This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Use of ISSR markers to assess genetic diversity of African edible seeded Citrullus lanatus landraces

Y. Djé a,*, C.G. Tahi c, A.I. Zoro Bi a, J.-P. Baudoin b, P. Bertin c

ABSTRACT

We used the twenty primers to evaluate the genetic variability of 80 individuals belonging to four accessions of edible seeded Citrullus lanatus originated from Côte d’Ivoire. Edible seeded C. lanatus, named “egusi” or “pistachio”, had a great importance in nutrition in West Africa. Nevertheless, due to its neglected status no study to our knowledge has been devoted to its genetic variability using DNA markers. The twenty ISSR primers generated 258 bands among which 252 were polymorphic (97.67%). On the whole, the bands generated revealed three types of profile sharply distinct from each other with minor differences within each type. One profile (P1) was most frequent with 252 polymorphic bands among which 252 were polymorphic (97.67%). The pairwise genetic distance between the 80 individuals varied from 0 to 0.61. The Factorial Component Analysis and the dendrogram clearly separated the 80 individuals into three clusters corresponding to the three types of profile. The results showed that clusters were well separated from each other whereas accessions were not. Our results suggest that high number of individuals should be taken into account for sampling missions and conservation strategies because accessions were not well differentiated from each other. Local agricultural practices consisting of frequent seeds exchanges between farmers and the conservation of harvested seeds for next year culture could be one explanation.

1. Introduction

The Cucurbitaceae family is one of the most economically important families worldwide. It is cultivated in all parts of the world and is used in many processes. For the most commercial species, several modern varieties are cultivated in intensive cropping systems with high inputs and only the flesh is consumed crude or as soup. Many traditional cucurbitaceous species are reported in several West and Central Africa where they are cultivated by women on a great diversity of soil types and rainfall, annually, as minor crops intercropped with staple food. For these species locally called “pistachio” or “egusi”, only seeds are transformed into paste for consumption (Badifu, 2001; Zoro Bi et al., 2003). These species share an important source of income for rural people. Nevertheless, due to their underused status, these locally adapted landraces have received little attention from scientists and deciders. To preserve the genetic richness of these landraces, efforts must be direct towards the implementation of reliable collecting and conservation strategies. Such plant materials are useful to set up breeding programs for cash crops. According to Zoro Bi et al. (2006), five species of Cucurbitaceae are cultivated for seed consumption in Côte d’Ivoire. Citrullus lanatus L. is one of the most important of them with respect to cultivation area, production and commercial transactions volume. This species is monoecious and naturally cross-pollinated (Jaskani et al., 2005; Ellul et al., 2007) with outcrossing rate around 75% (Ferreira et al., 2008). There is a need of urgent conservation program to avoid a serious depletion of the genetic richness of oilseed C. lanatus.

Several studies have been devoted to assess genetic variability of Cucurbitaceae based on molecular markers (Katzir et al., 1996; Levi et al., 2004; Behera et al., 2008). Among a wide range of DNA markers methods available, single sequence repeats (SSR) or microsatellites tend to be the most variable for discriminating among genotypes (Russell et al., 1997; Pejic et al., 1998). Microsatellites are widely represented in the genome but require complex methods to generate reliable markers and are highly
specific. A variant of microsatellites called ISSR (Inter Simple Sequence Repeat) involves the amplification of DNA region located between two microsatellites loci (Zietkiewicz et al., 1994). ISSR method combines the advantages of RAPD markers widely dispersed in species genome with high polymorphism and reliability of microsatellite. In addition, ISSR method is suitable for less known species like African edible seeded C. lanatus. This marker has a greater robustness in repeatability and show a high reliability of microsatellite. In addition, ISSR method is suitable for less known species like African edible seeded C. lanatus. We previously showed that ISSR could be powerful to evaluate genetic diversity of African edible seeded Cucurbitaceae (Bornet and Branchard, 2001). We previously showed that ISSR could be powerful to evaluate genetic diversity of African edible seeded C. lanatus. This marker has a greater robustness in repeatability and show a high variability (Borret and Branchard, 2001). We previously showed that ISSR could be powerful to evaluate genetic diversity of African edible seeded Cucurbitaceae (Bornet and Branchard, 2001). We previously showed that ISSR could be powerful to evaluate genetic diversity of African edible seeded C. lanatus. This marker has a greater robustness in repeatability and show a high variability (Borret and Branchard, 2001).

