Identification of *Stylosanthes guianensis* varieties using molecular genetic analysis

M. O. Santos-Garcia¹, C. T. Karia², R. M. S. Resende³, L. Chiari³, M. L. C. Vieira⁴, M. I. Zucchi⁵ and A. P. Souza¹,⁶*

¹ Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, CP 6010, 13083-970, Campinas, SP, Brazil
² Embrapa Cerrados, BR 020, Km 18, 73310-970, Planaltina, DF, Brazil
³ Embrapa Gado de Corte, BR 262, Km 4, 79002-970, Campo Grande, MS, Brazil
⁴ Departamento de Genética, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, CP 83, 13400-970, Piracicaba, SP, Brazil
⁵ Polo Regional Centro Sul APTA, SP 127, Km 30, 13400-970, Piracicaba, SP, Brazil
⁶ Departamento de Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, CP 6109, 13083-970, Campinas, SP, Brazil

Received: 2 June 2011; Returned for revision: 27 July 2011; Accepted: 5 January 2012; Published: 10 January 2012

Citation details: Santos-Garcia MO, Karia CT, Resende RMS, Chiari L, Vieira MLC, Zucchi MI, Souza AP. 2012. Identification of *Stylosanthes guianensis* varieties using molecular genetic analysis. *AoB PLANTS* 2012: pls001; doi:10.1093/aobpla/pls001

Abstract

**Background and aims**

The botanical classification of *Stylosanthes guianensis* is controversial, and few studies have used molecular markers to analyse this species. We used microsatellite markers to study the genetic diversity and population structure of *S. guianensis* and compare our results with the current infraspecific botanical classification.

**Methodology**

A representative sample from the *S. guianensis* Brazilian germplasm collection (150 accessions) was analysed using 20 microsatellite loci. A model-based Bayesian approach implemented in the software STRUCTURE was used to assign accessions into clusters. A dendrogram was constructed based on Rogers's genetic distances.

**Principal results**

The number of alleles per locus varied from 2 to 11, with an average of 4.7. The observed (Hₒ) and expected (Hₑ) heterozygosity values varied from 0 to 0.58 (mean of 0.18) and from 0.04 to 0.83 (mean of 0.55), respectively. Nine groups were assembled in STRUCTURE, and these groups were consistent with clusters inferred from the genetic distances and taxonomic varieties described for *S. guianensis*. The Gₛₜ among the nine groups was 0.46.

**Conclusions**

The low Hₒ and the Gₛₜ values observed are in agreement with the outcrossing rate (26 %) estimated for this species. The data indicate a high genetic diversity among and within the botanical varieties and suggest that microsatellite-based information can be combined with classical taxonomy to elucidate infraspecific levels.

Introduction

The genus *Stylosanthes* belongs to the family Fabaceae and consists of 48 species found in tropical and subtropical regions of the Americas, Africa and Southeast Asia (Costa and Ferreira 1984; Costa 2006). The genus has two centres of diversity, the most important being located in central Brazil (Ferreira and Costa 1979; Stace and Cameron 1984). It includes 45 % of all *Stylosanthes* species and exhibits the greatest degree of phenotypic variation and endemism (Costa 2006). Mexico and the

* Corresponding author’s e-mail address: anete@unicamp.br

Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Caribbean Islands are also major centres of *Stylosanthes* diversity (Stace and Cameron 1984).

*Stylosanthes guianensis* (AuBl.) Sw. \((2n = 20)\) is the most widespread *Stylosanthes* species and exhibits great phenotypic variation (Williams et al. 1984; Vieira et al. 1993). This species is native to South and Central America, where it is widely distributed, although not in the equatorial zone (Williams et al. 1984). *Stylosanthes guianensis* is considered to be a promising forage crop in the Brazilian savannahs, and many accessions produce a large amount of dry matter and retain their leaves during the dry season (Andrade and Karia 2000; Maass and Sawkins 2004).

The taxonomic classification of *S. guianensis* is controversial, and different taxonomic groups have been proposed based on different morphological characters. Mannetje (1977, 1984) recognized seven different *S. guianensis* varieties: *S. guianensis* var. *dissitiflora*, *S. guianensis* var. *gracilis*, *S. guianensis* var. *guianensis*, *S. guianensis* var. *intermedia*, *S. guianensis* var. *longiseta*, *S. guianensis* var. *marginata* and *S. guianensis* var. *robusta*. Ferreira and Costa (1979) considered Mannetje’s varieties to be different species. They proposed that *S. guianensis* was composed of the Mannetje *S. guianensis* var. *guianensis* but subdivided this species into four different botanical varieties. Analyses of random amplified polymorphic DNA (RAPD) markers (Kazan et al. 1993; Vieira et al. 1997), sequence-tagged site and internal transcribed spacer (ITS) sequences (Vander Stappen et al. 1998, 1999b) demonstrated a high genetic diversity among the varieties proposed by Mannetje (1984), supporting their separation into distinct species. More recently, Costa (2006) used ITS sequencing and classical taxonomy to review the genus *Stylosanthes*. The ITS data support the classification proposed by Ferreira and Costa (1979) and Brandão et al. (1985), according to whom there are four botanical varieties of *S. guianensis*: *S. guianensis* var. *guianensis*, *S. guianensis* var. *microcephala* and *S. guianensis* var. *pauciflora*. Vieira et al. (1993) analysed the karyotypes of the botanical varieties described by Ferreira and Costa (1979) and Brandão et al. (1985), and four species considered to be *S. guianensis* varieties by Mannetje (1977, 1984). Their results demonstrate that the varieties proposed by Ferreira and Costa (1979) and Brandão et al. (1985) have similar karyotypes, suggesting that *S. g. canescens*, *S. g. microcephala* and *S. g. pauciflora* may all have evolved from *S. g. guianensis*. The four species considered to be *S. guianensis* varieties by Mannetje (1977, 1984) have distinct karyotypes, supporting the classification proposed by Ferreira and Costa (1979) and Brandão et al. (1985).

