Investigation of the genetic structure of some common bean 
(Phaseolus vulgaris L.) commercial varieties and genotypes 
used as a genitor with SSR and SNP markers

Omer Avican · Behiye Banu Bilgen

Received: 6 September 2021 / Accepted: 24 April 2022 / Published online: 20 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract Common bean is a species belonging to the Phaseolus genus of the Leguminosae family. It has economic importance due to being rich in protein, vitamin A and C, and minerals. Being one of the most cultivated species of legumes, the determination of genetic diversity in bean genotypes or populations has an important role in terms of our genetic resources. The objective of this study was to evaluate the genetic structure of 94 genotypes which were cultivated in different parts of the world and our country with SSR and SNP markers. 10 SSR loci and 73 SNP primers were used for the determination of genetic structure in commercial cultivars and breeding lines. All of the SSR and SNP loci used in the study were found to be polymorphic. A total of 89 alleles were identified for 10 SSR loci. Mean number of alleles per locus (Na = 8.9), effective allele number (Ne = 3.731), Shannon information index (I = 1.468), observed heterozygosity (Ho = 0.023), and expected heterozygosity (He = 0.654) were calculated based on SSR analysis. According to the results of Bayesian-based STRUCTURE analysis using SSR and SNP data, 94 bean genotypes were genetically divided into three main clusters. According to genetic distance based UPGMA dendrogram obtained from SNP analysis, 94 bean genotypes were divided into 2 main clusters corresponding Mesoamerican and Andean gene pools. The obtained results provide important information about the genetic structures of the studied bean cultivars and breeding lines. With the obtained results, it will be possible to develop breeding programs to develop new cultivars by using our gene resources.

Keywords Bean breeding · Genetic diversity · Phaseolus vulgaris · SSR · SNP

Introduction

Common bean (Phaseolus vulgaris L.) belongs to the Leguminosae family which consists of 727 genera and approximately 19,000 species. Some economically important species such as beans, chickpeas, lentils, soya, broad beans, and peas are members of this family. Five species belonging to the genus Phaseolus were cultivated for human nutrition in the world (P. vulgaris L., P. coccineus
L., *P. acutifolius* A. Gray, *P. lunatus* L., and *P. polyanthus* Greenman (Bitocchi et al. 2017; Nadeem et al. 2021). Leguminosae is the second largest flowering plant family with 1013 species belonging to 71 genera in Turkey. Around 400 of these species are endemic to Anatolia with the rate of 40% endemism (Toksoy et al. 2015; Aydin-Kandemir and Demir 2021).

*Phaseolus vulgaris* is the most preferred type of bean species for economic and scientific purposes. The *P. vulgaris* is native to America and it was believed that domestication occurs from northern Andean and from Mesoamerican populations as two gene pools (Bitocchi et al. 2017; Cortes and Blair 2017; Assefa et al. 2019). Common bean was brought to Europe at the beginning of the sixteenth century as an ornamental plant. After the introduction of common bean lines, their agriculture increased over time and started to be grown in almost every part of the world (Rodino and Drevon 2004; Bitocchi et al. 2017). Common bean cultivation in Turkey, especially for fresh pod and dry seed, dates back to the seventeenth century with the introduction of common bean into Turkey by Asian traders from Europe (Bozoglu and Sozen 2007; Nadeem et al. 2018).

Common bean is an important part of the human diet due to its higher protein content (>22% of their dry weight) compared to some cereals such as rice and wheat (Chandrakanth and Hall 2008; Castro-Guerrero et al. 2016). The fresh pods and seeds of common bean have approximately 90% water, and they are rich in A and C vitamins. Due to the high nutritional value, being suitable for consumption in different ways (fresh, dry, canned, pickled, etc.), being rich in minerals such as phosphorus and iron besides its protein source, common bean is one of the vegetables with the highest consumption in our country (Akcin 1973; Nadeem et al. 2018). The bean, which has an important position in agricultural production, stands out because it contains the protein, vitamins, complex carbohydrates, and minerals (Ca, Mg, K, Cu, Fe, Mg, and Zn) necessary for a healthy life (Marotti et al. 2007; Blair 2013; Nadeem et al. 2021). In addition to being consumed as a nutrient, beans are an important type of plant due to enriching the structure of the soil, increasing the amount of organic matter in the soil, accumulating nitrogen, and using plant residues as a component of commercial feed mixtures (Bitocchi et al. 2017; Yuvaraj et al. 2020).

