Molecular characterization of common bean accessions using microsatellite markers

Caracterização molecular de acessos de feijão comum por meio de marcadores microssatélites

Taniele Carvalho de Oliveira[1] , Marco Antonio Aparecido Barelli[1] , Altacis Junior de Oliveira[1]* ,
Nilo Sander Leal[1] , Thiago Alexandre Santana Gilio[2] , Rafhael Felipin Azevedo[3]

[1]Universidade do Estado de Mato Grosso/UNEMAT, Faculdade Ciências Agrárias e Ambientais, Cáceres, MT, Brasil
[2]Universidade Federal de Mato Grosso/UFMT, Instituto de Ciências Agrárias e Ambientais, Sinop, MT, Brasil
[3]Centro Universitário de Várzea Grande/UNIVAG, Departamento de Agronomia, Várzea Grande, MT, Brasil
*Corresponding author: altacismarquesfig@hotmail.com
Received in March 3, 2023 and approved in June 19, 2023

ABSTRACT

The common bean, a legume of significant economic importance, is renowned for its extensive genetic variability. It is crucial to comprehend genetic diversity, analyze population structure, and understand relationships among commercial classes of accessions to facilitate genetic improvement. This study aimed to molecularly characterize 143 common bean accessions by employing 25 SSR molecular markers. The objectives were to estimate genetic diversity, analyze genetic structure, and cluster populations using the UPGMA and PCoA methods. A total of 105 alleles were amplified using microsatellite loci, and the observed heterozygosity was lower than expected across all loci, indicating inbreeding within the populations. Among the loci, 22 were highly informative, demonstrating their effectiveness and polymorphism in detecting genetic diversity. The genetic variability within the population was found to be the highest, while variation between populations was the lowest. The analysis of population structure revealed the presence of three populations with a notable rate of gene introgression. The UPGMA analysis categorized the accessions into 15 groups, but they did not form distinct clusters based on their geographic regions or gene pool. The first two principal coordinates accounted for 13.95% of the total variation among the accessions. The SSR markers employed effectively detected genetic variability among the common bean accessions, revealing that their genetic diversity was not correlated with their geographic distribution in this study.

Index terms: *Phaseolus vulgaris* L.; genetic variability; molecular marker; germplasm.

RESUMO

O feijão comum é uma das leguminosas de importância econômica e conhecida por sua ampla variabilidade genética. O conhecimento da diversidade genética, a análise da estrutura populacional e o entendimento das relações dentro e entre as classes comerciais a partir dos acessos são passos essenciais no melhoramento genético. O objetivo deste trabalho foi realizar a caracterização molecular de 143 acessos de feijoeiro usando 25 marcadores moleculares SSR para estimativa da diversidade genética, análise da estrutura genética e agrupamento das populações pelos métodos UPGMA e PCoA. Um total de 105 alelos foram amplificados usando loci microssatélites, e a heterozigosidade observada foi menor do que a esperada em todos os loci. O índice de fixação indicou endogamia dentro das populações. Dentre os locos, 22 foram considerados altamente informativos, mostrando-se eficazes e polimórficos na detecção da diversidade genética. A maior variabilidade genética observada dentro das populações, enquanto a menor variação entre as populações. A análise da estrutura populacional mostrou a presença de três populações, entre as quais foi observada uma alta taxa de introgressão de genes. A análise UPGMA permitiu a divisão dos acessos em 15 grupos, mas os acessos não formaram grupos distintos de acordo com suas regiões geográficas ou pool gênico. As duas primeiras coordenadas principais explicaram 13,95% da variação total entre os acessos. Os marcadores SSR utilizados mostraram-se eficientes em detectar a variabilidade genética entre os acessos, cuja diversidade genética demonstrou que a distribuição geográfica neste estudo não esteve relacionada à distância genética entre eles.

Termos para indexação: *Phaseolus vulgaris* L.; variabilidade genética; marcador molecular; germoplasma.

INTRODUCTION

The common bean (*Phaseolus vulgaris* L.) is indigenous to Central and South America, and its domestication occurred independently in the Andean and Mesoamerican regions, resulting in two distinct gene pools (Voysest, 2000; Gioia et al., 2013; Shamseldin;

Velázquez, 2020). Among the legumes belonging to the *Phaseolus* genus, the common bean holds significant global importance. In Brazil, it is a staple in daily diets and a primary source of vegetable protein (Pitura; Arntfield, 2019; Melo; Paixão, 2020).

This crop is renowned for its morphological variability and adaptability to diverse environments, resulting in a wide range of local varieties or landraces, each with distinct seed properties (Perina et al., 2014; Pereira et al., 2019). It is essential to characterize the genetic variability of common bean accessions from Germplasm Banks and working collections. Such studies facilitate understanding of the genetic relationships among accessions, identification of redundancies and mixtures in the germplasm, and determination of optimal genetic distances between pairs of parents (Coelho et al., 2010; Delfini et al., 2021).

Molecular markers are valuable tools for characterizing species, especially when morphologically similar individuals with agronomic potential need to be distinguished (El Kadri; Mimoun; Hormanza, 2019). Molecular markers, including microsatellites or simple sequence repeats (SSRs), have become the markers of choice for studying genetic diversity, species identification, phylogenetic analysis, gene mapping, and classification of genetic resources in various crops, including the common bean (Alghamdi et al., 2019).

SSR markers are frequently employed in common bean studies due to their abundant and relatively uniform distribution in the genome, codominant inheritance, high polymorphism, and reproducibility levels, and ease of analysis and comparison across studies and germplasm sets (Gioia et al., 2019). Previous studies conducted by Šajgalik et al. (2019), Mir et al. (2021), Oliveira et al. (2021), and Catarcione et al. (2023) have demonstrated that SSR markers are an effective genetic tool for assessing genetic diversity and subpopulation structure in common bean germplasm collections. Therefore, the objective of this study was to estimate the genetic diversity of common bean accessions using microsatellite markers.

MATERIAL AND METHODS

The research was carried out at the Laboratory of Genetic Resources and Biotechnology (LRG&B), located at the Cáceres, Mato Grosso State Campus of the State University of Mato Grosso "Carlos Alberto Reyes Maldonado" (UNEMAT), with coordinates at 16°05'08.1" latitude in the South and 57°39'00.3" longitude in the West.

A comprehensive assessment was conducted on a total of 143 accessions of the common bean sourced from the germplasm collection of the LRG&B (Table 1).