Microsatellites, or simple sequence repeats (SSRs), are useful markers for a variety of applications in plant genetics because they are codominant, multiallelic, easily detected by polymerase chain reaction, relatively abundant and provide good coverage of the genome (Powell et al. 1996a). Microsatellite markers are available for three species of *Stylosanthes*: *S. guianensis* (Vander Stappen et al. 1999a; Santos et al. 2009a), *S. capitata* (Santos et al. 2009b) and *S. macrocephala* (Santos et al. 2009c). Vander Stappen et al. (1999a) described 18 genomic microsatellites in *S. guianensis* and tested them in 65 genotypes of *S. guianensis* and in the related species *S. gracilis*, *S. hippocampoides*, *S. grandifolia*, *S. acuminata* and *S. longiseta*. All 18 microsatellites were polymorphic among *S. guianensis* and its related species, and 16 were polymorphic within *S. guianensis*. Karia (2008) tested the microsatellites developed by Vander Stappen et al. (1999a) to evaluate the relationships among 437 accessions of the *S. guianensis* Brazilian germplasm collection and found that only seven were polymorphic among the accesses. Because of the limited number of markers, many relationships were unclear, and few could be clarified with respect to the population structure and to the clustering of the botanical varieties based on the results of Karia (2008). Considering the small number of polymorphic microsatellites available for this species and the need to better evaluate the Brazilian germplasm collection, Santos et al. (2009a) developed 46 new genomic microsatellites, 20 of which were polymorphic when tested in 20 *S. guianensis* accessions.

Santos-Garcia et al. (2011) tested both sets of microsatellites (Vander Stappen et al. 1999a; Santos et al. 2009a) in 20 accessions that were used as maternal parents in a progeny array and found that only five were polymorphic and could be used to determine the mating system of *S. guianensis*. Based on the microsatellite data, these authors estimated the outcrossing rate in *S. guianensis* to be 26 %, indicating that the species presents a mixed mating system with predominance of autogamy.

Despite the importance of *S. guianensis* as forage in tropical areas and the fact that some Brazilian regions contain most of the diversity of this genus, few studies have been conducted to determine the genetic diversity in the Brazilian germplasm collection. The Brazilian *S. guianensis* collection was established in a series of collecting trips around Brazil and other South American countries, and it has not been subjected to any breeding selection. Genetic studies of this collection may provide valuable information for planning new collection trips as well as for the study of natural populations. Karia (2008) was the first to study this collection using molecular markers, but this analysis did not provide enough information to evaluate the genetic diversity of...
the collection or to clarify the botanical classifications and the relationships among the accessions.

Considering the classifications proposed by Ferreira and Costa (1979), Brandão et al. (1985) and Costa (2006), we have selected a representative sample of the S. guianensis Brazilian germplasm collection consisting of 150 accessions that were either classified as one of the four botanical varieties or remained unclassified. We are dealing with botanical varieties of S. guianensis in this study, and their classification is still controversial. Currently, microsatellite markers have been the most widely used for inferring population structure that ultimately defines variation at the infraspecific level.

Considering this, we used microsatellite markers to analyse the S. guianensis accessions with the goals of (i) estimating the genetic diversity and relationships present among the accessions, (ii) determining the population structure of the samples and (iii) comparing the observed population structure to the proposed botanical varieties. The microsatellite-based population structure observed in these samples was generally in agreement with the proposed botanical varieties of S. guianensis.

Materials and methods

Plant material

A total of 150 accessions from the Brazilian Institution Embrapa Cerrados were selected to evaluate the genetic diversity, the population structure and the relationships among them. A list of the accessions with their sample codes, accession numbers, places of origin and botanical varieties is provided in Table 1.

DNA extraction

DNA was extracted from the bulk leaf samples of four plants per accession using the cetyltrimethyl ammonium bromide method (Doyle and Doyle 1990) as modified by Bellon et al. (2007).

SSR amplification and detection

Twenty microsatellite loci were amplified from the DNA of all of the sampled accessions as previously described by Santos et al. (2009a). Amplified fragments were separated on 6% acrylamide gels and visualized by silver staining according to Creste et al. (2001).

Data analysis

The Genetic Data Analysis program (Lewis and Zaykin 2000) was used to estimate the observed and expected heterozygosities. Allele frequencies and Roger’s genetic distance, as modified by Wright (1978), were estimated between all pairs of accessions using Tools for Population Genetic Analysis (Miller 1997).

The STRUCTURE software package (Pritchard et al. 2000) was used to subdivide the accessions into the appropriate number of genetic clusters independent of any prior information concerning the geographic origin of the accessions. STRUCTURE uses a Bayesian approach to identify clusters based on their fit to Hardy–Weinberg proportions and their linkage equilibria. Ten STRUCTURE runs were performed by setting the number of populations (K) from 1 to 20. For each run, the burn-in time was set to 200 000, and the replication number was set to 300 000. The most likely number of distinct clusters (K) was determined according to the procedure described by Evanno et al. (2005). A graphical representation of the population assignments obtained from STRUCTURE was generated using R v.2.9.1.

DARwin 5.0 (Perrier and Jacquemoud-Collet 2006) was used to define the genetic relationships among accessions based on Roger’s genetic distance and the UPGMA (unweighted paired group method) clustering method. FSTAT (Goudet 2001) was used to estimate Nei’s $G_{ST}$ (Nei 1973) among the groups obtained by the STRUCTURE analysis.