### Table 1

List of 20 primers, melting (\(T_m\)) and annealing (\(T_a\)) temperatures optimized for PCR amplification, percentage of polymorphic bands generated per primer and the mean PIC per primer in C. lanatus.

<table>
<thead>
<tr>
<th>No</th>
<th>Primer</th>
<th>Sequence (5' → 3’)</th>
<th>(T_m) (°C)</th>
<th>(T_a) (°C)</th>
<th>Total</th>
<th>Polymorphic fragments</th>
<th>Polymorphic bands (%)</th>
<th>PIC mean (mean ± s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sola 1</td>
<td>BDB-(ACA)$_{12}$</td>
<td>50</td>
<td>51</td>
<td>15</td>
<td>15</td>
<td>100</td>
<td>0.16 ± 0.08</td>
</tr>
<tr>
<td>2</td>
<td>Sola 2</td>
<td>DD-(CGA)$_{12}$</td>
<td>56</td>
<td>55</td>
<td>10</td>
<td>9</td>
<td>90</td>
<td>0.19 ± 0.13</td>
</tr>
<tr>
<td>3</td>
<td>Sola 3</td>
<td>DBB-(GGA)$_{12}$</td>
<td>58</td>
<td>59</td>
<td>16</td>
<td>16</td>
<td>100</td>
<td>0.13 ± 0.07</td>
</tr>
<tr>
<td>4</td>
<td>Sola 4</td>
<td>VHV-(GTG)$_{12}$</td>
<td>56</td>
<td>55</td>
<td>13</td>
<td>13</td>
<td>100</td>
<td>0.22 ± 0.06</td>
</tr>
<tr>
<td>5</td>
<td>Sola 5</td>
<td>BDB-(ACA)$_{12}$</td>
<td>50</td>
<td>55</td>
<td>13</td>
<td>13</td>
<td>100</td>
<td>0.18 ± 0.13</td>
</tr>
<tr>
<td>6</td>
<td>Sola 6</td>
<td>BDB-(CAC)$_{12}$</td>
<td>60</td>
<td>59</td>
<td>15</td>
<td>15</td>
<td>100</td>
<td>0.24 ± 0.14</td>
</tr>
<tr>
<td>7</td>
<td>Sola 7</td>
<td>(AC)$_{12}$-YF</td>
<td>52</td>
<td>51</td>
<td>10</td>
<td>10</td>
<td>100</td>
<td>0.21 ± 0.06</td>
</tr>
<tr>
<td>8</td>
<td>Sola 8</td>
<td>(GA)$_{12}$-YF</td>
<td>56</td>
<td>55</td>
<td>13</td>
<td>12</td>
<td>92.3</td>
<td>0.15 ± 0.08</td>
</tr>
<tr>
<td>9</td>
<td>Sola 9</td>
<td>(AC)$_{12}$-YG</td>
<td>52</td>
<td>51</td>
<td>13</td>
<td>13</td>
<td>100</td>
<td>0.15 ± 0.13</td>
</tr>
<tr>
<td>10</td>
<td>Sola 10</td>
<td>(AC)$_{12}$-YG</td>
<td>56</td>
<td>55</td>
<td>11</td>
<td>10</td>
<td>90.9</td>
<td>0.17 ± 0.15</td>
</tr>
<tr>
<td>11</td>
<td>Sola 11</td>
<td>CAG-(CAA)$_{12}$</td>
<td>50</td>
<td>59</td>
<td>15</td>
<td>15</td>
<td>100</td>
<td>0.19 ± 0.11</td>
</tr>
<tr>
<td>12</td>
<td>Sola 12</td>
<td>CTG-(AG)$_{12}$</td>
<td>58</td>
<td>59</td>
<td>17</td>
<td>16</td>
<td>94.11</td>
<td>0.15 ± 0.09</td>
</tr>
<tr>
<td>13</td>
<td>B 1</td>
<td>(AG)$_{12}$-C</td>
<td>52</td>
<td>59</td>
<td>16</td>
<td>16</td>
<td>100</td>
<td>0.20 ± 0.17</td>
</tr>
<tr>
<td>14</td>
<td>B 3</td>
<td>(GA)$_{12}$-A</td>
<td>50</td>
<td>55</td>
<td>12</td>
<td>12</td>
<td>100</td>
<td>0.17 ± 0.14</td>
</tr>
<tr>
<td>15</td>
<td>B 4</td>
<td>(GA)$_{12}$-C</td>
<td>52</td>
<td>51</td>
<td>13</td>
<td>12</td>
<td>92.3</td>
<td>0.19 ± 0.11</td>
</tr>
<tr>
<td>16</td>
<td>B 5</td>
<td>(CA)$_{12}$-T</td>
<td>48</td>
<td>55</td>
<td>12</td>
<td>12</td>
<td>100</td>
<td>0.22 ± 0.06</td>
</tr>
<tr>
<td>17</td>
<td>B 6</td>
<td>(GAT)$_{12}$-C</td>
<td>49</td>
<td>51</td>
<td>12</td>
<td>12</td>
<td>100</td>
<td>0.16 ± 0.07</td>
</tr>
<tr>
<td>18</td>
<td>CBTC 1</td>
<td>(AC)$_{12}$-T</td>
<td>50</td>
<td>55</td>
<td>13</td>
<td>13</td>
<td>92.3</td>
<td>0.20 ± 0.14</td>
</tr>
<tr>
<td>19</td>
<td>CBTC 2</td>
<td>(AC)$_{12}$-G</td>
<td>52</td>
<td>59</td>
<td>13</td>
<td>13</td>
<td>100</td>
<td>0.15 ± 0.06</td>
</tr>
<tr>
<td>20</td>
<td>P 5</td>
<td>(TC)$_{12}$-A</td>
<td>50</td>
<td>51</td>
<td>6</td>
<td>6</td>
<td>100</td>
<td>0.19 ± 0.05</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>258</td>
<td>252</td>
<td></td>
<td>97.67</td>
<td></td>
<td>0.18 ± 0.02</td>
</tr>
</tbody>
</table>