The common bean accessions were planted in trays filled with the commercial substrate within the LRG&B greenhouse. Once the seedlings reached the V2 stage, characterized by the partial opening of the first trifoliate leaves, they were transferred to 2 mL microtubes. The collected plant samples were then taken to the LRG&B and stored in an ultra-freezer at –80 °C.

DNA extraction was conducted using the Promega *Wizard® Genomic DNA Purification* kit following the manufacturer's guidelines. The trifoliate leaves of the common bean were retrieved from the microtubes, and two beads were added to each tube. Approximately 600 µL of "*Nucleilysis Solution*" was then introduced into each microtube, which was macerated for 15 min using the tissue Lyser.

After the maceration period, the beads were removed from each microtube, and the tubes were incubated in a water bath at 65 °C for 15 min. Subsequently, 3 µL of RNAse was added gently to each microtube five times, and the microtubes were further incubated in a water bath at 37 °C for 15 min. Following this, the tubes were allowed to cool for 5 min at room temperature, and 200 µL of "*Protein Precipitation Solution*" was added. The microtubes were then vortexed at high speed for 20 s.

The microtubes were placed in a centrifuge at –4 °C and 14,000 rpm for 5 min, and the resulting supernatant was transferred to new 1.5 mL microtubes containing 600 µL of isopropanol. The solution was mixed by inversion at room temperature.

Subsequently, the microtubes were centrifuged again at 22 °C and 14,000 rpm for 5 min, and the supernatant was discarded. Then, 600 µL of 70% ethanol was added to the precipitate, and the microtubes were centrifuged at room temperature for 3 min at 14,000 rpm. The supernatant was removed by decanting, leaving only the pellet attached to the microtube wall. The microtubes were placed on absorbent paper for 15 min to dry the precipitate.

DNA was extracted from all common bean accessions, and the precipitated DNA was resuspended in autoclaved ultrapure water and refrigerated overnight for future use. DNA quality was assessed using electrophoresis on a 1.0% agarose gel stained with Gel Red and Blue Juice 6X, with 1% Tris borate EDTA used as the running buffer. The gel was run at 60 Volts for 1 h to ensure the extraction of high-quality DNA.

Table 1: Identification of a total of 143 accessions of common bean from the germplasm collection of LRG&B – UNEMAT.

N°	Common name	Gene pool[1]	Pop[2]	N°	Common name	Gene pool[1]	Pop[2]	N°	Common name	Gene pool[1]	Pop[2]
1	Miranda 72 - Peruano	M	AC	49	Branquinho 60 dias	M	TO	97	CHC 98–42	M	SC
2	Miranda 65	A	AC	50	Santa Helena	M	TO	98	CHP 01–238	M	SC
3	Miranda 65 - Peruano	A	AC	51	Enxofre	M	TO	99	CNFC 10762	M	GO
4	Bolinha	A	AC	52	Gurgutuba Vermelho	A	AM	100	CNFP 10794	M	GO
5	Jalinho Amarelo Claro	M	PA	53	Enxofre	A	AM	101	FT 08–75	M	PR
6	Rosinha Pé Curto	A	AC	54	Feijão Ashaninka	M	AC	102	LEC 01–11	M	PR
7	Paco Mineiro	A	AC	55	Branco - MT	M	MT	103	LEP 02–11	M	PR
8	Carioca (Mistura Varietal)	A	AC	56	UNEMAT-18	M	MT	104	LP 09–192	M	PR
9	Pardo (Mistura Varietal)	A	AC	57	UNEMAT-28	M	MT	105	LP 09–40	M	PR
10	Rosinha	M	AC	58	UNEMAT-11	A	MT	106	TB 02–23	M	RS
11	Bico de Ouro	A	AC	59	UNEMAT-27	M	MT	107	TB 03–13	A	RS
12	Peruano	M	AC	60	UNEMAT-22	M	MT	108	Rosinha Opaco	M	MS
13	Rosinha	M	AC	61	UNEMAT-40	M	MT	109	Rosinha A	M	MS
14	Mulatinho	M	AC	62	UNEMAT-17	M	MT	110	Roxinho A	M	MS
15	Mulatinho	A	AC	63	UNEMAT-32	M	MT	111	Carioquinha Limpo	M	MS
16	Peruano	A	AC	64	UNEMAT-04	M	MT	112	Rosado	M	MS
17	Jalinho	A	PA	65	UNEMAT-30	M	MT	113	Roxinho B	M	MS
18	Carioca	A	MT	66	UNEMAT-03	M	MT	114	Mulatão Lustroso	M	MS
19	Rosinha	M	MT	67	UNEMAT-06	M	MT	115	Bico de Ouro A	M	MS
20	Rosinha Cipó	M	MT	68	UNEMAT-29	M	MT	116	Mulatinho Vagem Roxa A	M	MS
21	Carioca	A	MT	69	UNEMAT-35	M	MT	117	Carioca Vagem Rosada	A	MS
22	Rosinha	M	MT	70	UNEMAT-42	M	MT	118	Jalo	A	MS
23	Rosinha	M	MT	71	UNEMAT-23	M	MT	119	Manteiguinha	A	MS
24	Carioca	A	MT	72	UNEMAT-26	M	MT	120	Uberabinha Preto	M	MS
25	Carioca Pé Curto	M	MT	73	UNEMAT-31	M	MT	121	Manteiguinha de Cipó	A	MS
26	Carioca	A	MT	74	UNEMAT-37	M	MT	122	Carioca Novo	M	MS
27	Carioca	A	MT	75	UNEMAT-38	A	MT	123	Jalo sem Cipó	A	MS
28	Rosinha	M	MT	76	UNEMAT-75	M	MT	124	Carioca sem Cipó	M	MS
29	Carioca	A	MT	77	UNEMAT-01	M	MT	125	Bodequena	A	MS
30	Carioca	A	MT	78	UNEMAT-56	M	MT	126	Chita Bonita	A	MS

Continue.

Table 1: Continuation.