Results

We tested 20 genomic microsatellite loci developed by Santos et al. (2009a) and found that all loci were polymorphic within the group of accessions studied here. These 20 microsatellites generated 94 alleles that were used to study the genetic diversity among 150 S. guianensis accessions. The number of alleles ranged from 2 to 11 (Table 2), with an average of 4.7 alleles per locus. The SG03G8 locus exhibited the greatest number of alleles (11), followed by the SG03E2 locus with 10 alleles.

The observed and expected heterozygosities values ($H_o$ and $H_e$) ranged from 0.00 to 0.58 (mean of 0.18) and from 0.04 to 0.83 (mean of 0.55), respectively. The markers with a higher observed heterozygosity were SG01D7 and SG01B9. In the case of bulk samples, a heterozygous pattern may be observed as a result of the presence of heterozygous plants in the DNA pool or variation within the accession (Zhang et al. 2008).

Roger’s genetic distances were estimated for each pair of accessions and varied from 0.00 (among several accessions) to 0.94 (between accessions 19 and 140), with an average of 0.66.

A model-based analysis performed by STRUCTURE was used to investigate the possible population structure within the collection sample (150 accessions). The
Table 1 The *S. guianensis* collection used to assess the genetic diversity at 20 microsatellite loci. Shown are the sample codes, the respective accession numbers in the Embrapa Cerrados and CIAT (International Center for Tropical Agriculture) germplasm collections, the geographic collection site and the botanical variety of each accession.

<table>
<thead>
<tr>
<th>Code</th>
<th>Embrapa Cerrados</th>
<th>CIAT number</th>
<th>Geographic site</th>
<th>Botanical variety</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>135</td>
<td>1297</td>
<td>Distrito Federal</td>
<td>pauciflora</td>
</tr>
<tr>
<td>2</td>
<td>464</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>n.a.</td>
</tr>
<tr>
<td>3</td>
<td>517</td>
<td>1705</td>
<td>Mato Grosso</td>
<td>n.a.</td>
</tr>
<tr>
<td>4</td>
<td>648</td>
<td>1950</td>
<td>Colombia</td>
<td>guianensis</td>
</tr>
<tr>
<td>5</td>
<td>662</td>
<td>136</td>
<td>Colombia</td>
<td>n.a.</td>
</tr>
<tr>
<td>6</td>
<td>1121</td>
<td>1507</td>
<td>Venezuela</td>
<td>pauciflora</td>
</tr>
<tr>
<td>7</td>
<td>1132</td>
<td>1890</td>
<td>Venezuela</td>
<td>pauciflora</td>
</tr>
<tr>
<td>8</td>
<td>1139</td>
<td>1975</td>
<td>Colombia</td>
<td>n.a.</td>
</tr>
<tr>
<td>9</td>
<td>1144</td>
<td>2034</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>10</td>
<td>1153</td>
<td>2315</td>
<td>Bahia</td>
<td>n.a.</td>
</tr>
<tr>
<td>11</td>
<td>1229</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>microcephala</td>
</tr>
<tr>
<td>12</td>
<td>1230</td>
<td>2950</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>13</td>
<td>1231</td>
<td>2951</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>14</td>
<td>1235</td>
<td>n.a.</td>
<td>Goiás</td>
<td>n.a.</td>
</tr>
<tr>
<td>15</td>
<td>1237</td>
<td>2615</td>
<td>Tocantins</td>
<td>pauciflora</td>
</tr>
<tr>
<td>16</td>
<td>1239</td>
<td>2659</td>
<td>Pará</td>
<td>n.a.</td>
</tr>
<tr>
<td>17</td>
<td>1354</td>
<td>2991</td>
<td>Espírito Santo</td>
<td>pauciflora</td>
</tr>
<tr>
<td>18</td>
<td>1360</td>
<td>2549</td>
<td>Piauí</td>
<td>pauciflora</td>
</tr>
<tr>
<td>19</td>
<td>1365</td>
<td>2812</td>
<td>Venezuela</td>
<td>pauciflora</td>
</tr>
<tr>
<td>20</td>
<td>1368</td>
<td>2742</td>
<td>Minas Gerais</td>
<td>pauciflora</td>
</tr>
<tr>
<td>21</td>
<td>1369</td>
<td>2529</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>22</td>
<td>1371</td>
<td>10 107</td>
<td>Goiás</td>
<td>microcephala</td>
</tr>
<tr>
<td>23</td>
<td>1372</td>
<td>2439</td>
<td>Pernambuco</td>
<td>pauciflora</td>
</tr>
<tr>
<td>24</td>
<td>1619</td>
<td>2748</td>
<td>Venezuela</td>
<td>n.a.</td>
</tr>
<tr>
<td>25</td>
<td>2203</td>
<td>2987</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>26</td>
<td>2464</td>
<td>10 993</td>
<td>Tocantins</td>
<td>guianensis</td>
</tr>
<tr>
<td>27</td>
<td>2725</td>
<td>2974</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>28</td>
<td>2734</td>
<td>2458</td>
<td>Pará</td>
<td>pauciflora</td>
</tr>
<tr>
<td>29</td>
<td>2738</td>
<td>2542</td>
<td>Ceará</td>
<td>pauciflora</td>
</tr>
<tr>
<td>30</td>
<td>2740</td>
<td>2708</td>
<td>Goiás</td>
<td>pauciflora</td>
</tr>
<tr>
<td>31</td>
<td>2741</td>
<td>2436</td>
<td>Alagoas</td>
<td>pauciflora</td>
</tr>
<tr>
<td>32</td>
<td>2761</td>
<td>2727</td>
<td>Distrito Federal</td>
<td>pauciflora</td>
</tr>
<tr>
<td>33</td>
<td>2771</td>
<td>2992</td>
<td>Espírito Santo</td>
<td>pauciflora</td>
</tr>
<tr>
<td>34</td>
<td>4142</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>guianensis</td>
</tr>
<tr>
<td>35</td>
<td>4144</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>canescens</td>
</tr>
<tr>
<td>36</td>
<td>4157</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>37</td>
<td>4172</td>
<td>n.a.</td>
<td>Goiás</td>
<td>canescens</td>
</tr>
</tbody>
</table>