Fresh common bean production in Turkey is 510,366 tons and dry bean production is 305,000 tons annually (TUIK 2021). Common bean (dry) global harvested area was 34,801,567 ha and production was 27, 545,942 tonnes in 2020 and common bean (green) global harvested area was 1,579,489 ha and production was 23,276,716 tonnes in 2020 (FAO 2021). The top three producers of fresh beans in the world are China, Indonesia, and Turkey according to average production from 1994 to 2020. In addition, India, Brazil, and Myanmar take the first three places in the average production of dry beans for the 1994–2020 range (FAO 2021). The large genetic diversity of beans is one of the reasons for such wide cultivation in the world and Turkey and also the increased usage of common bean in breeding studies over the years.

Turkey is significantly rich in plant genetic resources due to its location at the crossroads of the Mediterranean and the Near East gene centers. Moreover, Turkey has extensive biodiversity in terms of habitat types, geomorphological structure, climate, and topographic features (Özhatay et al. 2011; Noroozi et al. 2019). To protect the plant gene resources of our country, it is necessary to determine the genetic diversity, genetic and morphological characterizations of our plant resources and also evaluate the potentials of the species for various studies such as breeding. Especially with the advances in molecular biology and genetics in the last 50 years, the emergence and development of modern biotechnology have gained importance. The recent developments in DNA marker technology has reached high levels and has provided valuable tools in various genetic analyses, from phylogenetic analysis to the cloning of genes. It is possible to easily determine the genetic structure, create molecular maps, and label the characters of interest by PCR-based markers. Researchers have utilized various molecular markers to conduct molecular genetic studies in common bean populations/genotypes (Metais et al. 2002; Sicard et al. 2005; Galvan et al. 2006; Benchimol et al. 2007; Angioi et al. 2010; Buah et al. 2017; Carucci et al. 2017; Nadeem et al. 2018; Aydin and Baloch 2019; Nogueira et al. 2021; Savic et al. 2021; Sadohara et al. 2022).

In Turkey, the number of studies in which bean gene resources are defined by morphological or molecular methods has started to increase in recent years. In this study, the aims were: (1) to identify the
genetic structure of studied common bean cultivars and breeding lines by SSR and SNP markers, (2) to determine lines that can be used for the establishment of breeding programs by revealing genetic relationships between studied common bean cultivars and breeding lines.

Materials and methods

Plant materials

In this study, a total of 94 bean genotypes (P. vulgaris, P. acutifolius, and P. coccineus) were used. These included 79 commercial bean varieties that were cultivated in different regions of the world and Turkey and 15 breeding lines used in our breeding programs. The common bean genotypes included in the study were P. vulgaris nanus: Determinate-Bush and P. vulgaris comminus: Indeterminate-Climber bean forms (Table 1). Each bean genotype was grown in the greenhouse under controlled conditions. Fresh leaf samples (300 mg) from 94 genotypes were collected in a 96-well Qiagen tissue collection plate and they were stored at −80 °C until DNA extraction.

DNA extraction

DNA extraction was performed with Qiagen DNA Extraction Instrument (QIAcube-HT). Each tissue sample was ground using Tissue Lyser II for 2 min. The quantification and qualification of isolated DNAs were performed with Thermo Scientific™ NanoDrop™ One Microvolume UV–Vis Spectrophotometer. The DNA samples were preserved at −20 °C till PCR analysis.

SSR analysis

Ten SSRs (BM141, BM143, BM152, BM160, BM172, GATS91, PV-at002, PV-ctt001, PV-ag001, and PV-at007) were used for the genetic characterization of P. vulgaris genotypes (Yu et al. 2000; Gaitan-Solis et al. 2002). The fluorescently-labeled M13-tailed primer method was used for PCR amplification (Schuelke 2000). The characteristics of the SSR primers were indicated in Table 2.

The PCR amplifications were performed as described in Yu et al. (2000) and Gaitan-Solis et al. (2002) with the Applied Biosystems® Veriti™ Thermal Cycler. PCR products were controlled by 2% agarose gel electrophoresis (1X TBE buffer, 110 V, 120 min). Gel Imaging System Vilber Lourmat Quantum ST5 was used to visualize the agarose gels. The size of SSR fragments was determined by 3500 Genetic Analyzer (Applied Biosystems, Life Technologies, UK) capillary electrophoresis and GeneMapper Software 5.0 (Applied Biosystems).