N°	Common name	Gene pool[1]	Pop[2]	N°	Common name	Gene pool[1]	Pop[2]	N°	Common name	Gene pool[1]	Pop[2]
31	Rosado	M	MT	79	UNEMAT-50	M	MT	127	Rosinha sem Cipó	M	MS
32	Manteiguinha	M	MT	80	UNEMAT-49	M	MT	128	Manteiga sem Cipó	A	MS
33	Preto Pinheirinho	M	MT	81	UNEMAT-45	M	MT	129	Manteiga com Cipó	A	MS
34	Rosinha	M	MT	82	UNEMAT-46	M	MT	130	Carioca com Cipó	M	MS
35	Preto Caeté	M	MT	83	UNEMAT-72	M	MT	131	Bolinha	A	MS
36	Tibagi	M	MT	84	UNEMAT-55	M	MT	132	Mulatinho Vagem Roxa B	M	MS
37	Bico de Ouro	A	MT	85	UNEMAT-16	M	MT	133	Rosinha B	M	MS
38	Carioca	A	MT	86	UNEMAT-60	M	MT	134	Roxinho Mineiro	M	MS
39	Paquinha	M	MT	87	UNEMAT-58	M	MT	135	Preto Guamirim	M	MS
40	Carioca	M	MT	88	UNEMAT-07	A	MT	136	Manteigão	A	MS
41	Rosado	M	MT	89	UNEMAT-13	M	MT	137	Rosinha C	M	MS
42	Paquinha	M	MT	90	UNEMAT-20	M	MT	138	Rosinha Guaicurus	M	MS
43	Paquinha	M	MT	91	UNEMAT-21	M	MT	139	Cara Suja	M	MS
44	Carioca	A	MT	92	Pérola	M	GO	140	Manteiga	A	MS
45	Rosinha	M	MT	93	IPR Uirapuru	M	PR	141	Bico de Ouro B	M	MS
46	Roxinho Mineiro	M	MT	94	CNFP 10104	M	GO	142	Rosinha D	A	MS
47	Rosinha de 1984	M	MT	95	C 4–7–7–2–2	M	SP	143	BRS FC104	M	GO
48	Preto	M	TO	96	C 4–7–8–1–2	M	SP				

[1]A: Andean e M: Mesoamerican; [2]Population = AC: Acre, AM: Amazonas, GO: Goiás, MS: Mato Grosso do Sul, MT: Mato Grosso, PA: Pará, PR: Paraná, RS: Rio Grande do Sul, SC: Santa Catarina, SP: São Paulo, and TO: Tocantins.

To assess the genetic variability of common bean accessions from the collection, a set of 25 primers (Table 2) was employed for amplifying the SSR loci, as previously established by Valentini et al. (2018). DNA samples for the PCR were diluted to a concentration of 10 ng-µL^{-1} using autoclaved ultrapure water. The PCR mixture was prepared with a final reaction volume of 25 µL, comprising 2 µL of DNA (10 ng-µL^{-1}), 0.5 µL of deoxyribonucleotide mix (dATP, dCTP, dGTP, and dTTP) (10 mM), 1.25 µL of each forward and reverse primer (10 µM), 5 µL of buffer (5X) containing magnesium (7.5 mM), 0.2 µL of Taq polymerase (5U), and 14.8 µL of autoclaved ultrapure water.

The polymerase chain reaction (PCR) was conducted following the methodologies outlined by Williams et al. (1990), Gaitán-Solís et al. (2002), and Blair et al. (2003) using the Perkin Elmer Model 9600 thermocycler. Each PCR cycle consisted of an initial denaturation step at 94 °C for five minutes, followed by 35 denaturation cycles at 94 °C for 30 s. The annealing step was performed at the specific primer temperature (1 to 18) for 30 s, followed by an extension phase at 72 °C for 30 s. The final extension phase lasted five minutes at 72 °C.

For primers 19 to 23, the cycle began with an initial denaturation step at 92 °C for five minutes, followed by 30 cycles of denaturation at 92 °C for one minute. The annealing step was performed at the specific primer temperature for one minute, followed by an extension phase at 72 °C for two minutes. The final extension phase was conducted at 72 °C for five minutes. Primers 24 and 25 followed a cycle consisting of an initial denaturation step at 94 °C for two minutes, followed by 35 cycles of denaturation at 94 °C for 15 s. The annealing step was performed at the specific primer temperature for 15 s, followed by an extension phase at 65 °C for 15 s. The final extension phase lasted 15 s at 72 °C.

Table 2: SSR primers and sequence information used for genetic diversity analysis of common bean accessions.

N°	Locus	Repeat type	RA (bp)[1]	TA (°C)[2]	Primer sequence
1	PV-cct001	$(CCT)_7$	149	56.9	CCAACCACATTCTTCCCTACGTC
					GCGAGGCAGTTATCTTTAGGAGTG
2	PV-at006	$(AT)_5$	132	52.4	CCGTTGCCTGTATTTCCCCAT
					CGTGTGAAGTCATCGGAGTGGTC
3	BM143	$(GA)_{35}$	143	54.5	GGGAAATGAACAGAGGAA A
					ATGTTGGGAACTTTTAGTGTG
4	BM158	$(CT)_{22}$	130	57.5	CCGAGCACCGTAACTGAATGC
					CGCTCGCTTACTCACTGTACGC
5	BM170	$(CT)_5CCTT(CT)_{12}$	179	52.1	AGCCAGGTGCAAGACCTTAG
					AGATAGGGAGCTGGTGGTAGC
6	BM184	$(AC)_{11}$	160	50.5	AGTGCTCTATCAAGATGTGTG
					ACATAATCAATGGGTCACTG
7	BM187	$(CT)_{10}T(CT)^1_4$	191	48.1	TTTCTCCAACTCACTCCTTTCC
					TGTGTTTGTGTTCCGAATTATGA
8	PV-BR25	$(CT)_{23}$	158	50.5	GAGCTTCTCCGTCCTGTGT
					CGAACTGAATCAGAAAGGAA
9	PV-BR35	$(TC)_{10}$	214	52.1	TCTACGCGTTCCCTCTGTCT
					AGTGGATGTGTGGGAAAAGC
10	PV-BR60	$(CT)_9$	171	50.2	CTTAACAACCCCTCCGACAA
					AGTGTGAAAACCGCCCTCT
11	PV-BR112	$(TC)_{13}$	183	48.1	AACAACACCACCCTGGAGAC
					ACAAAAGCGAGGAATCACG
12	PV-BR167	$(AT)_7(AG)_9$	158	43.8	GGCAAAAACAAAACCATTTCA
					GCCATTTCTCCACTGTCTGG
13	PV-BR185	$(TC)_{11}$	153	50.5	TGGTAAAGCAAA ACGATGG
					GACAGAAGAGTGAGGGTGTGAA
14	SSR-IAC10	$(GA)_{12}(AG)_6(AG)_6$	290	50.5	AGGAACTAAAAGCCGAACTGG
					GCCTCCGCCGATCAACACTA
15	SSR-IAC14	$(GT)_7$	221	48.1	GCTGCATGTTTATCCACCTT
					TTGTTACTCACCCCACCATAC
16	SSR-IAC24	$(AC)_7(AT)_6$	165	43.8	TTGGGAAAATTATAGAGAACA
					AGCCACTGACCCTTACAT
17	SSR-IAC62	$(AG)_{14}$	211	49.5	AACCCGTGAATCTTTGAGG
					ATTGATGGTGGATTTTGAA
18	SSR-IAC74	$(CA)_9 (TA)_7$	271	48.1	GGAATCGAAGTTTGAAGTGAGG
					AAATGACCAAGCCAAGAATGTT

Table 2: Continuation.

