DIVERSITY AND GENETIC RELATEDNESS AMONG GENOTYPES OF *Vitis* spp. USING MICROSATELLITE MOLECULAR MARKERS

PATRÍCIA COELHO DE SOUZA LEÃO, COSME DAMIÃO CRUZ, SÉRGIO YOSHIMITSU MOTOIKE

**ABSTRACT** - The purpose of this research was to study the genetic diversity and genetic relatedness of 60 genotypes of grapevines derived from the Germplasm Bank of Embrapa Semiárido, Juazeiro, BA, Brazil. Seven previously characterized microsatellite markers were used: VVS2, VVMD5, VVMD7, VVMD27, VVMD3, ssrVrZAG79 and ssrVrZAG62. The expected heterozygosity (H_e) and polymorphic information content (PIC) were calculated, and the cluster analysis were processed to generate a dendrogram using the algorithm UPGMA. The H_e ranged from 81.8% to 88.1%, with a mean of 84.8%. The loci VrZAG79 and VVMD7 were the most informative, with a PIC of 87 and 86%, respectively, while VrZAG62 was the least informative, with a PIC value of 80%. Cluster analysis by UPGMA method allowed separation of the genotypes according to their genealogy and identification of possible parentage for the cultivars ‘Dominga’, ‘Isaura’, ‘CG 26916’, ‘CG28467’ and ‘Roni Redi’.

**Index terms:** Grape, multivariate analysis, genealogy, SSR.

DIVERSIDADE E RELAÇÕES GENÉTICAS ENTRE GENÓTIPOS DE *Vitis* spp. UTILIZANDO MARCADORES MOLECULARES MICROSSATÉLITES

**RESUMO** - O presente trabalho teve como objetivo estudar a diversidade genética e as relações de parentesco de 60 genótipos de videira procedentes do Banco de Germoplasma da Embrapa Semiárido, Juazeiro-BA. Sete marcadores microsatélites previamente caracterizados foram utilizados: VVS2, VVMD5, VVMD7, VVMD27, VVMD3, ssrVrZAG79 e ssrVrZAG62. Foram calculadas a heterozigosidade esperada (H_e), conteúdo de informação polimórfica (PIC), e as análises de agrupamento foram processadas para gerar um dendograma, utilizando-se do algoritmo UPGMA. A H_e variou de 81,8 a 88,1%, com média de 84,8%. Os lócus VrZAG79 e VVMD7 foram os mais informativos, com PIC de 87 e 86%, respectivamente, enquanto VrZAG62 foi o menos informativo, com um valor de PIC de 80%. A análise de agrupamento pelo método UPGMA permitiu separar os genótipos de acordo com sua genealogia e identificar possíveis parentais para as cultivares ‘Dominga’, ‘Isaura’, ‘CG 26916’, ‘CG28467’ e ‘Roni Redi’.

**Termos para indexação:** Uva, análise multivariada, genealogia, SSR.
INTRODUCTION

The development of new grapevine cultivars adapted to Brazilian tropical and subtropical conditions has been one of the main objectives of genetic breeding programs in the country. At EMBRAPA Semiárido, grapevine breeding studies are in an initial phase, and, in most recent decades, research efforts have been concentrated on morpho-agronomic enrichment and characterization of the genotypes of the Germplasm Bank (BORGES et al. 2008; LEÃO et al., 2010; LEÃO et al., 2011). This collection has begun in 1965 with cultivars collected in the Northeast region and later expanded with others imported from FAO in Italy; Instituto Agronomico do Campinas, in São Paulo; and Embrapa Uva e Vinho, in Rio Grande do Sul. It currently has 261 accessions, with most of the species being *Vitis vinifera* L..

Most of the grapevine cultivars are ancient and were derived from different processes, such as differentiation from wild grapevines, spontaneous crosses between wild grapevines and cultivars or crosses between cultivars. Ease of asexual propagation gave rise to an estimated number of 14,000 cultivars, for different purposes: fresh consumption, raisins, juices and wine, with the number of grapevine cultivars maintained in germplasm collections being estimated at around 10,000 (ALLEWELDT and DETTWEILER, 1994). Older wild cultivars involved in natural crosses were not able to be identified since they no longer exist naturally, but many parent species are still cultivated or are preserved in collections. Through the use of molecular methods, these parent cultivars and their descendents may be recognized, allowing the study of their geographic origin and evolutionary history.