Mean 12.9 12.6 97.67 0.18

By Djè et al. (2006). Briefly, leaf tissues were directly grounded in Eppendorf tubes in 500 μl of cethytrimethylamyl 2um NaCl, 2% PVP, 20 mM EDTA pH 8.2, 100 mM Tris–HCl pH 8.0, and 1% β-mercaptoethanol. The extracted DNA was quantified on 0.8% agarose gel with TBE 1 × buffer (10 × TBE: 108 g/l Tris base, 55 g/l boric acid, 8.3 g/l EDTA pH 8.0) and about 60 ng were used in amplification reactions.

Twenty primers (Eurogentec) were used in optimization process. Twelve and eight of them were 3'- and 5'-anchored, respectively (Table 1). Eleven of these primers had been previously optimized for African edible seeded C. lanatus landraces (Djè et al., 2006). The polymerase chain reactions (PCR) were performed in a mix containing 0.2 units per reaction of Taq DNA polymerase (Eurogentec), 1 μl of PCR buffer (Eurogentec buffer system II), 0.2 μl of 2 mM dNTPs, 0.16 μl of each ISSR primer, 60 ng of extracted DNA. PCR cycling conditions were as follows: 10 min initial denaturation at 95 °C, 36 cycles of amplification [30 s at 95 °C, 45 s at the annealing temperature (\(T_a\)), 2 min elongation at 72 °C] and final elongation of 5 min at 95 °C. PCR reactions were performed with a thermal cycler PTC-100™ Gradient 96 (MJ Research, Inc.). To determine the optimal annealing temperature (\(T_a\)), an interval of 10 °C around the melting temperature (\(T_m\)) was tested. Temperatures leading to clear patterns were then repeated until the best \(T_a\) was selected for routine procedure. The reproducibility of the technique was tested by carrying out two independent PCR with one positive control sample as well as one negative control.

The result of each amplification reaction was analysed on 1.5% agarose gel (Molecular grade II from Eurogentec) in 1 × TBE (Tris Borate EDTA) buffer. A smartladder provided by Eurogentec was used as standard marker to estimate sizes of PCR products by migration distance comparison (200–10,000 bp). The gels were stained by soaking 1.4 μl of ethidium bromide in 100 ml of agarose. Electrophoresis was run for 3 h at 90 V and DNA was visualised with an UV transilluminator and analysed with a video image analyser (Biocapt, Vilbert-Lourmat, Marne-la-Vallée, France).
2.3. Data collection and analysis

Only clearly and reproducible bands were scored. The percentage of reproducibility was determined by dividing the number of reproducible bands by the total number of bands observed (McGregor et al., 2000). The gels were scored for the presence (1) and absence (0) of amplification fragments using the computer program Gene Profiler V4.04 for Windows (Scanalytics, Inc.). Binary data from different primers were used to perform the analysis. The polymorphic information content (PIC) of a band was calculated according to the formula:

$$\text{PIC} = \sum [1 - (p_i^2 + a_i^2)]$$

where \( p_i \) and \( a_i \) are the presence and absence frequencies of the \( i \)th band, respectively.

