



## Molecular characterization of a corn core collection–landraces subgroup

### *Caracterização molecular de acessos de uma coleção núcleo de milho, subgrupo landraces*

Ronaldo Rodrigues Coimbra<sup>1</sup>, Glauco Vieira Miranda<sup>2</sup>, Newton Portilho Carneiro<sup>3</sup>, Cláudia Teixeira Guimarães<sup>3</sup>, Derly José Henriques da Silva<sup>2</sup>, Elaine Cristina Alves Martins Oliveira<sup>4\*</sup>

**Abstract:** The efficient use of genetic resources- stored in germplasm collections can be maximized if morpho-agronomic and molecular information on the accessions is made available. To achieve this, a collection that is well-structured, well-curated and easily accessible (the core collection) is required. Consequently, the objective of the current study was to characterize 80 landrace accessions from the maize core collection of the Federal University of Viçosa (UFV), and assay thenn genetic diversity of the various landraces, considering grain type and ecogeographic origin. For this, AFLP analysis was performed using 12 *primer* combinations. Genetic diversity of the collection was quantified with the UPGMA method, using the Jaccard Index to quantify dissimilarity. The core collection was divided into four sub-populations by grain type, and into six sub-populations based on ecogeographic origin. Genetic diversity analysis was performed both within and between sub-populations. A high level of genetic variability was found among the landrace accessions of UFV Core Collection, principally among those accessions with dentate type grains. Classification by grain type and ecogeographic origin allowed genetically divergent groups to be distinguished.

**Key words:** Diversity. Genetic Resources. AFLP. *Zea mays* L.

**Resumo:** A utilização eficiente de recursos genéticos, armazenados em coleções de germoplasma, pode ser maximizada se informações morfo-agronômicas e moleculares sobre os acessos forem disponibilizadas. Para isso, é necessária a existência de uma coleção racionalizada, refinada e estruturada, chamada de coleção núcleo. Assim, objetivou-se com este trabalho caracterizar molecularmente 80 acessos landraces da coleção núcleo de milho da Universidade Federal de Viçosa (UFV) e verificar a diversidade genética com relação à estratificação, considerando tipo de grão e origem ecogeográfica. Para isso, foi realizada análise de AFLP, utilizando-se 12 combinações de *primers*. A diversidade genética da coleção foi quantificada utilizando o método de UPGMA, tendo como medida de dissimilaridade o complemento do índice de Jaccard. A coleção núcleo foi estratificada em quatro sub-populações quanto ao tipo de grãos e seis quanto à origem ecogeográfica. Foi realizada análise de diversidade genética tanto entre quanto dentro das sub-populações. Foi verificada grande variabilidade genética entre os acessos landraces da Coleção Núcleo da UFV, principalmente entre os acessos com grãos do tipo dentado, e que a estratificação quanto ao tipo de grãos e origem ecogeográfica foi eficiente para a obtenção de grupos geneticamente divergentes.

**Palavras-chave:** Diversidade. Recursos Genéticos. AFLP. *Zea mays* L.

\*Corresponding author

Submitted for publication on 30/10/2017 and approved 02/12/2017

<sup>1</sup>Lecturer, Universidade Federal do Tocantins, Campus de Porto Nacional, Rua 03, Quadra 17, Lote 11, s/n° Setor Jardim dos Ipês, Porto Nacional, Tocantins, Brazil, ronaldo.rc@mail.uft.edu.br;

<sup>2</sup>Lecturer, Universidade Federal de Viçosa, Avenida P.H Rolfs, S/N, Campus Universitário, Viçosa, Minas Gerais, Brazil, glauco@gmail.com; derly@mail.ufv.br;

<sup>3</sup>Researcher, Embrapa Milho e Sorgo, Rodovia MG 424 Km 45, Sete Lagoas, Minas Gerais, Brazil, newton.caneiro@embrapa.com; claudia.guimaraes@embrapa.com;

<sup>4</sup>Lecturer, Universidade Federal do Tocantins, Rua Badejós, Lote 7, Chácara 69/72, Zona Rural, Campus de Gurupi, UFT, Tocantins, Brazil: biocris@mail.uft.edu.br

## INTRODUCTION

There are more than 6.1 million plant accessions preserved in some 1320 germplasm banks in the 157 countries that are members of the FAO's Plant Genetic Resources Commission (YETSIR *et al.*, 2008). Of the 47,000 accessions of corn stored worldwide in germplasm banks, around 27,000 make up the central corn germplasm collection, located at the International Corn and Wheat Improvement Center (CIMMYT), Mexico (YAN *et al.*, 2009).

In Brazil, there are more than 200,000 *ex situ* plant accessions preserved in some 160 germplasm banks at around 50 locations (VASCONSELOS *et al.*, 2007). There are also collections that aim to bring together the greatest representation of a species' genetic variability in a limited set of accessions. These are termed Core Collections. There are already core collections for several species, such as the rice core collection, comprising 550 accessions (5.6% of the base collection), 308 of which are traditional varieties, 94 improved lines/cultivars and 148 introduced cultivars/lineages (ABADIE *et al.*, 2005), and the Federal University of Viçosa (UFV) maize core collection, which has 104 accessions, 80 landraces, 12 improved accessions and 12 introduced from other countries. This collection was created from the active corn collection of Embrapa Corn and Sorghum, which, according to Netto *et al.* (2004), consisted of more than 3,500 accessions.

Generally, when creating a core collection, morphological and agronomic characteristics are considered, so that the full extent and nature of character variation in the core collection accessions can be compared with those in the base collection. However, due to environmental effects, the information on genetic diversity is not fully inclusive, and phenotypic analyses require genetic refinement. This has been achieved with molecular markers, which permit molecular characterization of the core collection. Such analysis is of great importance for effective conservation and use enhancement.

Among the molecular markers used in genetic diversity studies, those generated by Amplified Fragment Length Polymorphism (AFLP) analysis, described by Vos *et al.* (1995), are widely used to evaluate genetic diversity in plant species (EHRICH *et al.*, 2009; SUDHEER PAMIDIAMARRI *et al.*, 2009; TATIKONDA *et al.*, 2009; GARCÍA-PEREIRA *et al.*, 2010; MENDELSON *et al.*, 2010; GAUDEUL *et al.*, 2012).

## INTRODUÇÃO

Existem mais de 6,1 milhões de acessos de plantas conservados em, aproximadamente, 1320 bancos de germoplasma, dispostos nos 157 países que compõem a Comissão de Recursos Fitogenéticos da FAO (Food Agricultural Organization) (YETSIR *et al.*, 2008). Dos 47.000 acessos de milho conservados em bancos de germoplasma em todo o mundo, em torno de 27.000 compõem a coleção base de germoplasma de milho, localizada no Centro Internacional de Melhoramento de Milho e Trigo (CIMMYT), situado no México (YAN *et al.*, 2009).

No Brasil, existem mais de 200 mil acessos de plantas conservados *ex situ*, em cerca de 160 bancos de germoplasma implantados em mais de 50 locais (VASCONSELOS *et al.*, 2007). Existem também coleções cujo objetivo é reunir a maior parte da variabilidade genética de uma espécie em um conjunto limitado de acessos, denominadas de coleções núcleo. Dessa forma, já existem coleções núcleo de várias espécies, como a coleção núcleo de arroz, formada por 550 acessos (5,6% da coleção base), sendo 308 variedades tradicionais, 94 linhagens/cultivares melhoradas e 148 cultivares/linhagens introduzidas (ABADIE *et al.*, 2005), e a coleção núcleo de milho da Universidade Federal de Viçosa (UFV), que possui 104 acessos, sendo 80 landraces, 12 melhorados e 12 introduzidos de outros países. Essa coleção foi criada a partir da coleção ativa de milho da Embrapa Milho e Sorgo, que, segundo Netto *et al.* (2004), era composta por mais de 3500 acessos.