The use of microsatellite markers has allowed reconstruction of pedigrees of innumerable economically important grapevine cultivars, such as the discoveries that ‘Cabernet Sauvignon’ is a descendent of ‘Cabernet Franc’ and ‘Sauvignon Blanc’ (BOWERS and MEREDITH, 1997), ‘Durif’, known in California as ‘Petit Syrah’ (MEREDITH et al., 1999), resulted from the crossing of ‘Peloursin’ X ‘Syrah’, and ‘Muscat de Hamburgo’ is a descendent of ‘Moscato of Alexandria’ and ‘Schiava Grossa’ (CRESPLAN, 2003).

The purpose of this paper was to study the genetic diversity and genetic relatedness of 60 genotypes belonging to the grapevine Germplasm Bank of Embrapa Semiárido in Petrolina, PE.

MATERIAL AND METHODS

**Plant matter**

A group of 60 grapevine accessions were assessed, including 33 hybrids, 22 accessions of the species *Vitis vinifera* L., and five accessions of unknown origin. The accessions are part of the Germplasm Bank of Embrapa Semiárido, located in Juazeiro, state of Bahia, Brazil (9°24′S, 40°26′W, and 365.5m of altitude) and they are selected based on their importance for grape breeding and on previous studies (LEÃO et al., 2010; 2011).

**DNA extraction and amplification conditions**

Genomic DNA was extracted from young expanded leaves collected from the apex of shoots, as described by Lodhi et al. (1994). Vines were 8 years at the time to collect the samples. The leaves were homogenized by means of a mechanical homogenizer Homes 6 (Bioreba, Longmont, CO). In the final stage, the pellets were suspended in a 1X Tris EDTA buffer solution and stored at -20°C. The integrity of the DNA was verified in 0.8% agarose gel stained with ethidium bromide.

Seven previously characterized microsatellite markers were used: VVS2 (Thomas e Scott, 1993), VVMD5, VVMD7, VVMD27, VVMD31 (BOWERS et al. 1996; BOWERS et al. 1999), ssrVrZAG79 and ssrVrZAG62 (SEFC et al. 1999). They were chosen as being most recommended for characterization of cultivars in the grapevine germplasm collections (THIS et al., 2004). One of the primers of each pair was labeled at the 5’ end with fluorescent dyes such as 6-FAM (blue), HEX (yellow) or NED (green) (Applied Biosystems).

Amplifications were performed in a final reaction volume of 10 µl containing 2.5 ng/µl DNA, 10 pmoles of each primer, 2.5 mM of deoxynucleotides mix (Applied Biosystems); 1 µL of Gold 10X buffer (Applied Biosystems); 2mM MgCl₂ (Applied Biosystems) and 0.5 units of AmpliTaq Gold DNA polymerase (Applied Biosystems).

The PCR (polymerase chain reaction) was performed in a thermocycler PTC-100 (MJ Researcher) using a single program for all the primers. The amplification program consisted of an initial stage of 5 minutes at 95°C, followed by 35 cycles of 30 seconds at 95°C, 45 seconds at 60°C and 1 minute at 72°C and a final extension stage of 7 minutes at 72°C.

The amplifications were confirmed in 2% agarose gel stained with ethidium bromide (10 mg/mL). The PCR products were denatured at 94°C for
2 minutes and then applied in 6% acrylamide gel, performing electrophoresis in a DNA sequencer model ABI 377 (PE/Applied Biosystems).

The sizes of the alleles were assessed based on comparison with the known internal size standard Genescan Rox 500 (Applied Biosystems) and four grapevine cultivars (Carignane, Riesling, Thompson Seedless and Chardonnay), whose allelic profiles are known as database of Foundation Plant Service/USDA and University of California at Davis, to allow comparison of the allelic size profiles obtained. The individuals were genotyped using the software GeneScan™ version 3.1 and Genotyper™ version 2.5.2. (PE/Applied Biosystems).

Statistical analyses
The expected heterozygosity ($H_e$) was calculated according to the formula: $H_e = 1 - \sum p_i^2$, where $p_i$ is the frequency of the allele $i$ for the locus studied (Nei, 1987). The polymorphic information content (PIC) was calculated as $1 - \sum p_i^2 - \sum 2 p_i p_j$, where $p_i$ is equal to the frequency of the $i$th allele and $p_j$ is the frequency of the $(i + 1)$th allele (Botstein et al., 1980). Dissimilarity matrices were obtained using the arithmetic complement of the weighted index given by:

$$D_{ij} = 1 - \frac{1}{2} \sum_{l=1}^{L} p_i c_l$$

in which: $L =$ total number of loci; $C_l$ = number of common alleles between the pairs of accessions $i$ and $i'$, and $p_i = a_i / A$, in which: $a_i =$ total number of alleles of the locus $j$; $A =$ total number of alleles. 