Continued
Table 1 Continued

<table>
<thead>
<tr>
<th>Code</th>
<th>Embraopra Cerrados</th>
<th>CIAT number</th>
<th>Geographic site&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Botanical variety&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>4173</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>canescens</td>
</tr>
<tr>
<td>39</td>
<td>4174</td>
<td>n.a.</td>
<td>Maranhão</td>
<td>n.a.</td>
</tr>
<tr>
<td>40</td>
<td>4188</td>
<td>n.a.</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>41</td>
<td>4193</td>
<td>n.a.</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>42</td>
<td>4199</td>
<td>n.a.</td>
<td>Bahia</td>
<td>pauciflora</td>
</tr>
<tr>
<td>43</td>
<td>4227</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>canescens</td>
</tr>
<tr>
<td>44</td>
<td>4233</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>canescens</td>
</tr>
<tr>
<td>45</td>
<td>4237</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>46</td>
<td>4238</td>
<td>n.a.</td>
<td>Rio de Janeiro</td>
<td>guianensis</td>
</tr>
<tr>
<td>47</td>
<td>4239</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>48</td>
<td>4240</td>
<td>n.a.</td>
<td>Rio de Janeiro</td>
<td>guianensis</td>
</tr>
<tr>
<td>49</td>
<td>4262</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>50</td>
<td>4264</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>51</td>
<td>4286</td>
<td>n.a.</td>
<td>Goiás</td>
<td>n.a.</td>
</tr>
<tr>
<td>52</td>
<td>4292</td>
<td>n.a.</td>
<td>Goiás</td>
<td>pauciflora</td>
</tr>
<tr>
<td>53</td>
<td>4300</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>54</td>
<td>4302</td>
<td>n.a.</td>
<td>Rio de Janeiro</td>
<td>microcephala</td>
</tr>
<tr>
<td>55</td>
<td>4306</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>56</td>
<td>4308</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>57</td>
<td>4314</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>58</td>
<td>4315</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>59</td>
<td>4322</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>60</td>
<td>4323</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>61</td>
<td>4324</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>62</td>
<td>4331</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>63</td>
<td>4336</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>n.a.</td>
</tr>
<tr>
<td>64</td>
<td>4338</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>n.a.</td>
</tr>
<tr>
<td>65</td>
<td>4364</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>n.a.</td>
</tr>
<tr>
<td>66</td>
<td>4528</td>
<td>n.a.</td>
<td>Goiás</td>
<td>n.a.</td>
</tr>
<tr>
<td>67</td>
<td>4530</td>
<td>n.a.</td>
<td>n.a.</td>
<td>guianensis</td>
</tr>
<tr>
<td>68</td>
<td>5187</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>pauciflora</td>
</tr>
<tr>
<td>69</td>
<td>5192</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>pauciflora</td>
</tr>
<tr>
<td>70</td>
<td>5204</td>
<td>n.a.</td>
<td>Goiás</td>
<td>canescens</td>
</tr>
<tr>
<td>71</td>
<td>5206</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>72</td>
<td>5221</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>canescens</td>
</tr>
<tr>
<td>73</td>
<td>5233</td>
<td>n.a.</td>
<td>São Paulo</td>
<td>guianensis</td>
</tr>
<tr>
<td>74</td>
<td>5236</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>microcephala</td>
</tr>
<tr>
<td>75</td>
<td>5239</td>
<td>n.a.</td>
<td>Goiás</td>
<td>microcephala</td>
</tr>
<tr>
<td>76</td>
<td>5248</td>
<td>n.a.</td>
<td>Goiás</td>
<td>pauciflora</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Code</th>
<th>Embrapa Cerrados</th>
<th>CIAT number</th>
<th>Geographic site</th>
<th>Botanical variety</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
<td>5262</td>
<td>n.a.</td>
<td>Mato Grosso do Sul</td>
<td>microcephala</td>
</tr>
<tr>
<td>78</td>
<td>5277</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>pauciflora</td>
</tr>
<tr>
<td>79</td>
<td>5279</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>microcephala</td>
</tr>
<tr>
<td>80</td>
<td>5294</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>81</td>
<td>5295</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>82</td>
<td>5349</td>
<td>n.a.</td>
<td>Bahia</td>
<td>n.a.</td>
</tr>
<tr>
<td>83</td>
<td>1361</td>
<td>2445</td>
<td>Pernambuco</td>
<td>n.a.</td>
</tr>
<tr>
<td>84</td>
<td>5426</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>85</td>
<td>5428</td>
<td>10 488</td>
<td>Pará</td>
<td>n.a.</td>
</tr>
<tr>
<td>86</td>
<td>5429</td>
<td>10 416</td>
<td>Pará</td>
<td>n.a.</td>
</tr>
<tr>
<td>87</td>
<td>5430</td>
<td>2649</td>
<td>Maranhão</td>
<td>n.a.</td>
</tr>
<tr>
<td>88</td>
<td>5432</td>
<td>2700</td>
<td>Goiás</td>
<td>microcephala</td>
</tr>
<tr>
<td>89</td>
<td>5433</td>
<td>2677</td>
<td>Tocantins</td>
<td>microcephala</td>
</tr>
<tr>
<td>90</td>
<td>5435</td>
<td>101 24</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>91</td>
<td>5436</td>
<td>10 126</td>
<td>Espírito Santo</td>
<td>n.a.</td>
</tr>
<tr>
<td>92</td>
<td>5439</td>
<td>10 792</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>93</td>
<td>5440</td>
<td>10 793</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>94</td>
<td>5441</td>
<td>10 794</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>95</td>