Geralmente, para o desenvolvimento de uma coleção núcleo, são consideradas inicialmente características morfológicas e agrônomicas, visando identificar a ocorrência de alterações nas médias, frequências e variâncias de características da coleção núcleo em relação à coleção base. Entretanto, devido aos efeitos ambientais, as informações sobre diversidade genética, considerando essas características, não são totalmente conclusivas, sendo necessário o refinamento do trabalho. Isso vem sendo alcançado com os marcadores moleculares, que permitem também caracterizar molecularmente a coleção núcleo, o que é de grande importância para a eficiente conservação e estímulo à sua utilização.

Entre os marcadores moleculares utilizados em estudos de diversidade genética destacam-se os marcadores gerados pela análise de polimorfismo de comprimento de fragmento amplificado (Amplified Fragment Length Polymorphism - AFLP), descrito por Vos *et al.* (1995), que são amplamente utilizados para avaliar diversidade genética em espécies vegetais (EHRICH *et al.*, 2009; SUDHEER PAMIDIAMARRI *et al.*, 2009; TATIKONDA *et al.*, 2009; GARCÍA-PEREIRA *et al.*, 2010; MENDELSON *et al.*, 2010; GAUDEUL *et al.*, 2012).

In a study of lettuce, Jansen and Van Hintum (2007), used AFLP markers to develop a new sampling method for obtaining core collections. Liu and Hou (2010) developed a core collection with *Vicia faba* L., also using AFLP markers. Sharma *et al.* (2017) concluded that the level of genetic diversity detected in *Pongamia* accessions suggested that AFLP markers were an efficient technology with which to define genetic relationships between genotypes and populations.

The use of molecular markers and morpho-agronomic characteristics is important for characterizing germplasm collection accessions. As a result, the aim of the current study was to use molecular techniques to characterize landrace accessions of the Federal University of Viçosa, corn core collection and assay genetic diversity in relation to existing classifications that have used grain type and ecogeographic origin.

## MATERIAL AND METHODS

The analysis used 80 landraces from the Federal University of Viçosa corn core collection (Table 1). This collection was developed using passport data, characterization, evaluation and subsequent classification of the accessions according to the ecogeographic origin and grain type. A sampling strategy based on multivariate analyzes was used, with sampling 10% of the accessions of the active corn germplasm collection from Embrapa Corn and Sorghum. Laboratory. Analyzes were carried out at the Molecular Biology Laboratory, Applied Biology Centre of the Embrapa Corn and Sorghum Centre, Sete Lagoas, Minas Gerais state (MG), Brazil.

To obtain plant material, 150 seeds of each accession were planted in beds. At 30 days after planting, the third leaf of 50 seedlings from each accession was collected, and combined in a bulk sample. These were then labeled, frozen in liquid nitrogen and stored at -80°C.

DNA extraction was performed according to the protocol described by Saghai-Marooof *et al.* (1984). After quantification, a DNA stock was maintained at -20°C, and the diluted DNA, at a working concentration of 50 ng  $\mu\text{L}^{-1}$ , was stored at 4°C.

Use of the AFLP marker kit followed the manufacturer's guidelines (USA; www.invitrogen.com). DNA restriction reactions used *EcoRI* and *MseI* enzymes, and adapter binding were used the AFLP *Core Reagent Kit* (Invitrogen), following the manufacturer's instructions.

Jansen e Van Hintum (2007), trabalhando com alface, utilizaram marcadores AFLPs para o desenvolvimento de um novo método de amostragem para a obtenção de coleções núcleo. Liu e Hou (2010) desenvolveram uma coleção núcleo de *Vicia faba* L. também utilizando marcadores AFLP. Sharma *et al.* (2017) concluíram que o nível de diversidade genética detectado em acessos de *Pongamia* sugere que os marcadores AFLP representam uma tecnologia eficiente para delinear as relações genéticas entre genótipos e populações.

O uso de marcadores moleculares e as características morfo-agronômicas são importantes na caracterização de acessos em coleções de germoplasma. Diante do exposto, objetivou-se com este trabalho caracterizar molecularmente os acessos landraces da coleção núcleo de milho da Universidade Federal de Viçosa e verificar a diversidade genética em relação a estratificação, considerando tipo de grão e origem ecogeográfica.

## MATERIAL E MÉTODOS

Foram utilizados 80 acessos landraces pertencentes à coleção núcleo de milho da Universidade Federal de Viçosa (Tabela 1). Essa coleção foi desenvolvida utilizando-se dados de passaporte, caracterização, avaliação e posterior estratificação dos acessos de acordo com a origem ecogeográfica e o tipo de grão. Foi utilizada a estratégia de amostragem baseada em análises multivariadas, com intensidade de amostragem de 10% dos acessos da coleção ativa de germoplasma de milho da Embrapa Milho e Sorgo. As análises laboratoriais foram realizadas no Laboratório de Biologia Molecular do Núcleo de Biologia Aplicada (NBA) da Embrapa Milho e Sorgo, Sete Lagoas, MG – Brasil.

Para obtenção do material vegetal, foram semeadas em canteiros 150 sementes de cada acesso. Aos 30 dias após o plantio, foi coletada a terceira folha de 50 plântulas por acesso, sendo formado um *bulk*. Esses *bulks* foram identificados, congelados em nitrogênio líquido e armazenados a -80 °C.

A extração do DNA foi realizada segundo o protocolo descrito por Saghai-Marooof *et al.* (1984). Após a quantificação, o DNA estoque foi mantido a -20°C, e o DNA diluído, na concentração de trabalho de 50 ng  $\mu\text{L}^{-1}$ , foi armazenado a 4 °C.

O kit de marcadores AFLP utilizado seguiu as orientações do fabricante (USA; www.invitrogen.com). As reações de restrição de DNA com as enzimas *EcoRI* e *MseI* e a ligação dos adaptadores foram realizadas utilizando o AFLP *Core Reagent Kit* (Invitrogen), seguindo as informações do fabricante.

**Table 1** – Description of the 80 landrace accessions in the corn core collection at the Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil**Tabela 1** – Relação dos 80 acessos landraces da coleção núcleo de milho da Universidade Federal de Viçosa, Viçosa – MG

Nº	Name	Region <sup>1/</sup>	Grain Type <sup>2/</sup>	Nº	Name	Region <sup>1/</sup>	Grain Type <sup>2/</sup>
1	AC 003	4	1	41	PA 111	4	3
2	AC 013	4	2	42	PB 012	5	1
3	AL 002	5	1	43	PB 022	5	1
4	AL 005	5	2	44	PE 004	6	1
5	AM 002	4	3	45	PE 005	5	1
6	AM 012	4	2	46	PE 015	5	1
7	BA 012	5	3	47	PE 031	5	1
8	BA 045	5	1	48	PE 049	5	2
9	BA 072	6	1	49	PE 054	5	2
10	BA 073	6	2	50	PR 002	1	1
11	BA 126	6	2	51	PR 062	2	4
12	BA 173	2	1	52	RGS XIII	1	1
13	BA 178	2	1	53	RN 006	5	2
14	BA 195	2	2	54	RN 011	6	2
15	BA 211	5	2	55	RR 007	3	2
16	BA 226	5	1	56	RR 020	3	1
17	BA 233	5	3	57	RR 057	3	1
18	BA 237	6	4	58	RR 075	3	2
19	BA II – Cr	2	3	59	RR 108	4	2
20	CE 007	5	1	60	RR 117	4	4
21	CE 013	6	1	61	RR 123	4	2
22	CE 014	6	3	62	RR 172	3	2
23	CE 025	5	2	63	RS 001	1	3
24	DESC I	2	3	64	RS 013	1	1
25	GO 001	2	2	65	RS 057	1	1
26	Grã Turvo	2	2	66	RS 062	1	2
27	MA 012	5	1	67	RS 075	1	1
28	MG 013	2	1	68	RS 081	1	4
29	MG 094	2	2	69	RS 082	1	1
30	MG 100	2	3	70	RS 083	1	1
31	MG 108	2	1	71	RS 092	1	2
32	MG 111	2	2	72	SC 027	1	1
33	MG 112	2	1	73	SP 001	2	1
34	MG 119	2	3	74	SP 002	2	1
35	MT 038	2	4	75	SP 005	2	2
36	PA 012	4	2	76	SP 034	2	2
37	PA 070	4	2	77	SP 103	2	1
38	PA 071	4	2	78	SP 118	2	2
39	PA 076	4	3	79	SP 135	2	1
40	PA 087	4	1	80	Sul PA	4	1

<sup>1/</sup>Ecographic Region: 1 – South; 2 – Cerrado; 3 – Northern Cerrado; 4 – Amazônia; 5 – Caatinga; and 6 – Littoral.