Cluster analyses were processed to generate a dendrogram using the algorithm UPGMA (unweighted pair group method with arithmetic mean). 

Statistical analyses were performed on the Genes version 2007.0.0 computer program, developed in the Bioinformatics/Bioagro laboratory of the Universidade Federal de Viçosa, Minas Gerais, Brazil.

RESULTS AND DISCUSSION

Sixty accessions derived from the grapevine germplasm collection of Embrapa Semiárido in Juazeiro, state of Bahia, Brazil, were genotyped for the first time using seven reference microsatellite markers (This et al., 2004). This set of accessions is represented by cultivars and selections derived from different breeding programs. The largest group is composed by Brazilian cultivars from the breeding programs of Instituto Agronômico de Campinas-IAC (‘A Dona’, ‘Aurora’ or ‘IAC 77526’, ‘Isaura’, ‘Juliana’, ‘Patricia’, ‘Paulistinha’ and the root stock cultivars ‘IAC 313’ and ‘IAC 766’) and Embrapa Uva e Vinho (‘Moscato Embrapa’, ‘BRS Rubea’, ‘BRS Clara’, ‘BRS Linda’, ‘BRS Morena’ and ‘BRS Lorena’).

All the loci used simplified in a satisfactory manner in almost all the cultivars and were multiallelic. Eighty-nine alleles were obtained, ranging from 11 (VVS2) to 14 (VVMD31 and VVMD7), with an average and effective number of 12.7 alleles per locus (Table 1). This value is higher than those already mentioned by other authors working with grapevine accessions of diverse geographic origins. Almadanim et al. (2007) obtained a mean of 8.17 alleles per locus, Vouillamoz et al. (2006) observed a mean value of 11.9 alleles per locus and Martín et al. (2003) found 11 alleles per locus. Lamboy and Alpha (1998), analyzing the diversity of 110 accessions belonging to 21 species of Vitis and 4 hybrids found 24.4 alleles per locus, a greater quantity than that observed in this study.

The most frequent alleles in each locus were VVMD5 – 238 (25.8%), VVMD7 – 239 (21.6%), VVMD27 – 185 (26.7%), VVMD31 – 212 (29.2%), VVS2 -133 (20.0%), VrZAG62 – 189 (35.0%) and VrZAG79 – 255 (21.7%) (Figure 2). Forty-three alleles (48.3%) were rare, with frequencies of less than 5%. The frequency of rare alleles is in agreement with that observed by Martín et al., (2003). On the other hand, it was less than the quantity of 73% of rare alleles mentioned by Lamboy and Alpha (1998) upon analyzing a group of 110 taxonomically broader accessions.

Expected heterozygosity ($H_e$) ranged from 81.8% (VrZAG62) to 88.1% (VrZAG79), with a mean of 84.8% (Table 1). This value was greater than other $H_e$ results already obtained by Martinez et al. (2006) (81%), Martín et al. (2003) (80.6%) and Lamboy and Alpha (1998) (62.5%).

The high number of heterozygotes obtained in this study may be explained by the predominance of hybrids, with many interspecific hybrids among them, combining alleles of different species of Vitis. High heterozygosity of the cultivated grapevine is a consequence of selection throughout the domestication process and growing in favor of hybrid vigor which confers desirable characteristics such as high production and larger size of bunches and berries (Lamboy and Alpha, 1998).

The values obtained for polymorphic
information content were less than those of expected heterozygosity. All the loci may be considered highly informative (PIC > 70%). The loci VrZAG79 and VVMD7 were the most informative, with a PIC of 87.1 and 86.1%, respectively, while VrZAG62 was the least informative, with a PIC value of 80.1% (Table 1). In contrast, the locus VrZAG62 was the most informative (PIC=88%) according to Martínez et al., (2006) and VVMD7 was the least informative (PIC=70.5%), according to Martin et al. (2003).

The allelic combinations of the seven loci were sufficient to differentiate 55 genotypes. Correspondences were observed between the allelic profiles of the accessions ‘Aurora’ and ‘IAC 77526’, ‘A Dona’ and ‘Himoront’, ‘Júpiter’ and ‘CG 26858’, ‘Ferlongo’ and ‘Moscatel Nazareno’, and ‘A1581’ and ‘A1105’ (Table 2). Among them, the denomination ‘IAC 77526’ designates the number of selection for the ‘Aurora’ cultivar, and therefore they constitute a case of synonymy. The other genotypes constitute distinct cultivars and they may therefore be mistakenly identified in the collection, with the need for repeating the analyses using a greater number of microsatellite molecular markers.