<td>5445</td>
<td>10 799</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>96</td>
<td>5447</td>
<td>10 802</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>97</td>
<td>5454</td>
<td>10 825</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>98</td>
<td>5456</td>
<td>10 849</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>99</td>
<td>5458</td>
<td>10 852</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>100</td>
<td>5460</td>
<td>10 854</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>101</td>
<td>5462</td>
<td>10 876</td>
<td>Minas Gerais</td>
<td>canescens</td>
</tr>
<tr>
<td>102</td>
<td>5464</td>
<td>10 888</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>103</td>
<td>5469</td>
<td>10 808</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>104</td>
<td>5471</td>
<td>10 826</td>
<td>Minas Gerais</td>
<td>microcephala</td>
</tr>
<tr>
<td>105</td>
<td>5474</td>
<td>10 872</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>106</td>
<td>5476</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>microcephala</td>
</tr>
<tr>
<td>107</td>
<td>5477</td>
<td>10 855</td>
<td>n.a.</td>
<td>microcephala</td>
</tr>
<tr>
<td>108</td>
<td>5484</td>
<td>10 814</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>109</td>
<td>5488</td>
<td>10 821</td>
<td>Minas Gerais</td>
<td>pauciflora</td>
</tr>
<tr>
<td>110</td>
<td>5489</td>
<td>10 824</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>111</td>
<td>5493</td>
<td>10 830</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>112</td>
<td>5500</td>
<td>10 838</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>113</td>
<td>5505</td>
<td>10 843</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>114</td>
<td>5506</td>
<td>10 820</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>115</td>
<td>5509</td>
<td>10 847</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>Code</td>
<td>Embopa Cerrados</td>
<td>CIAT number</td>
<td>Geographic site(^a)</td>
<td>Botanical variety(^b)</td>
</tr>
<tr>
<td>------</td>
<td>----------------</td>
<td>-------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>116</td>
<td>5510</td>
<td>10 848</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>117</td>
<td>5513</td>
<td>10 858</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>118</td>
<td>5514</td>
<td>10 859</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>119</td>
<td>5515</td>
<td>10 862</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>120</td>
<td>5516</td>
<td>10 863</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>121</td>
<td>5518</td>
<td>10 866</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>122</td>
<td>5519</td>
<td>10 867</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>123</td>
<td>5523</td>
<td>10 882</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>124</td>
<td>5525</td>
<td>10 886</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>125</td>
<td>5532</td>
<td>10 896</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>126</td>
<td>5533</td>
<td>10 899</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>127</td>
<td>5536</td>
<td>10 904</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>128</td>
<td>5538</td>
<td>10 908</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>129</td>
<td>EPAMIG-493(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>130</td>
<td>2745</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>131</td>
<td>EPAMIG-1044(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>132</td>
<td>EPAMIG-1448(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>133</td>
<td>EPAMIG-443(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>134</td>
<td>NC 1099 A(^c)</td>
<td>n.a.</td>
<td>Minas Gerais</td>
<td>n.a.</td>
</tr>
<tr>
<td>135</td>
<td>NC 2270(^c)</td>
<td>n.a.</td>
<td>Pará</td>
<td>guianensis</td>
</tr>
<tr>
<td>136</td>
<td>n.a.</td>
<td>10 283</td>
<td>Minas Gerais</td>
<td>guianensis</td>
</tr>
<tr>
<td>137</td>
<td>n.a.</td>
<td>10 285</td>
<td>Minas Gerais</td>
<td>pauciflora</td>
</tr>
<tr>
<td>138</td>
<td>IPF 394/75(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>139</td>
<td>EPAMIG-1529(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>140</td>
<td>EPAMIG-1557(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>guianensis</td>
</tr>
<tr>
<td>141</td>
<td>EPAMIG-1670(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>guianensis</td>
</tr>
<tr>
<td>142</td>
<td>EPAMIG-1787(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>guianensis</td>
</tr>
<tr>
<td>143</td>
<td>EPAMIG-1691(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>guianensis</td>
</tr>
<tr>
<td>144</td>
<td>EPAMIG-1994(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>guianensis</td>
</tr>
<tr>
<td>145</td>
<td>2600</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>n.a.</td>
</tr>
<tr>
<td>146</td>
<td>2676</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>microcephala</td>
</tr>
<tr>
<td>147</td>
<td>2689</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>microcephala</td>
</tr>
<tr>
<td>148</td>
<td>2694</td>
<td>n.a.</td>
<td>Tocantins</td>
<td>microcephala</td>
</tr>
<tr>
<td>149</td>
<td>LC 4297(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>150</td>
<td>LC 4471(^c)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