<sup>2/</sup>Grain Type: 1 – Dentate; 2 – Semidentate; 3 – Hard; and 4 – Semihard.

<sup>1/</sup>Região Ecogeográfica: 1 – Sul; 2 – Cerrados; 3 – Cerrados-Norte; 4 – Amazônia; 5 – Caatinga; e 6 – Agreste-Litoral. <sup>2/</sup>Tipo de grãos: 1 – Dentado; 2 – Semidentado; 3 – Duro; e 4 – Semiduro.

Selective and pre-selective amplification reactions were performed using Applied Biosystems kits (USA, www.appliedbiosystems.com), with some modifications. Pre-selective amplification products were diluted 20-fold in TE, and subjected to selective amplification using 12 combinations of EcoRI-MseI primers, with three selective nucleotides at the 3' terminus. The EcoRI primer was labeled with one of three fluorescences 6-FAM, JOE or NED options. Combinations used were: *EcoRI-ACT/MseI-CTC*; *EcoRI-AGG/MseI-CTC*; *EcoRI-ACC/MseI-CTC*; *EcoRI-ACA/MseI-CTG*; *EcoRI-AGG/MseI-CTG*; *EcoRI-AGC/MseI-CTG*; *EcoRI-ACA/MseI-CAC*; *EcoRI-AGG/MseI-CAC*; *EcoRI-ACC/MseI-CAC*; *EcoRI-ACT/MseI-CTA*; *EcoRI-ACG/MseI-CTA* e *EcoRI-AAC/MseI-CTA*.

AFLP fragments were electrophoretically analyzed with a ABI Prism 377 sequencer at 1680 V for 5 h. Gel analysis, identifying amplified alleles with the fluorescent primers, was performed with Genescan 2.1 software. To obtain binary tables, Genotyper 2.0 software was used. Each tag was considered as a different locus, and was coded so that the presence of a band was coded as (1), and absence as (0), thus providing an array of binary data.

Prior to genetic analysis, the 80 analysed accessions from the core collection had been grouped with the Unweighted Pair Group Method using Arithmetic mean (UPGMA), with dissimilarity calculated using the Jaccard Index. To identify divergent groups, the mean dissimilarity matrix distance was considered as the cutoff point.

To use grain types and ecogeographic regions as grouping criteria to obtain genetically divergent groups, the 80 core collection accessions were divided into four grain type sub-populations, and six ecogeographic origin sub-populations. To calculate the genetic distance between accession groups genetic distance value (NEI 1972) was used to generate a dissimilarity matrix which was subsequently grouped using the UPGMA method.

Upon attaining a Hardy-Weinberg equilibrium, the frequency ( $p_{ij}$ ) of each allele (i) at each locus (j) was calculated. Using Nei (1978)'s methods for the 80 accessions, total diversity (HT), mean inter-group diversity ( $\bar{H}_s$ ), between-group diversity component (STD), and the proportion of genetic diversity attributed to the between-population component (GST) was calculated.

## RESULTS AND DISCUSSION

Most of the grains in the core collection accessions were serrate and semi-dentate types. In terms of ecogeographic origin, the majority of accessions came from the Cerrado, Caatinga and Amazonia, with smaller numbers from the South, northern Cerrados and Littoral (Table 1).

As reações de amplificação seletiva e pré-seletiva foram realizadas empregando-se os *kits da Applied Biosystems* (USA, www.appliedbiosystems.com) com algumas modificações. Os produtos da amplificação pré-seletiva foram diluídos 20 vezes em TE e submetidos a amplificações seletivas, utilizando-se 12 combinações dos *primers EcoRI-MseI*, com três nucleotídeos seletivos na extremidade 3'. O *primer EcoRI* foi marcado com uma das três fluorescências 6-FAM, JOE ou NED. As combinações utilizadas foram: *EcoRI-ACT/MseI-CTC*; *EcoRI-AGG/MseI-CTC*; *EcoRI-ACC/MseI-CTC*; *EcoRI-ACA/MseI-CTG*; *EcoRI-AGG/MseI-CTG*; *EcoRI-AGC/MseI-CTG*; *EcoRI-ACA/MseI-CAC*; *EcoRI-AGG/MseI-CAC*; *EcoRI-ACC/MseI-CAC*; *EcoRI-ACT/MseI-CTA*; *EcoRI-ACG/MseI-CTA* e *EcoRI-AAC/MseI-CTA*.

Os fragmentos de AFLP foram analisados no sequenciador ABI Prism 377 sob eletroforese, a 1.680 V, por 5 h. A análise dos géis para identificação dos alelos amplificados com os *primers* fluorescentes foi realizada com o software Genescan 2.1; para a obtenção das tabelas binárias, utilizou-se o software Genotyper 2.0. Cada marca foi considerada como um loco diferente, sendo codificado (1) para presença da banda e (0) para ausência, sendo obtida uma matriz de dados binários.

Foi realizado previamente o agrupamento dos 80 acessos da coleção núcleo de acordo com o método de ligação média (Unweighted Pair Group Method using Arithmetic mean - UPGMA), tendo como medida de dissimilaridade o complemento do índice de Jaccard. Para a identificação de grupos divergentes foi considerado como ponto de corte a média das distâncias da matriz de dissimilaridades.

Para avaliar o tipo de grão e a região ecogeográfica como critérios de agrupamento para obtenção de grupos geneticamente divergentes, os 80 acessos da coleção núcleo foram divididos em quatro sub-populações quanto ao tipo de grão e em seis sub-populações quanto à origem ecogeográfica. Neste caso, para calcular a distância genética entre grupos de acessos, foi utilizada a distância genética de Nei (1972) para gerar a matriz de dissimilaridade e posterior agrupamento pelo método de UPGMA.

Ao admitir o equilíbrio de Hardy-Weinberg, foi calculada a frequência ( $p_{ij}$ ) de cada alelo (i) em cada loco (j). Considerando os 80 acessos, de acordo com Nei (1978), foram estimadas a diversidade total (HT), diversidade média entre grupos ( $\bar{H}_s$ ), componente de diversidade entre grupos (DST) e a proporção da diversidade genética que é atribuída ao componente entre populações (GST).

## RESULTADOS E DISCUSSÃO

A maioria dos acessos da coleção núcleo possuem grãos do tipo dentado e semidentado. Em relação à origem ecogeográfica, a maioria dos acessos são das regiões Cerrado, Caatinga e Amazônia, existindo também acessos das regiões Sul, Cerrados-Norte e Agreste-Litoral (Tabela 1).