N°	Locus	Repeat type	RA (bp)[1]	TA (°C)[2]	Primer sequence
19	BMd-12	$(AGC)_7$	167	52.1	CATCAACAAGGACAGCCTCA
					GCAGCTGGCGGGTAAAACAG
20	BMd-20	$(TA)_5$	123	55.7	GTTGCCACCGGTGATAATCT
					GTGAGGCAAGAAGCCTTCAA
21	BMd-25	$(GAT)_6$	118	55.7	GCAGATCGCCTACTCACAAA
					CGTTGACGAGAAGCATCAAG
22	BMd-26	$(GAT)_6$	141	53.2	CTTGCCTTGTGCTTCCTTCT
					TCCATTCCCAACCAAGTTTC
23	BMd-36	$(TA)_8$	164	53.0	CATAACATCGAAGCCTCACAGT
					ACGTGCGTACGAATACTCAGTC
24	BM152	$(GA)_{34}$	127	50.5	CCGGGACTTGCCAGAAGAAC
					AAGAGGAGGTCGAAACCTTAAATCG
25	BM210	$(CT)_{15}$	166	50.5	CCCTCATCCTCCATTCTTATCG
					ACCACTGCAATCCTCATCTTTG

[1]Amplification region in base pairs (bp); [2]Annealing temperature.

The amplified PCR products were visualized on a 3% agarose gel stained with 6X Gel Red and Blue Juice. Electrophoresis was conducted using 1% Tris borate EDTA as the running buffer, and the gel was run at 60 Volts for 5 h. The gels were photographed using the Locus Biotechnology/LPix Image photo-documentation system, version 2.7. The heights of each allele (band) were measured using the LabImage 1D Revision 1.10 program (Locus Biotechnology), and DNA bands were compared to a 1000 base pair (bp) standard ladder.

Genetic diversity was estimated based on various parameters, including the number of alleles per locus (Na), observed (Ho) and expected average heterozygosity (He), inbreeding coefficient (f), and polymorphic information content (PIC), following the methods of Weir and Cockerham (1984). These data were analyzed using the Power Marker program, version 3.25 (Liu; Muse, 2005), and used to calculate the genetic distance matrix among the 143 accessions, following the approach of Nei, Tajima and Tateno (1983). The genetic distance matrix was then imported into the Genes program (Cruz, 2016) to construct a dendrogram using the Unweighted Pair Group Method with Arithmetic Average (UPGMA).

To analyze the distribution of genetic diversity between and within populations, the analysis of molecular variance (AMOVA) was performed, as described by Excoffier, Smouse and Quattro (1992). The significance of the differences between means was tested using 1,000 permutations with a 95% confidence interval. Additionally, genetic diversity was evaluated using principal coordinate analysis (PCoA) at the accession level, employing the GenAlEx6.5 software (Peakall; Smouse, 2012).

The Structure program (Pritchard; Stephens; Donnelly, 2000), employing Bayesian statistics, was utilized to estimate the number of groups (K). The analyses involved 20 runs for each K value, with a burn-in period of 5,000 and 100,000 Markov Chain Monte Carlo (MCMC) simulations. The criteria proposed by Evanno, Regnaut and Goudet (2005) were applied to determine the most probable value of K based on the groups defined by the program.

RESULTS AND DISCUSSION

Utilizing 25 SSR primers for genotyping 143 common bean accessions, a total of 105 alleles were successfully amplified, with an average of 4.20 alleles per primer. The maximum number of alleles per primer was seven (PV-cct001 and PV-BR185), while the minimum was three (BM143, BM187, PV-BR025, PV-BR60, PV-BR112, PV-BR167, SSR-IAC14, and BMd-20). Each evaluated accession demonstrated the presence of at least three alleles when assessed with all primers (Table 3).

Table 3: Genetic diversity of 143 common bean accessions assessed using 25 microsatellite markers.

SSR Locus	N_a	H_e	H_o	f	PIC
PV-cct001	7	0.7695	0.0070	0.9910	0.7334
PV-at006	4	0.2677	0.0000	1.0000	0.2418
BM143	3	0.6504	0.0000	1.0000	0.5751
BM158	4	0.4878	0.0000	1.0000	0.4470
BM170	6	0.7296	0.0000	1.0000	0.6810
BM184	4	0.6661	0.0000	1.0000	0.6037
BM187	3	0.6666	0.0000	1.0000	0.5925
PV-BR025	3	0.6130	0.0000	1.0000	0.5310
PV-BR35	5	0.7342	0.0000	1.0000	0.6898
PV-BR60	3	0.5892	0.0071	0.9880	0.5237
PV-BR112	3	0.4975	0.0000	1.0000	0.4422
PV-BR167	3	0.6289	0.0000	1.0000	0.5575
PV-BR185	7	0.6879	0.0000	1.0000	0.6319
SSR-IAC10	4	0.6793	0.0500	0.9269	0.6266
SSR-IAC14	3	0.5913	0.0000	1.0000	0.5100
SSR-IAC24	4	0.6331	0.0000	1.0000	0.5575
SSR-IAC62	5	0.7076	0.0000	1.0000	0.6591
SSR-IAC74	4	0.6136	0.0000	1.0000	0.5520
BMd-12	4	0.6994	0.0000	1.0000	0.6459
BMd-20	3	0.6195	0.0000	1.0000	0.5387
BMd-25	5	0.6961	0.0000	1.0000	0.6419
BMd-26	4	0.7111	0.0000	1.0000	0.6604
BMd-36	5	0.6706	0.0000	1.0000	0.6128
BM152	4	0.7087	0.0000	1.0000	0.6559
BM210	5	0.6959	0.0000	1.0000	0.6399
Media	4.20	0.6406	0.0026	0.9960	0.5821

N_a = The number of alleles; H_e = Expected heterozygosity; H_o = Observed heterozygosity; f = fixation index (endogamy); PIC = Polymorphic information content.