Genetic similarities were observed which ranged from 0 to 100% among the accessions assessed (Figure 1), with a mean similarity of 22% of common alleles, below the values mentioned in previous studies by other authors. Lopes et al. (1999; 2006), studying the diversity of the grapevine grown in Portugal, obtained 36% and 38% of common alleles respectively, while Seke et al. (1998) reported 40% and 43% of common alleles in the group of cultivars of V. vinifera and root stock from Austria respectively.

The formation of a large group with 52 accessions may be observed on the dendrogram (Figure 1) which was separated at the level of 57% of genetic dissimilarity from the cultivars Lake Emerald, Moscato Noir, ‘IAC 313’ and ‘IAC 766’ root stock, and from the group composed by the accessions ‘Blue Lake’, ‘Seyve Villard 12327’, ‘Seyve Villard 12375’ and ‘Regner’. The cultivar of German origin ‘Regner’ has an allele in common in all the loci with ‘Seyve Villard 12375’, showing a parent/progeny relationship between them. This shows that the denomination of this accession is incorrect in this collection because it did not correspond to the Regner cultivar (Vitis vinifera L.) whose pedigree is ‘Lugliena Bianca’ X ‘Gamay Precoce’. The root stock from Instituto Agronômico de Campinas exhibited the greatest diversity in relation to the other grapevine accessions (91%) and a genetic similarity between them of 37%.

Fatahi et al. (2003) also observed the formation of a group which gathered root stock from four different American species.

When a dissimilarity of 50% in the dendrogram was considered, five groups were formed and diverse accessions were isolated, not being included in any group.

‘Aurora’ and ‘IAC 77526’ exhibited the same allelic profile and formed the first group, with a mean dissimilarity of 84% in relation to the other accessions.

The second group was composed of seven selections from the INTA, Argentina and by the cultivars Júpiter, Angelo Pirovano, BR Sinda, Marrooo Seedless, Dominga, Marengo Pirovano, Paulistinha and Sovrano Pirovano. ‘Dominga’, a cultivar of Spanish origin, whose pedigree was unknown, exhibited a genetic similarity of 65% with the cultivar Marengo Pirovano and 42% with ‘Sovrano Pirovano’, sharing an allele in at least six markers, which may show a possible parent/progeny relationship between ‘Delizia de Vaprio’ and these three cultivars.

The third group was composed of six accessions, among them, two hybrids from the University of Arkansas, ‘Reliance’ and ‘Mars’, and two Brazilian hybrids, ‘Patricia’ and ‘BRS Lorena’. The accession called ‘Roni Redi’, also belonging to this group, has an unknown origin and pedigree, but shared an allele in six loci with the cultivar Beni Fugi, showing possible common parentage, and it also exhibited high genetic similarity with the American hybrids ‘Bordo’, ‘Mars’, ‘Reliance’ and ‘Vênus’, and it is possible that ‘Ontario’ is present in its genetic background.

The fourth group included only two accessions: ‘California’ and the hybrid ‘Juliana’, developed by the Instituto Agronômico de Campinas, with a 51% similarity between them.

The fifth group was made up of nineteen accessions, which may be separated into seven distinct subgroups.

‘A1105’, ‘A1581’, ‘A1118’ and ‘Saturn’ formed a subgroup; they have at least one allele in common in the seven loci studied, showing a genetic relatedness. The selections ‘A1581’ and ‘A1105’ exhibited the same allelic profile; both are seedless grapes that may be distinguished in the field by the color of their berries: ‘A1105’ has white berries and ‘A1581’ has dark berries. In Spain, the cultivars ‘Carrasquin’ and ‘Prieto Picudo Tinto’ also exhibited identical genotypes, in spite of having significant morphological and isoenzymatic differences (MARTÍN et al., 2003) with the authors confirming
that a greater number of microsatellite markers must be added to distinguish the genotypes.

Another subgroup was composed by hybrids from Instituto Agronômico de Campinas: ‘Isaura’, ‘Muscat Saint Vallier’ and ‘Beni Fugi’. ‘Isaura’, whose pedigree is unknown, and ‘Muscat Saint Vallier’ shared at least one allele in all seven loci, indicating genetic relatedness between them.