SNP analysis

128 SNP loci represented 11 P. vulgaris chromosomes were selected from Blair et al. (2013) for genotyping of 94 bean samples. For the SNP analysis Roche LightCycler® DNA Master HybProbe Master Mix and ROCHE –LightCycler® 480 Instrument II Real Time PCR were used.

Data analysis

For each SSR locus, observed allele size range (bp) and observed allele number were determined. In statistical analysis of SSR data, allele frequencies, allele numbers (Na), effective allele numbers (Ne), Shannon’s information index (I), heterozygosity levels (Ho and He), and polymorphic information contents (PIC) were estimated by the software GenAlEx Version 6.3 (Peakall and Smouse 2006).

Population structure based on SSR and SNP data was evaluated using STRUCTURE 2.3.4 (Pritchard et al. 2000) as described in Blair et al. (2009). Analyses had a burn-in length of 50,000 iterations and a
run length of 100,000 iterations after burning. Ten replicates were carried out for each K value (K = 1 to K = 10) (Evanno et al. 2005). STRUCTURE HARVESTER was used in order to determine the best K value (Earl and vonHoldt 2012). The genetic distances and similarity matrices generated using the Dice coefficient was used to construct a dendrogram with the UPGMA method using the DARwin 6.0 software (Perrier and Jacquemoud-Collet 2006).

In the SNP analysis, after the Real Time PCR process was completed, the genotypes in the SNP region were determined by Melting Point analysis. The binomial data matrix was created by scoring the raw data obtained in the SNP analysis according to present (1) or absent (0). The genetic distance matrix was used to create a dendrogram based on the UPGMA method using the software DARwin 6.0 software (Perrier and Jacquemoud-Collet 2006).

### Results

The polymorphism level of studied 10 SSR loci was estimated as 100%. Totally 89 alleles (mean value = 8.9 alleles/locus) were determined. Evaluating all studied commercial varieties and breeding lines, BM141, GATS91, and PV-at007 loci have the highest number of alleles (13 alleles), and BM160 has the lowest number of alleles (3 alleles). 11 alleles were observed in BM143 and BM152 loci. The remaining SSR loci have 7 or 6 alleles (Table 3). Table 3 indicates genetic diversity parameters estimated in the studied *P. vulgaris* genotypes with 10 SSR loci.

PIC values were calculated for SSR loci ranged from 0.854 to 0.289. The mean PIC value was estimated relatively high (0.621). Based on SSR analysis, the overall mean number of effective alleles per locus (Ne) was 3.731 ± 0.628 (varied from 1.440 to 7.539).

| Primer | Sequence 5′ 3′ |
|--------|----------------|
| M13-FAM | 5′-FAM-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM141-F | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM141-R | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-PET | 5′-PET-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM143-F | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM143-R | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-NED | 5′-NED-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM152-F | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM152-R | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-VIC | 5′-VIC-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM160-F | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM160-R | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-FAM | 5′-FAM-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM172-F | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| BM172-R | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-FAM | 5′-FAM-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| GATS91-F | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| GATS91-R | 5′-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-VIC | 5′-VIC-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| PV-at002-F | 5′-VIC-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| PV-at002-R | 5′-VIC-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-PET | 5′-PET-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| PV-ct001-F | 5′-PET-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| PV-ct001-R | 5′-PET-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-NED | 5′-NED-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| PV-ag001-F | 5′-NED-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| PV-ag001-R | 5′-NED-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |
| M13-NED | 5′-NED-TGTAAGCAGGGCCGAGATCTCGGTAATGAGGCTGACGACGAGGAAA 3′ |

### Table 2

SSR and M13 primers used in the study.
The overall average value of Shannon’s information index (I) was calculated as 1.468. The highest value of I was observed in the GATS91 locus (2.211) and the lowest in the BM160 locus (0.579). Estimated values of mean expected heterozygosity (He) and observed heterozygosity (Ho) were 0.654 and 0.023, respectively (Table 3).