In the current study, the choice of a bulk-sample of 50 seedlings to represent each accession from the core collection was made, following the studies of Dubreuil *et al.* (1998) and Rebourg *et al.* (1999), each of whom, concluding that bulk analysis can be used to represent corn populations, used 30 plant bulk-samples to represent accessions. Malosetti *et al.* (2000), working with simulations and AFLP markers from bulk samples based on 50 individuals, reported that the probability of detecting alleles with frequencies between 0.9 and 0.4 was 100%, and that it was highly likely that alleles with frequencies of 0.3 would be detected in such bulk samples. Garcia *et al.* (2004), also working with accessions, used 16 plants for sampling purposes. Therefore, due to the great variability in corn accessions, there appears to be no defined optimal number of seedlings that may be sampled to represent an accession.

The presence of an allele in an accession whose DNA has been extracted from a bulk-sample of 50 seedlings does not mean that all seedlings in the sample possess this allele. This is because corn is an allogamic species. Consequently, the level of genetic variability within each accession may be great. Since it is important to be certain whether or not the accession has a given allele, when a given allelic band was seen on a gel it was assumed that all 50 individuals forming the bulk-sample had it, or that the allele was, at least, present in the sampled population. For plant breeding purposes, the presence of an allele is sufficient to guarantee the use of the accession, although different numbers of backcrosses may be required to introduce the allele into any improved population or lineage that might be developed.

Analysis amplified 720 bands from the 80 used accessions, an average of 60 bands per primer combination (Table 2), ranging from 40, for the combination of EcoRI ACG + MseI CTA primers, to 79 bands for EcoRI ACT + MseI CTC. The observed degree of polymorphism was high, with 660 of the 720 bands being polymorphic (92% polymorphism). The degree of polymorphism for each combination of primers ranged from 82.3 to 100%, with combinations of 4, 5 and 12 showing 100% polymorphism (Table 2). Fragment amplitudes, taking into account all 12 primer combinations, ranged from 63.39 to 476.68 base pairs (bp). These results corroborate observations by Meudt and Clarke (2007), that AFLPs markers can generate a large number of polymorphic bands.

When working with landraces, and using a larger number of primers, a greater number of polymorphism-related bands can be obtained, as observed by Garcia *et al.* (2004), who used 20 combinations of AFLP primers on corn landraces and obtained 774 bar polymorphisms. Clearly, the higher the number of primers used, the more efficient the obtained results are likely to be. In consequence, the number of primers used in practice depends on the time and financial resources available to carry out a given study. However, it is possible to optimize the number of primers used, so that polymorphisms present in the study population will be identified representatively.

Neste trabalho, a escolha de um *bulk* de 50 plântulas para representar cada acesso da coleção núcleo foi feita com base nos trabalhos de Dubreuil *et al.* (1998) e Rebourg *et al.* (1999), que utilizaram um *bulk* de 30 plantas para representar acessos de milho, chegando à conclusão que a análise em *bulk* pode ser utilizada para representar populações de milho. Malosetti *et al.* (2000), trabalhando com simulações e marcadores AFLP e *bulk* de 50 indivíduos, verificaram que a probabilidade de serem detectados alelos com frequências entre 0,9 e 0,4 foi de 100%, sendo que alelos com frequências de 0,3 teriam a probabilidade de 98% de serem detectados nesse *bulk*. Portanto, devido à grande variabilidade existente em acessos de milho, não existe um número ótimo de plântulas a ser amostrado para representar um acesso. Garcia *et al.* (2004), também trabalhando com linhagens, utilizaram 16 plantas para o mesmo propósito. Percebe-se que mesmo utilizando linhagens o número de indivíduos amostrados é variável.

A presença de um alelo em um acesso cujo DNA foi extraído de um *bulk* de 50 plântulas não significa que as plântulas possuam este alelo. Isso ocorre porque o milho é uma espécie alógama e apresenta grande variabilidade dentro de cada acesso. Como o que interessa é verificar se o acesso possui ou não determinado alelo, quando na presença da banda visualizada no gel, foi considerado que todos os 50 indivíduos do *bulk* a possuíam, ou que, pelo menos, o alelo está presente na população. Para fins de melhoramento de plantas, a presença do alelo é suficiente para a utilização do acesso, apesar que diferentes números de retrocruzamentos serão necessários para introgridir o alelo na população melhorada ou na linhagem a ser desenvolvida.

Para os 80 acessos utilizados, foram amplificadas 720 bandas, com uma média de 60 bandas por combinação de *primers* (Tabela 2), variando de 40, para a combinação de *primers* EcoRI ACG + MseI CTA, a 79 bandas, para a combinação EcoRI ACT + MseI CTC. Foi observado elevado grau de polimorfismo, sendo obtidas 660 bandas polimórficas, o que equivale a 92 de polimorfismo. O grau de polimorfismo para cada combinação de *primers* variou entre 82,3 a 100%, destacando-se as combinações 4, 5 e 12 com 100% de polimorfismo (Tabela 2). A amplitude dos fragmentos, considerando todas as 12 combinações de *primers*, variou de 63,39 a 476,68 pares de bases (pb). Estes resultados corroboram com a afirmação de Meudt e Clarke (2007), os quais afirmam que marcadores AFLPs podem gerar um grande número de bandas polimórficas.

Mesmo trabalhando com linhagens, ao se utilizar um maior número de *primers*, pode-se obter maior número de marcas, como observado por Garcia *et al.* (2004), que utilizaram 20 combinações de *primers* AFLP em linhagens de milho, obtendo 774 marcas. Fica evidente que quanto maior o número de *primers* utilizados mais eficiente o trabalho e que na prática a quantidade utilizada depende da disponibilidade de tempo e de recurso financeiro para a realização do trabalho. Porém, pode-se obter um número ótimo de primers que seja o suficiente para identificar o polimorfismo.

**Table 2** – Fragment size by base-pair numbers, total number of bands, number of polymorphic bands, and percentage of polymorphism for each combination of primers and EcoRI/MseI**Tabela 2** – Tamanho dos fragmentos em pares de bases, número total de bandas, número de bandas polimórficas e porcentagem de polimorfismo para cada combinação de primers seletivos EcoRI/MseI

Combination	Selected Nucleotides		Fragment variation (pb)	Total no. bands	No. Polymorphic bands	% Polimorphism
	EcoRI	MseI				
1	ACT	CTC	68.00 – 490.00	79	65	82.3
2	AGG	CTC	64.14 – 494.52	73	68	93.2
3	ACC	CTC	59.80 – 476.88	58	51	87.9
4	ACA	CTG	51.67 – 451.75	59	59	100.0
5	AGG	CTG	56.11 – 482.43	59	59	100.0
6	AGC	CTG	56.11 – 481.49	58	55	94.8
7	ACA	CAC	70.72 – 490.59	63	54	85.7
8	AGG	CAC	59.58 – 500.00	66	61	92.4
9	ACC	CAC	70.14 – 490.94	45	42	93.3
10	ACT	CTA	62.57 – 491.35	72	64	88.9
11	ACG	CTA	64.08 – 495.86	40	34	85.0
12	AAC	CTA	77.66 – 373.91	48	48	100.0
Average				60	55	92.0
Total				720	660	

Netto *et al.* (2004), using AFLP markers and six primer combinations to molecularly characterize corn core collection accessions of the hard endosperm subgroup of the Embrapa Maize and Sorgum collection, obtained 207 fragments ranging in size from 52.43 to 498.79 bp, with an average of 32.8 polymorphic bands per primer combination, and fragment numbers ranging from 24 to 50. They too considered that molecular characterization allowed an adequate assessment of the genetic variability present in the accessions.

When using the UPGMA method to group 80 landrace accessions, with dissimilarity measured by the Jaccard Index, six affinity groups were found (Figure 1), with groups I and II each having a single accession. These two accessions differed greatly from all other accessions in the core collection. Groups III and IV contained 16 and 27 accessions, respectively. Group V was formed by only three accessions, and group VI by 32 accessions. Groups III, IV and VI contained members of several subgroups. These results illustrate the great genetic diversity in the UFV corn core collection.