\(^a\)Geographic sites indicate the Brazilian state or other country where the plant was collected.

\(^b\)According to Ferreira and Costa (1979) and Brandão et al. (1985).

\(^c\)For accessions without a number in the Embrapa Cerrados collection, the identification numbers in other germplasm collections or collector numbers are shown.
The results of the analysis of 150 accessions of *S. guianensis* using 20 microsatellite loci (Santos et al. 2009a). Shown are the locus names, primer sequences, repeat motifs, number of alleles per locus (A), observed heterozygosity (H₀) and expected (Hₑ) heterozygosity.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequence (5′ – 3′)</th>
<th>Repeat motif</th>
<th>A</th>
<th>H₀</th>
<th>Hₑ</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG03A7</td>
<td>5′TACGGAAGTCCTACATTTGAGG3′ 5′GGCTGCGGAAGTGGACG3′</td>
<td>(AG)₈</td>
<td>6</td>
<td>0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>SG03B5</td>
<td>5′AGGTGAGTCAGGTGTTCTT3′ 5′TCCCTTCTACCGGATGTC3′</td>
<td>(TG)₇</td>
<td>2</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>SG03D4</td>
<td>5′GATCGGTCGGTTGGCCTACTA3′ 5′CCTATTATCCCCATCCCTCAC3′</td>
<td>(GT)₈</td>
<td>3</td>
<td>0.03</td>
<td>0.59</td>
</tr>
<tr>
<td>SG03E1</td>
<td>5′CCAAAGCGAGGAATGATGAG3′ 5′CAATGGGACGGAAGGACA3′</td>
<td>(GT)₈</td>
<td>2</td>
<td>0.00</td>
<td>0.43</td>
</tr>
<tr>
<td>SG03D1</td>
<td>5′GTGCGAAAAAATCTAAATGTTC3′ 5′GTTGAAATCTCTAAGTGAAG3′</td>
<td>(GT)₆</td>
<td>4</td>
<td>0.11</td>
<td>0.59</td>
</tr>
<tr>
<td>SG03G2</td>
<td>5′GTGAGAAATGAGGAGGAGAGA3′ 5′GAGGAAAAAAAGGCAAGC3′</td>
<td>(GT)₆</td>
<td>7</td>
<td>0.11</td>
<td>0.78</td>
</tr>
<tr>
<td>SG03G4</td>
<td>5′TTGCTCTTTATCTTTGCTACTA3′ 5′ATCAAGAATCTCAATCAGAGA3′</td>
<td>(TG)₈</td>
<td>5</td>
<td>0.06</td>
<td>0.68</td>
</tr>
<tr>
<td>SG01B9</td>
<td>5′TTACCGGAAACCCGAAAC3′ 5′GCACCTAACAAAGCCTACAC3′</td>
<td>(AC)₇</td>
<td>5</td>
<td>0.54</td>
<td>0.58</td>
</tr>
<tr>
<td>SG03G8</td>
<td>5′AATTTAGAGGAGGAGGAGGAATG3′ 5′TGAGAAGAATTAAAGACAGTGAG3′</td>
<td>(CT)₁₁ … (CA)₉</td>
<td>11</td>
<td>0.10</td>
<td>0.78</td>
</tr>
<tr>
<td>SG01C2</td>
<td>5′TGAAAGCACAAGGATAAGGA3′ 5′CGAACCACCAACCACAT3′</td>
<td>(GT)₇</td>
<td>2</td>
<td>0.05</td>
<td>0.36</td>
</tr>
<tr>
<td>SG01B12</td>
<td>5′ATGTTAGGAGATAGTGATTA3′ 5′TCGAGTTGATGTTGATTATG3′</td>
<td>(AC)₇ … (CT)₄ (TC)₄</td>
<td>2</td>
<td>0.15</td>
<td>0.44</td>
</tr>
<tr>
<td>SG01D7</td>
<td>5′ACATAACTGAGCCGATGAT3′ 5′ATTGTTGGCTGTAAGTCT3′</td>
<td>(AC)₇</td>
<td>4</td>
<td>0.58</td>
<td>0.51</td>
</tr>
<tr>
<td>SG01D3</td>
<td>5′AGATGCGTCAATCCGAGATA3′ 5′TCGAGTTGTGGTTGTGTTG3′</td>
<td>(AC)₆ … (AC)₇</td>
<td>2</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>SG03A9</td>
<td>5′AGTCCCCGATCCGAAAAAC3′ 5′AACCCCTTTAAAACACACAA3′</td>
<td>(CA)₇</td>
<td>2</td>
<td>0.00</td>
<td>0.45</td>
</tr>
<tr>
<td>SG03E10</td>
<td>5′TCCACGCTGCATGAAGAAATG3′ 5′GGACCCGAGACCCATCT3′</td>
<td>(GT)₇</td>
<td>7</td>
<td>0.20</td>
<td>0.83</td>
</tr>
<tr>
<td>SG03B10</td>
<td>5′CTTAACCTTAACTGAAATGAG3′ 5′AAAGGAAACAAAGGACGAGAGAC3′</td>
<td>(GT)₂ (GT)₄ (TG)₃ (GT)₆</td>
<td>7</td>
<td>0.35</td>
<td>0.79</td>
</tr>
<tr>
<td>SG03E2</td>
<td>5′GTGCCCTTGACGCCCTTATTG3′ 5′GAGGGGCGATCGGAGTGT3′</td>
<td>(AC)₈</td>
<td>10</td>
<td>0.35</td>
<td>0.68</td>
</tr>
<tr>
<td>SG03E7</td>
<td>5′AAGATGCGGCAAAGAAGAAAC3′ 5′GGTGGGCTCGCTACGATAC3′</td>
<td>(AC)₆ … (CT)₃</td>
<td>4</td>
<td>0.03</td>
<td>0.73</td>
</tr>
<tr>
<td>SG01A7</td>
<td>5′TACCTTGAATCCGCACTGATG3′ 5′ACCCGAAACCTCTTCTCATAA3′</td>
<td>(AG)₅ … (GT)₆</td>
<td>5</td>
<td>0.49</td>
<td>0.56</td>
</tr>
<tr>
<td>SG03E9</td>
<td>5′GGTCGAAATGGGGGAAAGAAA3′ 5′ATCGAAGAGAGGAAGGCTACT3′</td>
<td>(AC)₅ … (CA)₅</td>
<td>4</td>
<td>0.11</td>
<td>0.40</td>
</tr>
</tbody>
</table>
The number of clusters was set to nine, following the statistics presented in Evanno et al. (2005). In the ancestry plot (Fig. 1B), each accession is represented by a horizontal bar, and the length of each segment of the bar is proportional to the accession’s estimated ancestry fraction from each of the nine groups. The same ancestry plot is shown in more detail in Supplementary Figure S1 [see Additional Information], which includes the sample codes for the accessions. The majority of the model-based groups were in agreement with the currently proposed *S. guianensis* taxonomic classification (Ferreira and Costa 1979; Brandão et al. 1985). Some of the groups could also be correlated with the geographic origins of the accessions.