The observed heterozygosity was consistently lower than expected across all loci, ranging from 0.0000 to 0.0500, with a mean of 0.0026. In contrast, the expected heterozygosity ranged from 0.2677 to 0.7695, with a mean of 0.6406. These high expected heterozygosity values, coupled with low observed heterozygosity values, suggest a substantial degree of genetic diversity among the accessions and a notable inbreeding rate, as indicated by the elevated fixation indices estimated from each analyzed locus.

Gioia et al. (2019) assessed the degree of genetic variance and relatedness among 192 advanced cultivars of common bean, using 58 SSR markers. Their findings indicated that the anticipated heterozygosity exceeded the observed heterozygosity. On the contrary, Savić et al. (2021) performed an examination of the genetic diversity inherent in Serbia's common bean germplasm, employing 27 SSR markers. They noticed that for 17 primers, the observed heterozygosity surpassed the expected heterozygosity, suggesting a surplus of heterozygotes within the population.

It is significant to note an apparent excess of homozygotes among the common bean populations. This is evidenced by the positive fixation coefficient

observed across all loci, with an average of 0.9960. This statistic points to the occurrence of inbreeding within the populations. These findings align with those from a study by Catarcione et al. (2023), where they investigated the genetic diversity and population structure of common bean landraces in Italy's Lazio region, they observed positive values ranging from 0.951 to 1.

The PIC value varied between 0.2418 (PV-at006) and 0.7334 (PV-cct001), with an average of 0.5821. Out of the loci evaluated, 22 were deemed highly informative, whereas two were identified as moderately and slightly informative. Overall, the SSR loci under evaluation proved to be efficient and polymorphic, reflecting the genetic diversity within the common bean accessions. It is important to understand that the PIC value correlates with the count of alleles, which in itself is directly linked to the genetic divergence and the quantity of accessions evaluated in the study.

Findings comparable to those previously discussed were reported by Vidak et al. (2021), who investigated the origin and diversity of the Croatian common bean germplasm. They employed phaseolin type, SSR and SNP markers, along with morphological traits, and reported PIC values spanning from 0.310 to 0.862, with an average of 0.497. These results are consistent with those reported by Catarcione et al. (2023), where observed values ranged from 0.263 to 0.868, and an average value of 0.572 was calculated.

To measure the variation both between and within populations, AMOVA was undertaken. This showed that the most significant variability, accounting for 84% of the total, was detected within populations. In contrast, the least variation, contributing 16% to the total, was discerned between different populations. Additionally, the F_{ST} index, a measure of genetic differentiation between populations, was determined to be low (F_{ST} = 0.157), indicative of a relatively low degree of population structure (Table 4).

These results are in concordance with the relatively closed reproduction in this specie, similar results were reported by Özkan et al. (2022) on the genetic diversity and population structure of common bean landraces in Turkey using SSR markers, AMOVA results revealed that the within-population variance (66%) was higher than the between population variance (34%).

The Bayesian model-based clustering method was employed to evaluate the population structure of 143 common bean accessions. The number of subpopulations (known as the K value) was identified based on maximum likelihood and Delta K values. This approach divided the accessions into three distinct groups (K = 3) (Figure 1). Group I contained several Andean accessions, while Group III incorporated several other Mesoamerican accessions. Group II was composed of a mixture of gene pools.

Specifically, Group I was comprised of 19 accessions (18 Andean and 1 Mesoamerican) sourced from the Acre, Amazonas, Mato Grosso, Mato Grosso do Sul and Rio Grande do Sul, populations. Group II contained 57 accessions from the Acre, Amazonas, Goiás, Mato Grosso, Mato Grosso do Sul, Pará, and Tocantins populations, including 20 Andean and 37 Mesoamerican. The most populous was Group III, with 67 accessions (66 Mesoamerican and 1 Andean), originating from the Goiás, Mato Grosso, Mato Grosso do Sul, Paraná, Santa Catarina, and São Paulo populations.

Corroborative results were evident in a study conducted by Delfini et al. (2021), who explored the population structure, genetic diversity, and genomic selection within the Brazilian common bean germplasm. Their observations revealed a tripartite division of accessions (K = 3). Conversely, Vidak et al. (2021) studied the origin and diversity of Croatian common bean germplasm using phaseolin types, SSR and SNP markers, and morphological traits. Their population structure analysis yielded K values of 2 and 3 for the different molecular markers, respectively.

Table 4: Analysis of molecular variance (AMOVA) of 11 common bean populations using 25 microsatellite markers.

Source of Variation	DF	SQ	CV	VT (%)	P
Between populations	10	342.520	1.344	16	<0.000
Within populations	275	1986.008	7.222	84	
Total	285	2328.528	8.566	100	
Fixation Index	F_{ST}	0.157			

DF = degree of freedom; SQ = Sum of squares; CV = Variance components; VT = Total variance; P = probability of having a variance component greater than the values observed randomly; The probabilities were calculated based on 1000 random permutations.
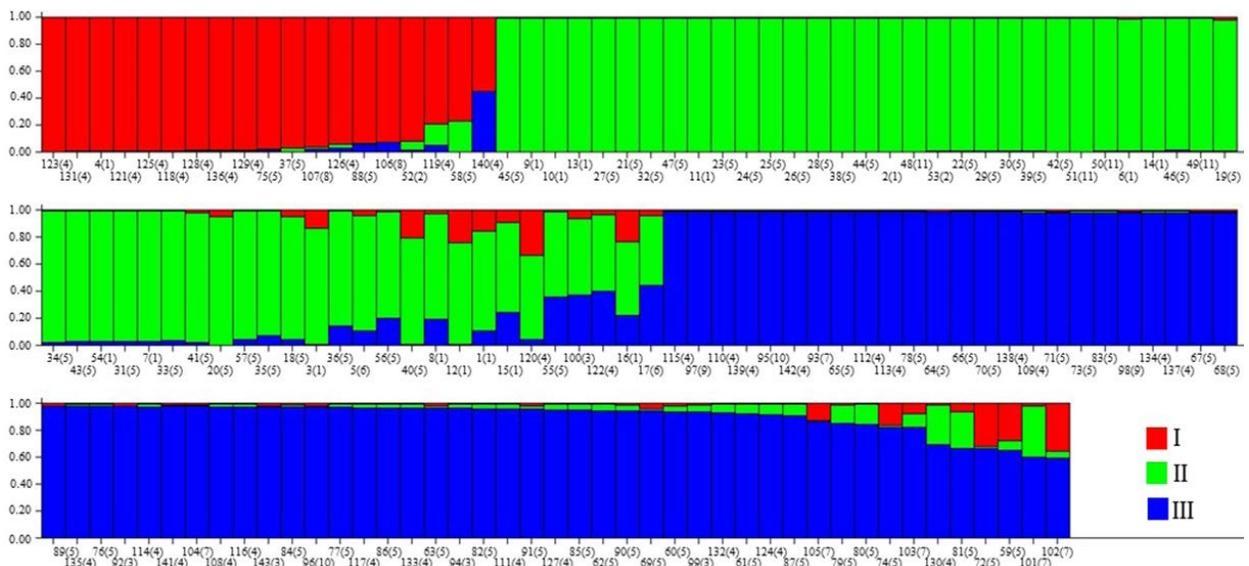
Figure 1: Population structure of 143 common bean accessions evaluated based on 25 SSR molecular markers (K = 3); Each vertical bar represents access and the percentage of adhesion to each group.