Netto *et al.* (2004), utilizando marcadores AFLPs e seis combinações de *primers* para caracterizar molecularmente os acessos da coleção núcleo de milho da Embrapa Milho e Sorgo, subgrupo endosperma duro, obtiveram 207 fragmentos com tamanho variando de 52,43 a 498,79 pb, com média de 32,8 bandas polimórficas por combinação de *primers*, tendo o número de fragmentos variando de 24 a 50. Consideraram ainda que a caracterização molecular possibilitou a identificação adequada da variabilidade genética entre os acessos.

Ao realizar o agrupamento dos 80 acessos landraces com base no método de UPGMA, tendo como medida de dissimilaridade o complemento do índice de Jaccard, foi notada a formação de seis grupos de similaridade (Figura 1), sendo os grupos I e II formados por um acesso cada. Esses dois acessos se mostraram bem divergentes em relação aos demais acessos da coleção núcleo. Os grupos III e IV foram constituídos por 16 e 27 acessos, respectivamente. O grupo V foi formado por apenas três acessos e o grupo VI por 32 acessos. Os grupos III, IV e VI foram formados por vários subgrupos. Esses resultados evidenciam a grande diversidade genética existente na coleção núcleo de milho da UFV.

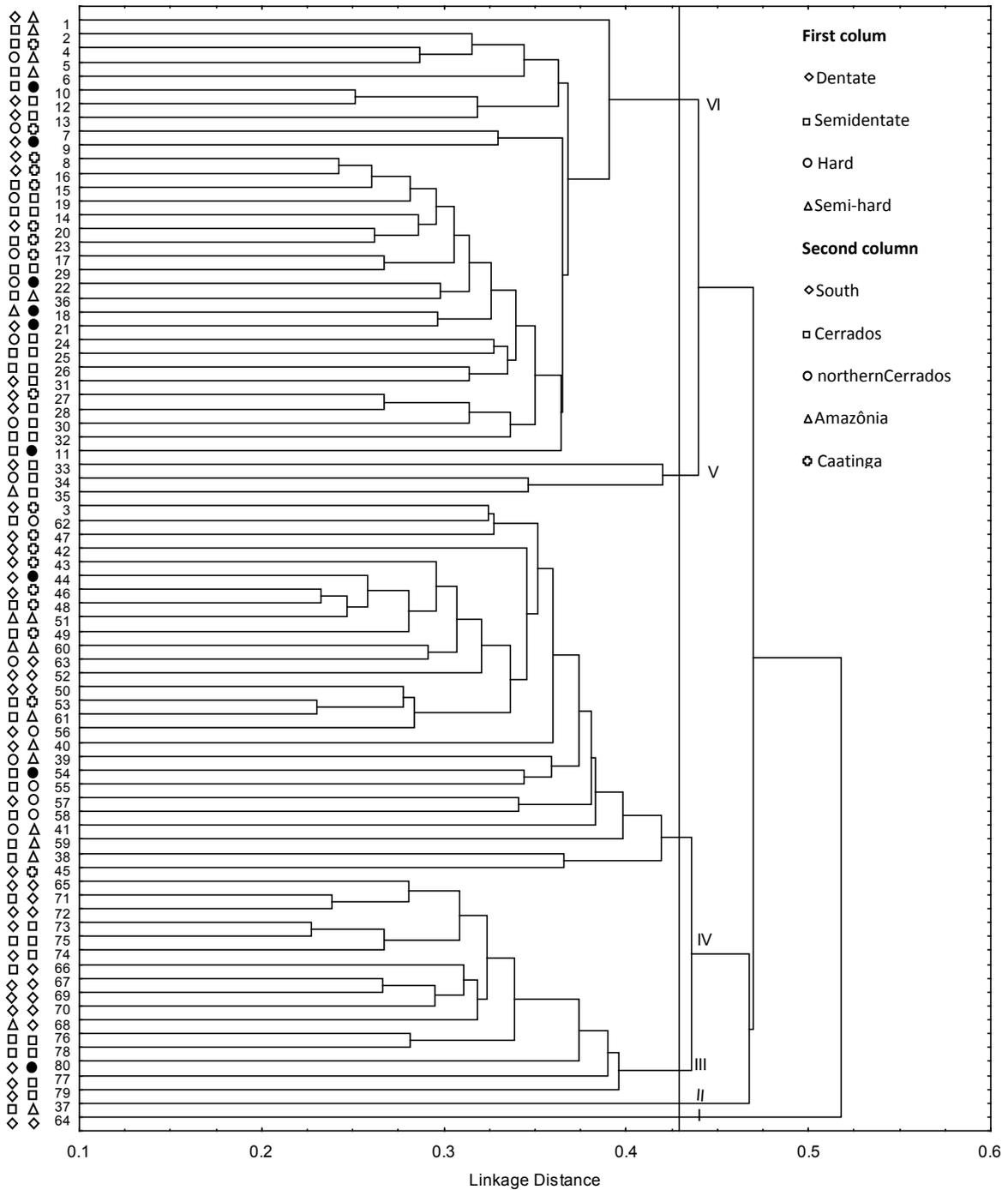


Figure 1 – Grouping of 80 core collection accessions by mean linkage method between groups, using Jaccard Index mean dissimilarity.

Figura 1 – Agrupamento dos 80 acessos da coleção núcleo, de acordo com o método de ligação média entre grupos, tendo como medida de dissimilaridade o índice de Jaccard.

A number of studies have used clustering methods based on molecular data, and reported positively on their efficiency (OLIVEIRA *et al.*, 2013, MARTINS *et al.*, 2014, PONTES *et al.*, 2017). For example, in the study carried out by Odong *et al.* (2011), the authors confirmed the effectiveness of UPGMA methods for determining the genetic structure of germplasm collections. In the current study, it was observed that within the groups formed, accessions differed in terms of grain type and ecogeographic origin. However, there was a notable tendency for accessions with the same type of grain to appear in the same subgroup, this was also true for ecogeographic origins. In a few subgroups, accessions with the same grain type and from the same ecogeographic region co-occurred. This was also true for a subgroup of group III, which was comprised of accessions 67 (RS075), 69 (RS082) and 70 (RS083), all from the South region and with dentate grains.

When evaluating grain type as a grouping criterion, it was found that the four resulting sub-populations showed similar genetic diversity, despite their different sizes (Table 3). Only 4.69% of diversity was associated with differences between sub-populations.

A utilização de métodos de agrupamento com base em dados moleculares tem sido utilizada e sua eficiência verificada em vários trabalhos (OLIVEIRA *et al.* 2013; MARTINS *et al.* 2014; PONTES *et al.* 2017), como, por exemplo, no estudo realizado por Odong *et al.* (2011), os autores confirmaram a eficiência de métodos como o UPGMA para determinar a estrutura genética de coleções de germoplasma. Foi observado que dentro dos grupos formados, os acessos diferiram em relação ao tipo de grão e origem ecogeográfica. Entretanto, foi notada a tendência de que acessos com mesmo tipo de grão sejam agrupados em um mesmo subgrupo, o mesmo ocorreu em relação a origem ecogeográfica. Em poucos subgrupos observam-se acessos com o mesmo tipo de grãos e oriundos de uma mesma região ecogeográfica, como é o caso do subgrupo do grupo III, composto pelos acessos 67 (RS075), 69 (RS082) e 70 (RS083) que são da região Sul e possuem grãos dentados.

Ao avaliar o tipo de grão como critério de agrupamento, foi detectado que as quatro subpopulações formadas apresentaram diversidade genética semelhante, apesar do diferente tamanho (Tabela 3). Apenas 4,69% da diversidade foi associada a diferenças entre sub-populações.