**Fig. 1** (A) The Roger’s genetic distance dendrogram of 150 *S. guianensis* accessions constructed using the UPGMA method implemented in DARwin. The colours in the dendrogram indicate the accession’s group as assigned by STRUCTURE analysis. (B) The bar plot obtained from the model-based ancestry analysis of the same *S. guianensis* accessions implemented in the STRUCTURE software.
The number of accessions belonging to each variety that were assigned to one of the nine groups with >80% probability is shown in Table 3. Of the 40 accessions belonging to *S. g. guianensis*, 11 were assigned to Group C, 5 to Group H and 14 to Group I. These groups are formed mostly by accessions belonging to this variety, which showed the greatest genetic diversity among all *Stylosanthes* varieties. The remaining 10 accessions from *S. g. guianensis* were distributed along with *S. g. pauciflora* accessions. Three of the accessions were assigned to Group B, two were assigned to Group G, and five were not assigned to any group with >80% probability. Of the 28 *S. g. pauciflora* accessions, 13 were assigned to Group A (formed mostly by this variety), seven were assigned with *S. g. guianensis* in Groups B and G, two were in distinct Groups (E and I), and six were not assigned to any group with >80% probability.

The 15 *S. g. microcephala* were mostly clustered in Groups D and E, both with six accessions. Group D contained six accessions belonging to *S. g. microcephala* and two non-classified accessions, and Group E contained six accessions belonging to *S. g. microcephala*, one accession from *S. g. pauciflora* and three non-classified accessions. Of the remaining three *S. g. microcephala* accessions, two were assigned to Group I and one to Group A.

The *S. g. canescens* accessions were clustered in a distinct Group (F) composed exclusively of this variety. Of the 20 *S. g. canescens* accessions studied, 14 were assigned to Group F.

In the dendrogram constructed using DARwin and the UPGMA method (Fig. 1A), the clusters, with few exceptions, were consistent with the groups generated using the Bayesian approach in STRUCTURE. The same dendrogram labelled with the sample code for the accessions is provided in Supplementary Figure S2 [see Additional Information].

The genetic differentiation among the *S. guianensis* groups, as clustered by the Bayesian approach, was estimated based on Nei’s *G_ST* as 0.46, indicating that 46% of differences resulted from the variation among groups. The remaining 54% was a function of the genetic variation within groups.

**Discussion**

All 20 microsatellite markers from Santos et al. (2009a) were polymorphic among the 150 accessions analysed. The same set of microsatellite markers was tested previously (Santos-Garcia et al. 2011), but only three were found to be polymorphic in that study, probably because the previous study was based on samples from breeding material, and many of the genotypes were closely related. In contrast, we have studied a diverse germplasm collection of accessions obtained in collecting trips around Brazil and other countries in South America.

The mean number of alleles per locus observed in this study (4.7) was higher than that reported by Vander Stappen et al. (1999a) for 65 *S. guianensis* genotypes and some related species but lower than that observed by Karia (2008) among 437 *S. guianensis* accessions; those authors reported mean numbers of alleles per locus of 3.7 and 6.43, respectively. These differences can be attributed to the number of accessions studied by each of the authors.

Considering the observed and expected heterozygosities, our data revealed a deficit in heterozygosity that was consistent with the *S. guianensis* outcrossing rate (26%) that was estimated based on microsatellite data (Santos-Garcia et al. 2011). The predominance of autogamy reduced the number of heterozygous samples, but some heterozygosity was maintained as a result of the 26% outcrossing.

The genetic distances found among the studied accessions were higher than those previously reported for the species. The highest genetic distance (0.94) was observed between accessions 19 and 140. Accession 19 belongs to *S. g. pauciflora* and was collected in
Venezuela; accession 140 belongs to S. g. guianensis and was collected in the Brazilian state of Minas Gerais. Accession 140 was one of the most divergent, and its mean genetic distance (0.75) is greater than the overall mean genetic distance (0.66). Kazan et al. (1993) studied 31 S. guianensis accessions and observed a mean genetic distance of 0.26 using RAPD. Faleiro et al. (2003) also used RAPD to study 35 accessions and observed genetic distances varying from 0.04 to 0.54. Microsatellite markers are codominant and multiallelic with a high degree of polymorphism, making them more useful than dominant markers for revealing diversity (Laborda et al. 2005). Powell et al. (1996b) compared the expected heterozygosities and the estimated genetic similarities based on different molecular marker types for the evaluation of a set of soybean accessions. Their results show that microsatellites have higher expected heterozygosity (0.61) and lower estimated genetic similarities (0.45) among accessions relative to RAPD markers (0.31 and 0.72, respectively).

STRUCTURE analysis was performed to investigate the population structure in the germplasm collection to compare the genetic population structure with the described botanical varieties. The nine groups generated based on microsatellite analysis were mostly consistent with the botanical classification. The S. g. microcephala and S. g. canescens varieties were well differentiated and formed individual groups. Although most of the accessions belonging to S. g. guianensis and S. g. pauciflora formed distinct groups, some of them were mixed together in other groups. These mixed groups generally contained small numbers of individuals, many of which were not classified into varieties. Stylosanthes g. pauciflora was recognized as a new botanical variety in 1985 (Brandão et al. 1985); the three other varieties were recognized in 1979 (Ferreira and Costa 1979). This may have affected the classification process, and some of the accessions belonging to S. g. pauciflora may have been incorrectly classified as S. g. guianensis.