Introgression presence is correlated with accessions, which are depicted by varying colors within the same bar plot (Figure 1). Among the 88 accessions evaluated for introgression, the most pronounced intensity was noted for accession 16 (Peruano) in Group II. Moreover, it was found that Mesoamerican accessions exhibited the highest rate of introgression.

As Blair et al. (2013) highlight, the notion of pure gene pools for the common bean germplasm from Brazil is not necessarily accurate, given that many accessions result from introgression. Thus, certain Brazilian varieties (jalo and carioca) cannot be considered representatives of a single gene pool. Almeida et al. (2020) assessed genetic diversity, population structure, and Andean introgression in common bean cultivars, noting the presence of introgression among evaluated accessions following population structure analysis. Here, Mesoamerican accessions displayed the highest rate of introgression relative to their Andean counterparts.

Hierarchical clustering analysis, employing the UPGMA method, unveiled the formation of 15 groups based on a similarity criterion of 78% (Figure 2). Group I was the largest, comprising 50 accessions from states such as Acre, Mato Grosso, Mato Grosso do Sul, Pará, Paraná, and Rio Grande do Sul. Within this group,

41 were Mesoamerican and 9 were Andean. This was followed by Group V, housing 25 accessions from Acre, Amazonas, Mato Grosso, Mato Grosso do Sul, and São Paulo populations, among which 17 were Andean and 8 Mesoamerican, as well as Group IV, with 14 accessions from Mato Grosso and Mato Grosso do Sul, including 5 Andean and 9 Mesoamerican.

Groups IX and XII each consisted of 10 accessions, with those of IX coming from Mato Grosso and Tocantins, and those of XII from Acre and Mato Grosso. All accessions of Group IX were Mesoamerican, while in Group XII, only one was Andean, with the rest being Mesoamerican. Group II encompassed 9 accessions from Mato Grosso do Sul and Paraná, 6 of which were Mesoamerican and 3 Andean, whereas the accessions within Group XI hailed from Mato Grosso and Mato Grosso do Sul populations, including 4 Andean and 2 Mesoamerican.

Groups X and XIV each contained 5 accessions. Those from Group X originated from Acre, Mato Grosso, and Santa Catarina, while those from Group XIV were from Mato Grosso and Mato Grosso do Sul. In both groups, 4 accessions were Mesoamerican, with just 1 being Andean. Groups VI and VIII comprised 4 accessions each, deriving from Goiás and Mato Grosso do Sul (all Mesoamerican), and Mato Grosso (1 Andean and 3 Mesoamerican), respectively.
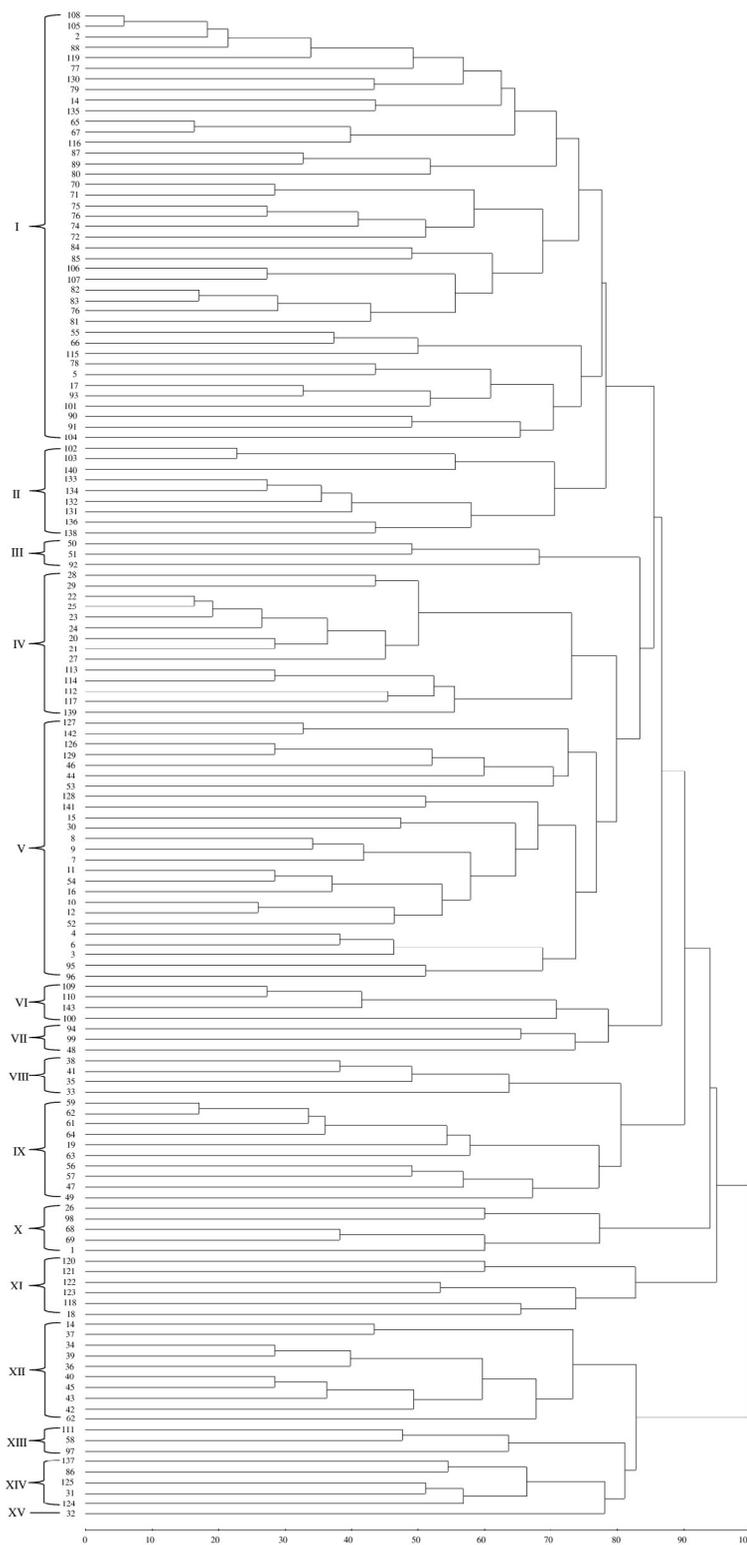
Figure 2: Dendrogram for 143 common bean accessions obtained by the UPGMA hierarchical clustering method using SSR markers.