**Table 3** – Total number of polymorphic loci (%), sample size, within sub-population diversity component (HS), total diversity (HT) and proportion of diversity attributed to the between sub-populations component (GST), for populations with different grain types and ecogeographic origins

**Tabela 3** – Número total e de locos polimórficos (%), tamanho da amostra, componente de diversidade dentro de subpopulação (HS), diversidade total (HT) e proporção da diversidade atribuída ao componente entre sub-populações (GST) das populações com diferentes tipos de grãos e diferentes origens ecogeográficas

Subpopulation	Number of polymorphic loci (%)	Sample size (accessions)	HS	HT	GST (%)
<b>Grain type</b>					
Dentate	617 (85.69)	35	0.2322		
Semidentate	596 (82.78)	29	0.2278		
Hard	482 (66.94)	11	0.2215		
Semi-hard	388 (53.89)	5	0.2051		
<b>All accessions</b>	<b>659 (91.53)</b>	<b>80</b>	<b>0.2216 1/</b>	<b>0.2325</b>	<b>4.69</b>
<b>Ecogeographic origin</b>					
1 - South	424 (58.89)	12	0.1866		
2 - Cerrado	563 (78.19)	23	0.2101		
3 - Northern Cerrado	303 (42.08)	5	0.1697		
4 - Amazônia	506 (70.28)	14	0.2261		
5 - Caatinga	538 (74.72)	18	0.2228		
6 -Littoral	430 (59.72)	8	0.1966		
<b>All accessions</b>	<b>659 (91.53)</b>	<b>80</b>	<b>0.2020/1</b>	<b>0.2302</b>	<b>12.25</b>

<sup>1/</sup> HS mean.

<sup>1/</sup> Média do HS.

UPGMA analysis gave three distinct groups, the first being composed of the semi-hard sub-population, the second by the hard sub-population and the third by the dentate and semi-hard sub-populations (Figure 2). It may be suggested that, since the allelic frequency for the different accessions or populations are largely the same, differences are probably related to survival and adaptation metabolism, and that only a few alleles are rare. These rare alleles may be related to specific characteristics selected naturally, either by the environment or by man during decades of cultivation. These results corroborate those of Brieger *et al.* (1958), who stated that genotypes with different endosperm types are genetically divergent because they have different evolutionary origins.

Ao se realizar o agrupamento pelo método de UPGMA, foram formados três grupos distintos, sendo o primeiro composto pela sub-população semiduro, o segundo pela sub-população duro e o terceiro pelas subpopulações dentado e semidentado (Figura 2). Pode se sugerir que a base da frequência alélica dos diferentes acessos ou populações é a mesma, possivelmente relacionada ao metabolismo de sobrevivência e adaptação e que apenas alguns alelos são raros. Esses alelos raros podem estar relacionados à características específicas selecionadas naturalmente pelo ambiente ou pelo homem durante décadas de cultivo. Esses resultados corroboram com os obtidos por Brieger *et al.* (1958), que afirmaram que genótipos com diferentes tipos de endosperma são divergentes geneticamente por possuírem origem evolucionária diferente.

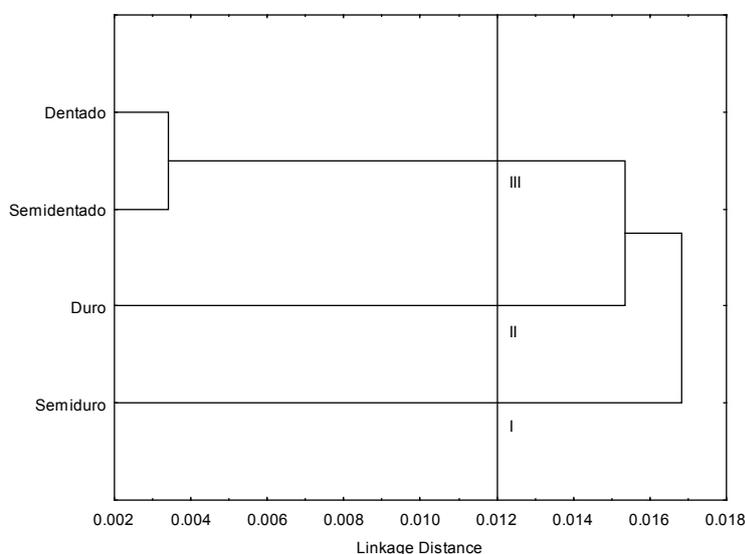


Figure 2 – Grouping of 80 core collection accessions for grain type, using the mean inter-group connection method, using mean dissimilarity and Nei (1972)'s measure of genetic distance.

Figura 2 – Agrupamento dos 80 acessos da coleção núcleo quanto ao tipo de grão pelo método de ligação média entre grupos, tendo como medida de dissimilaridade a distância genética de Nei (1972).

Regarding the presence of specific alleles, 16, 4, 1 and 1 alleles were found, to occur, respectively, in only the dentate, semi-dentate, hard and semi-hard sub-populations. Accessions with dentate and semi-dentate grain types had the highest number of polymorphic bands, and among the dentate accessions there appears to be a greater diversity in relation to the other sub-populations in the core collection. It was also noted that the dentate and semidentate accessions are more similar to each other than to the hard and semi-hard grain accessions (Figure 2). In Brazil, rural agriculturalists have always been preferred dentate maize types to hard forms for their ease of handling when making typical local dishes such as *angu* (cornmeal porridge) and cornmeal cakes. However, with the increase of commercial agriculture and corn processing of by large

Em relação à ocorrência de alelos específicos, foram encontrados 16, 4, 1 e 1 alelos ocorrendo somente nas subpopulações dentado, semidentado, duro e semiduro, respectivamente. Os acessos com tipo de grãos dentado e semidentado foram os que apresentaram maior número de bandas polimórficas, sendo possível inferir que entre os acessos dentados existe maior diversidade em relação às demais subpopulações da coleção núcleo. Também foi observado que os acessos dentados e semidentados são mais similares entre si do que os acessos com grão do tipo duro e semiduro (Figura 2). No Brasil, as populações de milho dentadas sempre foram preferidas em relação às populações duras pela facilidade de manipulação para fazer pratos típicos locais como *angu* e bolos de fubá. No entanto, com o aumento da agricultura empresarial e o processamento do milho por grandes indústrias, o milho tipo duro começou a ter importância na cadeia produtiva.

industries, hard corn types have become more important in the production chain.

Because of the importance of making new germplasm available to broaden the genetic base of breeding programs, core collections are characterized fairly frequently. For example, Malvar *et al.* (2007) assayed the capacity of accessions in the European Union Millet Landrace Collection to survive infestation by millet borers (*Ostrinia nubilalis* Hbn.), an important regional pest, while Ferro *et al.* (2007) worked with the same germplasm collection under field conditions, and analyzed productivity in high- and low-dose nitrogen situations, a high-percentage contributor to overall production costs, and a potential environmental pollutant whose presence in waterbodies can result in their eutrophication.

In the current study, sorting core collection accessions by ecogeographic region resulted in two similarity groups. The first was composed of northern-Cerrado, Cerrado, Caatinga, Littoral and Amazônia sub-populations, the second contained only the South sub-population (Figure 3). The most diverse sub-populations came from the Amazon and the Caatinga regions (higher  $H_s$ ), while the northern Cerrado accessions had the lowest diversity. The Cerrado sub-population had the highest percentage of polymorphism (78.19%), but it was not the one with the highest diversity. Greater polymorphism does not necessarily imply greater diversity, since while only one individual can possess a particular band, it requires only a single individual not to possess it for the locus to be considered polymorphic (Table 3).

A caracterização de coleções núcleos têm sido realizada com certa frequência em função da importância de disponibilizar novo germoplasma para ampliar a base genética de programas de melhoramento. Malvar *et al.* (2007) avaliaram a performance produtiva de acessos da Coleção Núcleo de Milho Landrace da União Europeia em relação à infestação por brocas (*Ostrinia nubilalis* Hbn.), importante praga regional. Ferro *et al.* (2007) trabalharam com a mesma coleção de germoplasma em condições de campo, analisaram a produção nas situações de alta e baixa dose de nitrogênio, nutriente de maior valor para o custo de produção e que pode ser poluidor de ambiente, resultando na eutrofização de corpos d'água.