A factorial correspondence analysis (FCA) based on ISSR markers were performed using the SAS statistical program. The analysis of the molecular variance AMOVA at two levels (within and between the accessions) was performed using the GenAlEx6 software (Peakall and Smouse, 2006) from the molecular data provided by ISSR markers. The genetic distances (Nei and Li, 1979) between individuals were calculated and a dendrogram was generated using the UPGMA (unweighted pair-group method with arithmetic averaging) cluster analysis with Treecon software (Van de Peer and De Wachter, 1994). Bootstrap analysis was performed with a set of 20 primers to test the reliability of the tree obtained.

3. Results

3.1. ISSR markers optimized and polymorphism detected

For all of the 20 ISSR markers tested, we obtained clear and reproducible bands as showed for primer Sola 4 (Fig. 1). Nine optimal annealing temperatures \( T_a \) were lower and eleven higher than \( T_m \) (one degree below and from one to nine degree above \( T_m \), respectively). These temperatures could be grouped in three groups (51 °C, 55 °C and 59 °C). The 20 ISSR primers amplified a total of 258 bands resulting in an average of 12.9 bands per primer. The lowest number of bands was observed with primer P 5 (6) which was the only primer amplifying less than 10 bands. The highest number of bands was observed for primer Sola 12 (17). The molecular weight of the bands varied from 192 pb (Sola 1) to 2015 pb (Sola 9). Of the 258 bands amplified, 252 showed polymorphism corresponding to 97.67% of polymorphic bands. The average number of polymorphic bands per primer was 12.6. The Polymorphic Information Content (PIC) per primer ranged from 0.13 (Sola 3) to 0.24 (Sola 6), with a mean of 0.18 (Table 1).

3.2. Genetic diversity

For the entire experiment on 80 individuals from the four different accessions analysed with 20 primers, three types of profiles were clearly highlighted (Fig. 2). Minor differences were revealed inside each type of profile. Three accessions (NI084, NI127 and NI145) exhibited the three types of profiles, while the accession NI076 presented only one type of profile (P1). The first two axis of the factorial component analysis (FCA) explained 59% of the total variation (38.29 and 20.71% of the variation, respectively.)
Fig. 4. UPGMA dendrogram using Nei (Nei & Li, 1979) distance of ISSR band profiles from samples of 80 C. lanatus individuals. The scale bar on the top of figure represents the pairwise genetic distances between individuals.
for axis 1 and 2). Three main clusters could be distinguished in FCA (Fig. 3). The first component separated clusters 2 and 3 from cluster 1. The second axis clearly separated the three groups from each other. The first cluster contained 65 individuals belonging to the four accessions, whereas clusters 2 and 3 contained 9 and 6 individuals respectively, belonging to 3 accessions only. The dendrogram UPGMA (Fig. 4) divided the 80 individuals into 3 main clusters that clearly corresponded to the structure obtained with the FCA. The genetic structure inside the three clusters revealed a sub-clustering by accession: individuals from same accession were grouped into one or a few number of sub-clusters. This is particularly true for clusters 2 and 3 for which no superimposing was observed between individuals from different accessions. We found 83% of the variation between clusters using AMOVA ($F_{st} = 0.83$, $P$-value < 0.1), indicating that clusters were highly different from each other. The genetic distances between clusters were $0.507 \pm 0.048$ (clusters 1 and 2), $0.627 \pm 0.032$ (clusters 1 and 3) and $0.617 \pm 0.040$ (clusters 2 and 3). In contrast, the pairwise genetic diversity within clusters was fairly weak ($0.000 \pm 0.024$ for cluster 1, $0.193 \pm 0.059$ for cluster 2 and $0.218 \pm 0.060$ for cluster 3). The molecular analysis of variance of accessions indicated that 9 and 91% of the observed variation was between and within the accessions respectively. This value indicated that the accessions were little distinct from each other, but high variations were observed within them. $F_{st}$ was very weak ($0.092$), confirming a very weak distance between accessions ($P = 0.002$). The pairwise average genetic distance was calculated within and between accessions. The mean distance within accession (between individuals of the same accession) was very weak for accession NI076 ($0.055 \pm 0.017$) compared to others ($0.246-0.275$). The distances between accessions were identical and were the same order of magnitude as those related to individuals within accessions NI084, NI127 and NI145. Thus, this result confirmed those from FCA and AMOVA. The four accessions could be distinguished in two groups. The first were more diverse with individuals dispersed into three clusters and relatively medium genetic diversity (NI084, NI127 and NI145). The second group was represented by the only accession NI076 for which all individuals fell into cluster P1 with the lowest genetic diversity.