‘A Dona’, ‘Himoront’, ‘BRS Clara’ and ‘BRS Rubea’ formed a subgroup. ‘A Dona’ and ‘Himoront’ exhibited the same allelic profile, proving to be the same genotype. Genetic relatedness was observed between the cultivars Perlona and ‘A Dona’, with a genetic similarity of 49%; they have the cultivar Muscat of Alexandria in their genetic background. The high genetic similarity between the cultivars ‘BRS Clara’ and ‘BRS Rubea’, developed by Embrapa Uva e Vinho, was not expected as they had distinct genealogies (CAMARGO et al., 2010), indicating an error in denoting ‘BRS Rubea’.

The accessions ‘Damarim’, ‘Dacari’, ‘Moscatuel’, ‘Emperatriz’ and ‘Baviera’ were confirmed as synonyms for the selections ‘CG 40016’, ‘CG 102024’, ‘CG 102295’, ‘CG 28467’ and ‘CG26916’ respectively, developed by Gargiulo in the INTA, Argentina. The selections of Gargiulo were divided into two groups; they have the cultivar Thompson Seedless in common in their genetic background. The selections ‘CG26916’, ‘CG28467’ and ‘CG87908’, exhibited an allele in common with the cultivar Thompson Seedless in at least six loci, showing a parent/progeny relationship between them.

The accessions ‘Vênus’, ‘Feal’, ‘CG 351’, ‘Perlona’, ‘Donia Maria’, ‘Tampa’, ‘Moscat Embrapa’, ‘BRS Morena’ and ‘Neptune’ were not included in any group, when 50% of genetic dissimilarity was considered as the cutoff point.

A separation of the accessions according to their geographic origin, or even according to their use, was not observed: table grapes or wine production, or morphological traits, as, for example, the color of the berries. Martín et al. (2003), studying 272 grapevine accessions from the Iberian Peninsula in Portugal and Spain, did not find a separation according to the region of origin. Martinez et al. (2006) evaluating the diversity of ‘Criollas’ grapevines from Peru and Argentina, also did not observe separation of groups according to the country of origin. Nevertheless, Dangl et al. (2001) obtained separation of cultivars of V. vinifera from the USDA Germplasm Repository according to their geographic origin. In this study, genealogy was the primordial factor for formation of groups, grouping accessions with similar genetic backgrounds, as may be observed in the group that joined the accessions that had ‘Thompson Seedless’ as a parent in the first or second generation. These results obtained allow better understanding regarding the diversity, origin and genetic relatedness among the accessions, providing subsidies to grapevine genetic breeding programs in Brazil, as well as correct identification of the cultivars for the purposes of genetic certification, traceability and legal protection of new cultivars.

### TABLE 1- Genetic parameters of seven SSR loci evaluated in 60 grapevine accessions.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Alleles Number</th>
<th>He1(%)</th>
<th>PIC2(%)</th>
<th>F maximum 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVMD5</td>
<td>12</td>
<td>83.9</td>
<td>81.9</td>
<td>25.8</td>
</tr>
<tr>
<td>VVMD7</td>
<td>14</td>
<td>87.4</td>
<td>86.1</td>
<td>21.6</td>
</tr>
<tr>
<td>VVMD27</td>
<td>13</td>
<td>83.6</td>
<td>81.7</td>
<td>26.7</td>
</tr>
<tr>
<td>VVMD31</td>
<td>14</td>
<td>83.3</td>
<td>81.4</td>
<td>29.2</td>
</tr>
<tr>
<td>VVS2</td>
<td>11</td>
<td>85.9</td>
<td>84.3</td>
<td>20.0</td>
</tr>
<tr>
<td>VrZAG62</td>
<td>12</td>
<td>81.8</td>
<td>80.1</td>
<td>35.0</td>
</tr>
<tr>
<td>VrZAG79</td>
<td>13</td>
<td>88.1</td>
<td>87.1</td>
<td>21.7</td>
</tr>
<tr>
<td>Média</td>
<td>12,7</td>
<td>84.8</td>
<td>83.2</td>
<td></td>
</tr>
</tbody>
</table>

1 Expected heterozygosity; 2 polymorphic information content; 3 Maximum frequency of alleles.
FIGURE 1 - Dendogram representing the clustering of 60 grapevine accessions by UPGMA method using seven microsatellite markers.
FIGURE 2 - Frequency of alleles of seven microsatellite loci in 60 accessions of grapevine.
**CONCLUSION**

1- The study of genetic diversity by means of cluster analysis by the UPGMA method allowed separation of the genotypes according to their genealogy, facilitating understanding of the genetic relatedness among them;

2- Microsatellite markers allowed identification of possible parentage for the cultivars ‘Dominga’, ‘Isaura’, ‘CG 26916’, ‘CG28467’ and ‘Roni Redi’, whose pedigrees are not available in the literature and databases consulted;


**REFERENCES**