A estratificação dos acessos da coleção núcleo em relação à região ecogeográfica formou dois grupos de similaridade. O primeiro grupo foi composto pelas subpopulações Cerrados-Norte, Cerrados, Caatinga, Agreste-Litoral e Amazônia, o segundo grupo foi formado apenas pela subpopulação Sul (Figura 3). As populações que apresentaram maiores diversidades foram as da Amazônia e da Caatinga (maiores  $H_s$ ), sendo a população Cerrados-Norte aquela com menor diversidade. A subpopulação Cerrados foi a que apresentou maior porcentagem de polimorfismo (78,19%), mas não foi a que apresentou maior diversidade. O maior grau de polimorfismo não implica necessariamente em maior diversidade, pois apenas um indivíduo pode possuir determinada banda, ou somente um indivíduo não possuí-la para que o loco seja considerado polimórfico (Tabela 3).

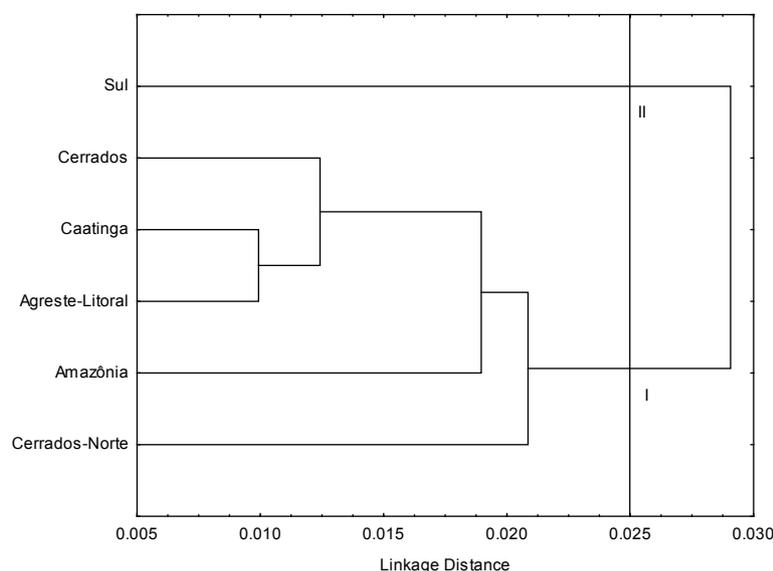


Figure 3 – Grouping of 80 accessions from the core collection by ecogeographic region, using a mean between-group connection method, using the dissimilarity and genetic distance measures of Nei (1972).

Figura 3 – Agrupamento dos 80 acessos da coleção núcleo quanto a região ecogeográfica pelo método de ligação média entre grupos, tendo como medida de dissimilaridade a distância de Nei (1972).

The greatest divergence was observed between the accessions from the South and northern-Cerrados regions. The most similar regions were the Caatinga and Littoral. These results suggest that the development of corn core collections should take into account the type of grain and the ecogeographic origin when seeking to access genetically divergent groups.

In a study of the microsatellites (SSR) of 124 accessions originating from the mountain region of Youling, China, Yao *et al.* (2008) also reported that using ecogeographic information was a highly efficient means of obtaining genetically divergent groups.

The proportion of diversity attributed to differences between sub-populations with different ecogeographic origins was 12.25%, higher than that observed for grain type. In the South, Cerrado and northern Cerrados subpopulations, a specific allele was found in each sub-population. In the Caatinga sub-population, three specific alleles were found, while no specific alleles were observed in the Amazonia and Littoral sub-populations. It was notable among the landraces accessions that a greater diversity occurred in relation to grain type than in relation to ecogeographic origin, while rare alleles were present in all sub-populations.

## CONCLUSIONS

There is great genetic variability among the landrace accessions in the Corn Core Collection of UFV, especially among the dentate grain accessions;

Differentiation by grain type and ecogeographic origin revealed genetically divergent groups, and these characters should be used in conjunction when developing corn core collections;

There is genetic variability within landraces, and rare alleles have been found in all sub-populations.

A maior divergência foi observada entre os acessos das regiões Sul e Cerrados-Norte. As regiões mais similares foram a Caatinga e Agreste-Litoral. Tais resultados sugerem que no desenvolvimento de coleções núcleo de milho deve-se levar em consideração o tipo de grão e a origem ecogeográfica para a obtenção de grupos geneticamente divergentes.

Em um trabalho com microsatélites (SSR) e 124 acessos originários da região da montanha Youling na China, Yao *et al.* (2008) também verificaram a eficiência de utilização de informações ecogeográficas para a obtenção de grupos geneticamente divergentes.

A proporção da diversidade atribuída às diferenças entre as subpopulações com diferentes origens ecogeográficas foi de 12,25%, valor superior ao observado quanto ao tipo de grão. Nas subpopulações Sul, Cerrados e Cerrados-Norte foi verificada a ocorrência de um alelo específico em cada subpopulação. Na subpopulação Caatinga, foi constatada a presença de três alelos específicos, e nas subpopulações Amazônia e Agreste-Litoral não foram observados alelos específicos. Foi notada entre os acessos landraces uma maior diversidade em relação ao tipo de grão do que em relação à origem ecogeográfica e que alelos raros são encontrados em todas as subpopulações.

## CONCLUSÕES

Existe grande variabilidade genética entre os acessos landraces da Coleção Núcleo de milho da UFV, principalmente entre os acessos com grãos do tipo dentado;

A estratificação quanto ao tipo de grãos e origem ecogeográfica contribuiu para a obtenção de grupos geneticamente divergentes e devem ser utilizadas em conjunto no desenvolvimento de coleções núcleo de milho;

Existe variabilidade genética para os acessos landraces, entretanto, foram encontrados alelos raros em todas as subpopulações.

## CITED SCIENTIFIC LITERATURE

ABADIE, T.; CORDEIRO, C. M. T.; FONSECA, J. R.; ALVES, R. B. N.; BURLE, M. L.; BRONDANI, C.; RANGEL, P. H. N.; CASTRO, E. M.; SILVA, H. T.; FREIRE, M. S.; ZIMMERMANN, F. J. P.; MAGALHÃES, J. R. Construção de uma coleção nuclear de arroz para o Brasil. **Pesquisa Agropecuária Brasileira**. 2005;40:2:129-136. doi:10.1590/s0100-204x2005000200005

BRIEGER, A. H. D.; GRUGEL, J. T. A.; PATERNIANI, E.; ALLEONI, M. R. Races of maize in Brazil and other Eastern South American countries. ed. 593. National Academy, 283 p.

