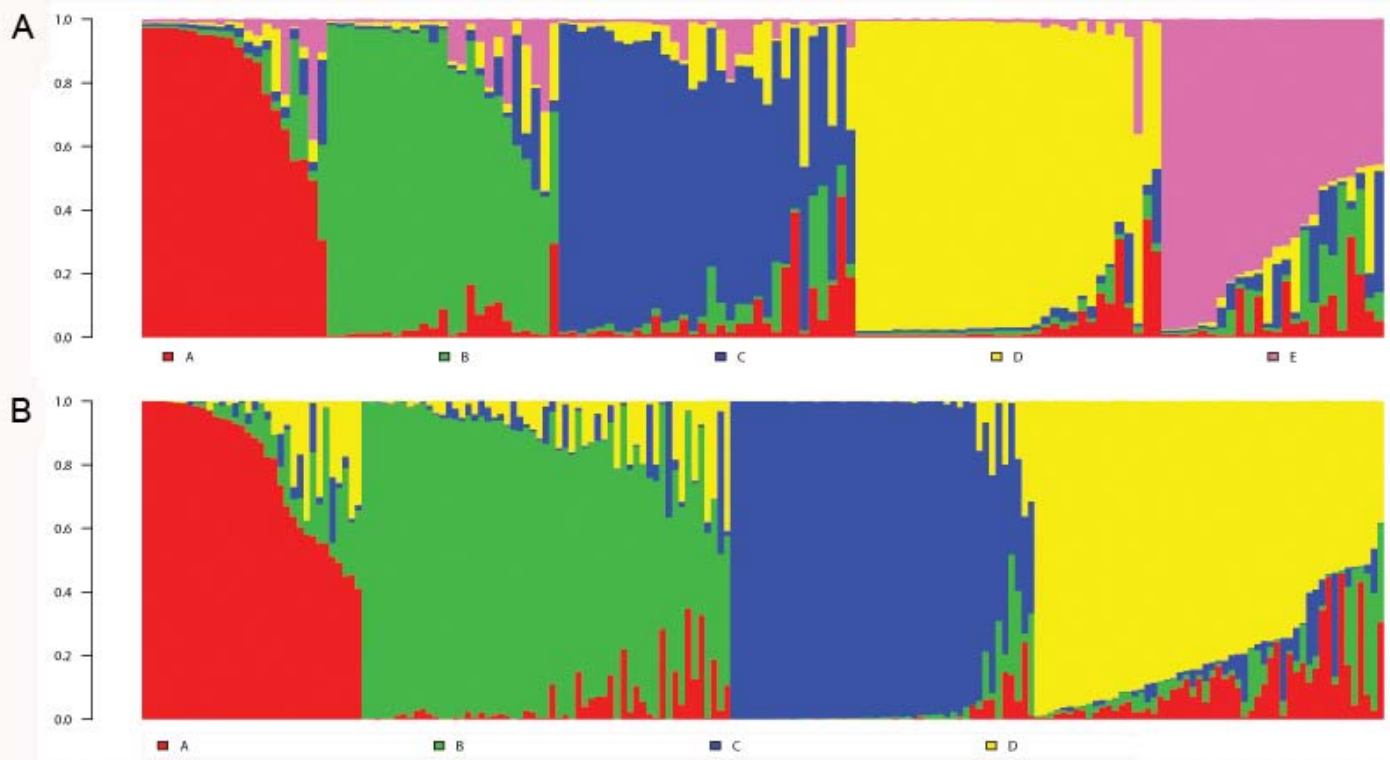

Figure 1. Barplot obtained from the model-based analysis of ancestry of the same *S. macrocephala* (A) and *S. capitata* (B) accessions implemented in the STRUCTURE software.

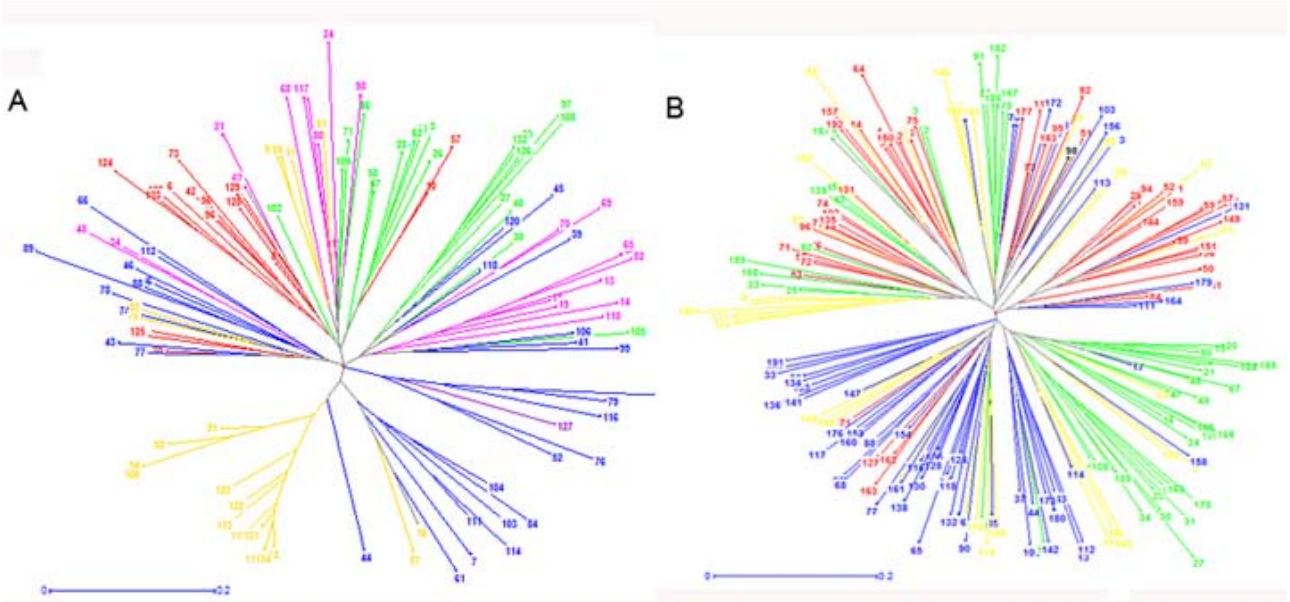

Figure 2: The Roger's genetic distance dendrogram of *S. macrocephala* (A) and *S. capitata* (B) accessions constructed using the Neighbor-Joining method implemented in DARwin.