Within the 15 groups established, 3 groups were each made up of 3 accessions (III, VII, and XIII). Accessions within groups III and VII originated from Goiás and Tocantins populations and were all Mesoamerican. In contrast, the ones in Group XIII were from the states of Mato Grosso, Mato Grosso do Sul, and Santa Catarina, consisting of 1 Andean and 2 Mesoamericans. Uniquely, Group XV was comprised of a single Mesoamerican accession from Mato Grosso, suggesting its distinct divergence from the other evaluated groups. The creation of groups containing only a single accession implies greater divergence when contrasted with others, offering potential avenues for the execution of improvement programs.

It was noted that accessions did not cluster specifically based on their geographical regions of acquisition/collection or their gene pool. This observation aligns with the findings reported by Özkan et al. (2022), who found no significant correlation between geographical origin and genetic similarity in their cluster analysis of genotypes, indicating a recent introduction from a common source population.

Ekbiç and Hasancaoğlu (2019) carried out the morphological and molecular characterization of 33 local common bean genotypes from the Ordu province in Turkey via SSR markers. Their UPGMA cluster analysis, based on geography (city), revealed geographical separation of genotypes due to local bean producers typically producing their own seeds, with seeds from a single locality being supplied to various local markets.

The clustering method was implemented according to the cophenetic correlation coefficient (CCC) via the t-test. The UPGMA hierarchical clustering method revealed a significant value (P ≤ 0.01), with (r ≥ 0.66); the distortion of 1.86% was deemed minimal, and the attained stress level (13.65%) suggested a high degree of accuracy in the graphic projection of genetic distance among populations. These results established that the generated dendrogram suitably mirrored the genetic dissimilarity matrix, thus, validating the efficacy of the statistical methods utilized and illustrating the dissimilarity among the studied common bean populations using SSR markers.

The CCC is utilized to assess the reliability of the generated clusters. Vaz Patto et al. (2004) posited that values exceeding 0.56 can be deemed ideal. Santos et al. (2022) evaluated genetic diversity among local varieties and common bean cultivars, observing significant CCC values of 0.77 (P ≤ 0.01) based on the t-test, indicating that the dendrogram suitably mirrored the genetic dissimilarity matrix.

Employing principal coordinate analysis (PCoA), it was feasible to discern the spatial distribution of the 11 populations and the 143 common bean accessions within those populations. The first two coordinates accounted for 13.95% of the total variation among accessions, with dimensions 1 and 2 accounting for 8.18% and 5.77%, respectively (Figure 3).

Accessions from Paraná, Rio Grande do Sul, Santa Catarina, São Paulo, and Tocantins populations coalesced, signifying population homogeneity, while other populations appeared separately on the graph. This suggests internal divergence within the population itself, as some accessions reside at the top and others at the bottom of the axis, implying considerable genetic variation within populations. These results echo the findings of Gyang, Muge, and Nyaboga (2020), who investigated the
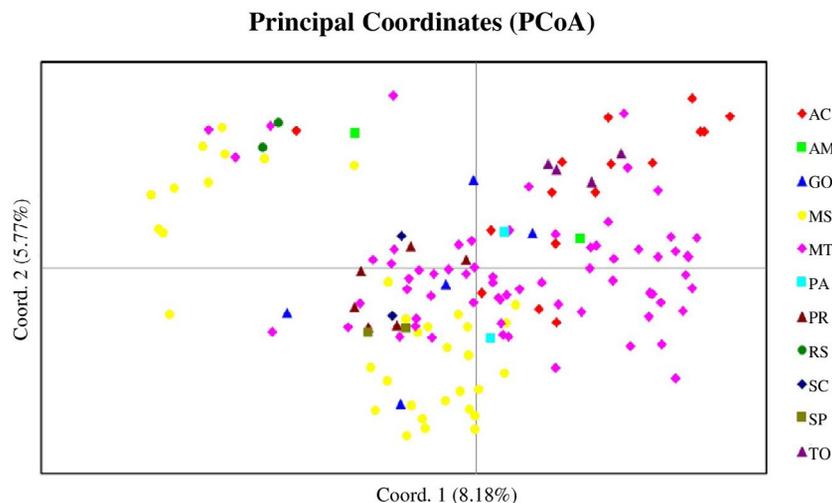
Figure 3: Analysis of principal coordinates obtained using 25 SSR markers in 143 common bean accessions.

genetic diversity and population structure of common bean germplasm in Kenya. They noticed that genotypes did not cluster into distinct groups, thus, providing evidence of population intermingling in the scatter plot of the principal coordinate analysis (PCoA).

The PCoA analysis outcomes endorsed the genetic variability as inferred from AMOVA and the UPGMA clustering analysis, illustrating a blend of populations. Furthermore, the populations Acre, Goiás, Mato Grosso, Mato Grosso do Sul, and Pará coalesced based on both gene pools (Andean and Mesoamerican), hinting at the possibility of these accessions originating from other geographical areas.

CONCLUSIONS

The examined common bean accessions displayed considerable genetic variability, which bolsters their use in crop enhancement programs. The SSR molecular markers implemented in this study demonstrated efficacy in quantifying genetic variability; hence, they can be deemed as supplemental instruments to foster a comprehensive comprehension of diversity among accessions from the BAG families and working collections.

AUTHORS CONTRIBUTION

Conceptual idea: Oliveira, T.C.; Barelli, M.A.A.; Sander, N.L.; Methodology design: Oliveira, T.C.; Barelli, M.A.A.; Sander, N.L.; Gilio, T.A.S.; Data collection: Oliveira, T.C.; Oliveira, A.J.; Sander, N.L.; Gilio, T.A.S.; Data analysis and interpretation: Oliveira, T.C.; Sander, N.L.; Gilio, T.A.S.; Felipin-Azevedo, R.; Writing and editing: Oliveira, T.C.; Barelli, M.A.A.; Oliveira, A.J.; Sander, N.L.; Gilio, T. A.S.; Felipin-Azevedo, R.