DUBREUIL, P.; CHARCOSSET, A. Genetic diversity within and among maize populations: a comparison between isozyme and nuclear RFLP loci. **Theoretical Applied Genetics**. 1998;96:577-587. doi:10.1007/s001220050

EHRICH, D.; EIDEN, P. B.; ALSOS, I. G.; BROCHMANN, C. An AFLP clock for absolute dating of shallow-time evolutionary history - too good to be true? **Molecular Ecology**. 2009;18:4526-4532. doi:10.1111/j.1365-294x.2009.04387.x

- FERRO, R. A.; BRICHETTE, I.; EVGENIDIS, G.; KARAMALIGKAS, C. H.; GONZALEZ, J. M. Variability in European maize (*Zea mays* L.) landraces under high and low nitrogen inputs. **Genetic Resources and Crop Evolution**. 2007;54:295-308. doi:10.1007/s10722-005-4500-x
- GARCIA, A. A. F.; BENCHIMOL, L. L.; BARBOSA, M. A. M.; GERALDI, I. O.; SOUZA, J. R. C. L.; SOUZA, A. P. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. **Genetics and Molecular Biology**. 2004;24:4:579-588. doi:10.1590/s1415-47572004000400019
- GARCÍA-PEREIRA, M. J.; CABALLERO, A.; QUESADA, H. Evaluating the relationship between evolutionary divergence and phylogenetic accuracy in AFLP data Sets. **Molecular Biology and Evolution**. 2010;27:988-1000. doi:10.1093/molbev/msp315
- GAUDEUL, M.; ROUHAN, G.; GARDNER, M. F.; HOLLINGSWORTH, M. P. AFLP markers provide insights into the evolutionary relationships and diversification of New Caledonian *Araucaria* species (**Araucariaceae**). **American Journal of Botany**. 2012;99:1:68-81. doi:10.3732/ajb.1100321
- JANSEN, J.; VAN HINTUM, T. H. Genetic distance sampling: a novel sampling method for obtaining core collections using genetic distances with an application to cultivated lettuce. **Theoretical Applied Genetics**. 2007;114:421. doi:10.1007/s00122-006-0433-9
- LIU, Y.; HOU, W. Genetic Diversity of Faba Bean Germplasms in Qinghai and Core Germplasm Identified based on AFLP Analysis. **Legume Genomics and Genetics**. 2010;1:1:1-6. doi:10.5376/lgg.2010.01.0001
- MALVAR, R. A.; BUTRON, A.; ALVAREZ, A.; PADILLA, G.; CARTEA, M. E.; REVILLA, P.; ORDAS, A. Yield performance of the European Union Maize Landrace Core Collection under multiple corn borer infestations. **Crop Protection**. 2007;26:775-781. doi:10.1016/j.cropro.2006.07.004
- MALOSETTI, M.; DIAZ, J.; ABADI, T. Evaluación preliminar del proceso de regeneración de la colección de maíz. Informe de actividades desarrolladas en el marco del proyecto. Sete Lagoas, MG: Embrapa Milho e Sorgo. 9 p. 2000.
- MARTINS, E. C. A.; PELUZIO, J. M.; COIMBRA, R. R.; DA SILVEIRA, M. A.; OLIVEIRA, J. D. D.; DE OLIVEIRA JUNIOR, W. P. Diversidade genética em batata-doce no Tocantins. **Bioscience Journal**. 214;30:2:429-435. < <http://www.seer.ufu.br/index.php/biosciencejournal/article/view/17999/13758>>
- MENDELSON, T. C.; WONG, M. K. AFLP phylogent of the snubnose darters and allies (Percidae: *Etheostoma*) provides resolution across multiple levels of divergence. **Molecular Phylogenetics and Evolution**. 2010;27:1253-1259. doi:10.1016/j.ympev.2010.10.007
- MEUDT, H. M.; CLARKE, A. C. Almost forgotten or latest practice? AFLP applications, analyses and advances. **Trends in plant Science**. 2007;12:3:106-117. doi:10.1016/j.tplants.2007.02.001
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics**. 1978;89:583-590. < <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1213855/>>
- NEI, M. Genetic distance between populations. **American Naturalist**. 1972;106:283-292. < <http://www.jstor.org/stable/2459777>>
- NETTO, D. A. M.; SOUZA, I. R. P.; OLIVEIRA, A. C.; PINTO, C. A. B. P.; ANDRADE, R. V. Avaliação agrônômica e molecular de acessos da coleção núcleo de milho, subgrupo endosperma duro. **Revista Brasileira de Milho e Sorgo**. 2004;3:1:92-107. < <https://www.embrapa.br/en/web/mobile/publicacoes/-/publicacao/487786/avaliacao-agronomica-e-moleculares-de-acessos-da-colecao-nucleo-de-milho-subgrupo-endosperma-duro>>
- ODONG, T. L.; VAN HEERWAARDEN, J.; JANSEN, J.; VAN HINTUM, T. J. L.; VAN EEUWIJK, F. A. Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data? **Theoretical Applied Genetics**. 2011;123:195-205. doi:10.1007/s00122-011-1576-x
- OLIVEIRA, E. C. A. M.; OLIVEIRA JUNIOR, W. P. O.; OLIVEIRA, J. D. D.; DE CASTRO, H. G. Divergência genética entre acessos de mentrasto do Estado do Tocantins através de marcadores RAPD. **Journal Of Biotechnology And Biodiversity**. 2013;4:4:290-298. < <https://sistemas.uft.edu.br/periodicos/index.php/JBB/article/view/632/367>>
- PONTES, L. C. G.; MOURA, E. F.; MOURA, M. F.; RODRIGUES, S. M.; OLIVEIRA, M. S. P.; CARVALHO, J. E. U.; THERRIER, J. Molecular characterization of progenies of bacurizeiro (*Platonia insignis*) from Marajó Island, northeastern Amazon. **Acta Amazonica**. 2017;47:4:293-300. < <https://www.embrapa.br/en/busca-de-publicacoes/-/publicacao/1078387/molecular-characterization-of-progenies-of-bacurizeiro-platonia-insignis-from-marajao-island-northeastern-amazon>>
- REBOURG, C.; DUBREUIL, P.; CHARCOSSET, A. Genetic diversity among maize populations: Bulk RFLP analysis of 65 accessions. **Maydica**. 199;44:237-249. < <http://agris.fao.org/agris-search/search.do?recordID=IT2001060966>>

- SAGHAI MAROOF, M. A.; SILIMAN, K. M.; JORGENSON, R.; ALLARD, R. W. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance chromosomal location and population dynamics. **Proceedings National Academy Sciences**. 1984;81:8014-8018 < <https://www.ncbi.nlm.nih.gov/pubmed/6096873>>
- SHARMA, S. S.; ISLAM, M. A.; SINGH, V. K.; NEGI, M. S.; TRIPATHI, S. B. Genetic diversity, population structure and association study using TE-AFLP markers in *Pongamia pinnata* (L.) Pierre germplasm. **Tree Genetics & Genomes**. 2017;13:1-6 doi:10.1007/s11295-016-1088-6
- SUDHEER PAMIDIAMARRI, D. V. N.; PANDYA, N.; REDDY, M. P. Comparative study of interspecific genetic divergence and phylogenetic analysis of genus *Jatropha* by RAPD and AFLP. **Molecular Biology Reports**. 2009;36:901-901-907. doi:10.1007/s11033-008-9261-0
- TATIKONDA, L. AFLP-based molecular characterization of an elite germplasm collection of *Jatropha curcas* L., a biofuel plant. **Plant Science**. 2009;176:4:505-513. doi:10.1016/j.plantsci.2009.01.006
- YETISIR, H.; SAKAR, M.; SERCE, S. Collection and morphological characterization of *Lagenaria siceraria* germplasm from the Mediterranean region of Turkey. **Genetic Resources and Crop Evolution**. 2008;55:1257-1266. doi:10.1007/s10722-008-9325-y
- VASCONCELOS, E. S.; CRUZ, C. D.; BHERING, L. L.; FERREIRA, A. Estratégias de amostragem e estabelecimento de coleções nucleares. **Pesquisa Agropecuária Brasileira**. 2007;42:4:507-514. doi:10.1590/s0100-204x2007000400008
- YAO, Q.; FAN, P.; ZOU, S. Constructing a Core Collection for Maize (*Zea mays* L.) Landrace from Wuling Mountain Region in China. **Agricultural Sciences in China**. 2008;7:12:1423-1432. doi:10.1016/S1671-2927(08)60398-3
- YAN, J.; SHAH, T.; WARBURTON, M.L.; BUCKLER, ES; McMullen, MD; Crouch, J. Genetic Characterization and Linkage Disequilibrium Estimation of a Global Maize Collection Using SNP Markers. **PLoS ONE**. 2009;4:12:1-14. doi:10.1371/journal.pone.0008451