In previous studies, most S. g. microcephala accessions did not group together (Faleiro et al. 2003; Karia 2008). In the present work, these samples clustered into two diverse groups. Of the 15 S. g. microcephala accessions, five from the state of Minas Gerais were assigned to Group D, and six accessions (two from the state of Goiás and four from the state of Tocantins) were assigned to Group E. In both groups, the majority of the accessions were S. g. microcephala. The soil in Goiás and Tocantins is considered less fertile than the soil in Minas Gerais, and this difference in soil fertility may be the main driver of the observed genetic differentiation (Karia 2008). Our data reinforce the idea that the presence of S. g. microcephala in Minas Gerais is associated with more fertile soils, as proposed by Costa (2006). However, further studies are necessary to address this question and should include the collection and analysis of more plants from those regions. If genetic differentiation based on soil conditions is confirmed by further studies, this information could be useful for the development of new commercial varieties that are adapted to specific soil conditions.

The genetic data obtained in this study are also consistent with previously published karyological findings. All varieties have similar total chromosome lengths except for S. g. microcephala, which has visibly smaller chromosomes and the most asymmetrical karyotype (Fig. 2) (Vieira et al. 1993). This variety predominantly formed Groups D and E (Fig. 1B). Stylosanthes g. canescens and S. g. microcephala are quite close in the dendrogram (Fig. 1A); in those varieties, chromosome 10 is submetacentric (2.58 and 2.55 as arm ratios, respectively) (Table 4), which is distinct from chromosome 10 in the other varieties. However, S. g. canescens has a uniquely submetacentric chromosome 8 (1.76 arm ratio). As revealed from the UPGMA-based dendrogram and the model implemented in the STRUCTURE software, S. g. guianensis and S. g. pauciflora share alleles at many of the microsatellite loci under investigation. The karyotypes of these two varieties are also very similar, except for chromosome 10, which is significantly smaller in S. g. guianensis than in S. g. pauciflora (7.08 and 8.56 relative lengths, respectively). As described above, this taxon was separated from S. g. guianensis.

The G_{ST} observed in the present study was higher than that observed with allozymes in plants that exhibit a mixed mating system (Hamrick and Godt 1996), probably because autogamy is predominant in S. guianensis (26 % of outcrossing, as described above). In general, G_{ST} values in autogamous or predominantly autogamous
species are higher than those in allogamous species (Hamrick and Godt 1996; Maki et al. 2003). High variation values (30 %) were observed between groups of Mexican and South American accessions of *Stylosanthes humilis* using amplified fragment length polymorphism (AFLP) (Vander Stappen et al. 2000). The variation was even higher in AFLP studies of *S. humilis* (59 %) and *S. viscosa* (66 %) (Sawkins et al. 2001).

*Stylosanthes g. microcephala* and *S. g. guianensis* were distributed in more than one group of the STRUCTURE analysis. The genetic differentiation (GST) between the groups formed by *S. g. microcephala* was low (18 %) and, as discussed above, may be related to the soil conditions. However, the genetic differentiation among the three clusters formed mostly by the *S. g. guianensis* variety was 49 %, which is similar to the overall GST (46 %). Because most of the accessions from this variety that were included in this study were collected in the same Brazilian state, no conclusion can be drawn about the correlation between these data and the geographic origin of the accessions. More botanical and genetic studies of this variety should be conducted to determine whether it should be subdivided into multiple varieties based on the existing variation.

The accessions that were not classified into botanical varieties were randomly distributed in the genetic groups. Although groups were generally dominated by one botanical variety under both clustering methods (STRUCTURE and UPGMA), some superposition was observed. Considering these facts, the classification of individuals in each botanical variety based exclusively on the genetic groups cannot be accomplished with confidence, showing that taxonomic information is fundamental for correct classification. The consistency between the molecular analysis and botanical classification could be increased by the analysis of a larger number of microsatellite markers.

### Conclusions and forward look

This study has revealed valuable information about the relationships among a large number of *S. guianensis* accessions, showing a population structure that is generally consistent with the taxonomic classification proposed by Ferreira and Costa (1979), Brandão et al. (1985) and Costa (2006). These data show that molecular markers such as microsatellites can provide complementary information to address botanical questions at the infraspecific level. These data are important for existing germplasm conservation efforts and will help in planning new collecting trips and studies of natural populations of *S. guianensis* in its diversity centre. Moreover, information about the genetic diversity of germplasm collections is essential for their use in breeding programmes, mainly in guiding controlled crosses, as well as for the identification of natural and induced hybrids, and for monitoring the variability in subsequent generations.

### Additional information

The following additional information is available in the online version of this article –

**File 1.** Fig. 1: The bar plot obtained from the model-based ancestry analysis of the *S. guianensis* accessions.
implemented in the STRUCTURE software. The numbers indicate the sample code given in Table 1.

File 2. Fig. 2: The Roger’s genetic distance dendrogram for 150 S. guianensis accessions constructed using the UPGMA method implemented in DARwin. The colours in the dendrogram indicate the accession’s group as assigned by STRUCTURE analysis. The numbers indicate the sample code given in Table 1.

Sources of funding
The work was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant no. 2005/51010-0). M.O.S.-G. received a scholarship from FAPESP (2005/52211-9), and A.P.S. and M.L.C.V. received research fellowships from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Contributions by the authors
M.O.S.-G. performed experimental and statistical analyses and drafted the manuscript; M.I.Z. participated in statistical analysis; C.T.K. and L.C. participated in the germplasm selection and the design and implementation of the study; M.L.C.V. participated in the cytogenetic experiments and their interpretation; R.M.S.R. and A.P.S. conceived of the study and participated in its design and coordination. A.P.S. helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements
We thank the State of São Paulo Research Foundation (FAPESP) for financial support (proc 2005/51010-0) and scholarships, and the Brazilian National Council for Scientific and Technological Development (CNPq) for the research fellowships awarded to A.P.S. and M.L.C.V. We would also like to thank Dr Alexandre Franco Garcia for his help with the figures.

Conflict of interest statement
None declared.

References