4. Discussion

Annealing temperature had a great influence in the optimization of PCR-based genetic markers (Bornet and Branchard, 2001). Some studies used a unique and low annealing temperature ($T_a$) for ISSR amplifications with different primers (Sanchez de la Hoz et al., 1996). Some others used primer-specific $T_a$ always higher than $T_m$ (Bornet and Branchard, 2001). In our case, $T_a$ was primer-specific, below or above $T_m$. The three groups of $T_a$ generated could facilitate routine PCR amplification procedure. To obtain a reliable result from genetic analyses, a high number of primers represent an important advantage (Verma et al., 2007; Behera et al., 2008). Twenty ISSR primers are now available to examine edible seeded C. lanatus genetic variability. Wang et al. (2005) founded high number of alleles by locus in Chinese watermelon (198 alleles for 8 primers) but only 120 of them showed polymorphism (60.60%). In the present study, the percent of polymorphic bands (97.67%) clearly indicated that the ISSR markers are highly polymorphic and especially informative for estimating genetic relationships. High percentage of polymorphism based on ISSR markers was obtained by Paris et al. (2003) in Cucurbita pepo (74%); Behera et al. (2008) in Momordica charantia and Verma et al. (2007) in Benincasa hispida (80%). These studies demonstrated that ISSR markers frequently detected a higher level of polymorphism than that detected with other dominant markers (RAPD or AFLP) among Citrullus or other Cucurbitaceae (Paris et al., 2003; Senseny et al., 2007; Behera et al., 2008). Stepansky et al. (1999) showed in Cucumis melo L. the same trend i.e. a low polymorphism with RAPD (6.9 bands/alleles and 70% of polymorphic alleles) than with ISSR (9 bands/alleles and 90% of polymorphic alleles).

The different analyses carried out showed a low level of structure inside accession zone whereas the level of structuring by cluster or profiles was high. Profiles were thus more significant than accessions in term of genetic diversity. The low level of genetic diversity observed in the present study, notably in cluster 1 and particularly in the accession NI076 could be explained in several ways. The genetic basis of the cultivated C. lanatus is known to be narrow (Maggs-Kölling et al., 2000; Levi et al., 2004). The accessions used in this study did not escape to previous general conclusions. One hypothesis which could explain the low level of diversity would be the origin of the cultivation. The entire set of accessions of C. lanatus could be the resultant of one single entry point and then gradually by migration, the seeds were widespread in the entire country. Similar observations were reported by Romão (2000) from a study related to the evolution of watermelon after its introduction in Northeast Brazil by African slaves. The factors like genetic drift, sampling effect or bottleneck can cause big losses of genetic variation for small populations (Ellstrand and Elam, 1993) like C. lanatus populations dispersed from women stocks and fields in Côte d’Ivoire. Another hypothesis could be attributed to peasant’s agricultural practices. Two approaches are currently observed for the treatment of harvested seeds. In one hand, the peasants preserve a part of their harvest for the next sowing, thus, possibly losing part of their diversity at each cultivation period. The less variable accession NI076 could be a result of many self-reproduction of the same seeds stock from several generations. In the second hand, women currently take a part of reserve of their parents or neighbours or at the market in order to establish a new culture or to restore losing seeds of precedent harvests. Thus, it may happen that accessions NI084, NI127 and NI145 for which individuals failed into three clusters and had medium genetic diversity resulted from frequent seeds exchange between women.

5. Conclusion and perspective

For underutilized species like edible seeded C. lanatus, the optimization of any new technology could be powerful for management. The ISSR marker system was quick and highly reproducible. It generated sufficient polymorphism to have large-scale DNA fingerprinting purposes. This study enabled us to find twenty ISSR primers, which could be used directly for the further molecular studies on this species.

The genetic structure observed was characterised by three clusters that did not correspond to accessions analysed. To confirm such a pattern and then use reliability the results for genetic resources collection and management purposes, further investigation using more accessions are needed.

Acknowledgments

This research was financed by the Direction Générale de la Coopération au Développement (DGD, Brussels, Belgium) and supervised by the Comité Interuniversitaire pour le Développement (CUD, Brussels, Belgium). Thank to Willy Dussard for his technical assistance.

References