ACKNOWLEDGMENTS

REFERENCES

ALGHAMDI, S. S. et al. Biochemical and molecular characterization of cowpea landraces using seed storage proteins and SRAP marker patterns. Saudi Journal of Biological Sciences, 26(1):74-82, 2019.

ALMEIDA, C. P. D. et al. Genetic diversity, population structure, and Andean introgression in Brazilian common bean cultivars after half a century of genetic breeding. Genes, 11(11):1298, 2020.

BLAIR, M. W. et al. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics, 107(8):1362-1374, 2003.

BLAIR, M. W. et al. Diversity and population structure of common bean from Brazil. Crop Science, 53(5):1983-1993, 2013.

CATARCIONE, G. et al. Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces in the Lazio region of Italy. Plants, 12(4):744, 2023.

COELHO, C. M. M. et al. Características morfo-agronômicas de cultivares crioulas de feijão comum em dois anos de cultivo. Semina: Ciências Agrárias, 31(1):1177-1186, 2010.

CRUZ, C. D. Genes software-extended and integrated with the R, matlab and selegen. Acta Scientiarum. Agronomy, 38(4):547-552, 2016.

DELFINI, J. et al. Population structure, genetic diversity and genomic selection signatures among a Brazilian common bean germplasm. Scientific Reports, 11:2964, 2021.

EKBIÇ, E.; HASANCAOĞLU, E. M. Morphological and molecular characterization of local common bean (*Phaseolus vulgaris* L.) genotypes. Applied Ecology and Environmental Research, 17(1):841-853, 2019.

EL KADRI, N.; MIMOUN, M. B.; HORMANZA, J. I. Genetic diversity of Tunisian male date palm (*Phoenix dactylifera* L.) genotypes using morphological descriptors and molecular markers. Scientia Horticulturae, 253(6):24-34, 2019.

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular ecology, 14(8):2611-2620, 2005.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 131(2):479-491, 1992.

GAITÁN-SOLÍS, E. et al. Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, characterization, and ross-species amplification in *Phaseolus ssp.* Crop Science, 42(6):2128-2136, 2002.

GIOIA, T. et al. Evidence for introduction bottleneck and extensive intergene pool (Mesoamerican x Andean) hybridization in the European common bean (*Phaseolus vulgaris* L.) germplasm. PloS One, 8(10):e75974, 2013.

GIOIA, T. et al. Evolution of SSR diversity from wild types to US advanced cultivars in the Andean and Mesoamerican domestications of common bean (*Phaseolus vulgaris*). PLoS One, 14(1):e0211342, 2019.

GYANG, P. J.; MUGE, E. K.; NYABOGA, E. N. Genetic diversity and population structure of Kenyan common bean (*Phaseolus vulgaris* L.) germplasm using peroxidase gene markers. Proceedings of the National Academy of Sciences, 90(2):293-301, 2020.

LIU, K.; MUSE, S. V. PowerMarker: An integrated analysis environment for genetic marker analysis. Bioinformatics, 21(9):2128-2129, 2005.

MELO, W. H. M.; PAIXÃO, M. V. S. Disponibilidade de ferro em diferentes espécies de feijões cultivadas na região de Adustina-BA. Revista Ifes Ciência, 6(2):163-170, 2020.

MIR, R. R. et al. Allelic diversity, structural analysis, and Genome-Wide Association Study (GWAS) for yield and related traits using unexplored common bean (*Phaseolus vulgaris* L.) germplasm from Western Himalayas. Frontiers in Genetics, 11:609603, 2021.

NEI, M.; TAJIMA, F.; TATENO, Y. Accuracy of estimated phylogenetic trees from molecular data. Journal of Molecular Evolution, 19(2):153-170, 1983.

OLIVEIRA, L. S. G. et al. Genetic diversity of recombinant inbred lines of *Phaseolus vulgaris* L. Revista Ciência Agronômica, 52(3):e20207292, 2021.

ÖZKAN, G. et al. Determining genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from Türkiye using SSR markers. Genes, 13(8):1410, 2022.

PEAKALL, R.; SMOUSE, P. E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research: An update. Bioinformatics, 28(19):2537-2539, 2012.

PEREIRA, H. S. et al. Genetic diversity among common bean cultivars based on agronomic traits and molecular markers and application to recommendation of parent lines. Euphytica, 215(2):1-16, 2019.

PERINA, E. F. et al. Technological quality of common bean grains obtained in different growing seasons. Bragantia, 73(1):14-22, 2014.

PITURA, K.; ARNTFIELD, S. D. Characteristics of flavonol glycosides in bean (*Phaseolus vulgaris* L.) seed coats. Food Chemistry, 272:26-32, 2019.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. Genetics, 155(2):945-959, 2000.

ŠAJGALÍK, M. et al. Higher effectiveness of new common bean (*Phaseolus vulgaris* L.) germplasm acquisition by collecting expeditions associated with molecular analyses. Sustainability, 11(19):5270, 2019.

SANTOS, G. G. et al. Genetic diversity among bean landraces and cultivars for agronomy traits and selection of superior parents. Revista Ciência Agronômica, 53:e20217873, 2022.

SAVIĆ, A. et al. Genetic diversity of common bean (*Phaseolus vulgaris* L.) germplasm from Serbia, as revealed by single sequence repeats (SSR). Scientia Horticulturae, 288:110405, 2021.

SHAMSELDIN, A.; VELÁZQUEZ, E. The promiscuity of *Phaseolus vulgaris* L. (common bean) for nodulation with rhizobia: A review. World Journal of Microbiology and Biotechnology, 36:63, 2020.

VALENTINI, G. et al. Population structure and genetic diversity of common bean accessions from Brazil. Plant Molecular Biology Reporter, 36(5):897-906, 2018.

VAZ PATTO, M. C. et al. Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. Euphytica, 137:63-72, 2004.

VIDAK, M. et al. Assessment of the origin and diversity of Croatian common bean germplasm using phaseolin type, SSR and SNP markers and morphological traits. Plants, 10(4):665, 2021.

VOYSEST, O. V. Mejoramiento genetico del frijol (*Phaseolus vulgaris*) legado de variedades de América Latina (1930-1999). Cali, Colômbia: Centro Internacional de Agricultura Tropical, 2000. 195p.

WEIR, B. S.; COCKERHAM, C. Estimating F-statistics for the analysis of population structure. Evolution, 38(6):1358-1370, 1984.

WILLIAMS, J. G. K. et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research, 18(22):6531-6535, 1990.