The ideal K value was calculated by STRUCTURE HARVESTER (Earl and vonHoldt 2012) program in the SSR based STRUCTURE analysis for 94 different genotypes (Pritchard et al. 2000). The optimal number of subpopulations was found for K=3 (Fig. S1). Bayesian-based STRUCTURE analysis showed that 94 bean genotypes were distributed between 3 main groups (Fig. 1). DARwin 6.0 software was used to evaluate the distance matrix and dendrogram (Fig. 2). Genetic diversity was found in the range of 0.17–1.0 for all studied bean genotypes. Some of the genotypes such as OT1/SK15, SK11/SB2, OT28/SK7/SB7, OB3A/ST3 had genetic diversity value of 1.0. Three

| Locus   | N   | Observed allele size range (bp) | Most frequent allele frequency (allele size, bp) | Na  | Ne | I    | Ho   | He   | PIC  |
|---------|-----|--------------------------------|-----------------------------------------------|-----|----|------|------|------|------|
| BM141   | 94  | 196–252                         | 0.510 (243)                                    | 13  | 3.096 | 1.560 | 0.021 | 0.677 | 0.644 |
| BM143   | 94  | 133–187                         | 0.432 (149)                                    | 11  | 4.001 | 1.756 | 0.011 | 0.750 | 0.723 |
| BM152   | 94  | 104–150                         | 0.489 (108)                                    | 11  | 3.458 | 1.689 | 0.021 | 0.711 | 0.686 |
| BM160   | 94  | 198–203                         | 0.805 (198)                                    | 3   | 1.479 | 0.579 | 0.042 | 0.324 | 0.289 |
| BM172   | 94  | 94–128                          | 0.474 (112)                                    | 6   | 2.967 | 1.292 | 0.021 | 0.663 | 0.608 |
| GATS91  | 94  | 232–276                         | 0.208 (248)                                    | 13  | 7.539 | 2.211 | 0.031 | 0.867 | 0.854 |
| PV-at002| 94  | 258–268                         | 0.828 (262)                                    | 6   | 1.440 | 0.696 | 0.010 | 0.306 | 0.294 |
| PV-ctt001| 94 | 160–193                          | 0.354 (193)                                    | 7   | 3.672 | 1.469 | 0.042 | 0.728 | 0.683 |
| PV-ag001| 94  | 157–175                         | 0.422 (175)                                    | 6   | 2.934 | 1.264 | 0.021 | 0.659 | 0.597 |
| PV-at007| 94  | 210–234                         | 0.274 (212)                                    | 13  | 6.723 | 2.166 | 0.011 | 0.851 | 0.836 |
| Mean    | 94  | –                               | –                                             | 8.9 | 3.731 | 1.468 | 0.023 | 0.654 | 0.621 |
| Standard Error | – | –                              | –                                             | 1.169 | 0.628 | 0.171 | 0.004 | 0.061 | 0.058 |

Table 3 Genetic diversity parameters of studied 10 SSR loci (N=Sample size, Na=mean number of alleles per locus, Ne=effective number of alleles, I=Shannon’s information index, Ho=observed Heterozygosity, He=expected heterozygosity (Nei 1987), PIC=Polymorphic information contents)
main clusters were observed for 94 bean genotypes in the diversity matrix based dendrogram. As a result of the analysis using 10 SSR markers, it was seen that this marker system did not clearly distinguish genotypes of Andean and Mesoamerican origin.

According to SNP analysis, 73 out of 128 studied SNP primers were polymorphic for the studied *P. vulgaris* genotypes. Evaluating polymorphic SNP primers’ melting peak profiles, with each peak as an allele, revealed highest allele group for SNP51 and SNP65 (5 alleles) primers. SNP13, SNP22, SNP28, SNP61, SNP62, SNP69 and SNP72 primers had the lowest allele groups (2 alleles). The number of allele groups and PIC values for SNP primers were given in Table 4. Calculated PIC values were varied between 0.042 and 0.523. The mean PIC value was 0.337. SNP59 has the highest PIC value (0.523), whereas SNP61 and SNP69 had the lowest PIC value (0.042).

The binomial data matrix was created by scoring the raw data obtained in the SNP analysis according

---

Fig. 2 Distance matrix based dendrogram constructed by using 10 SSRs
to the presence (1) or absence (0). The ideal K value was calculated by STRUCTURE HARVESTER (Earl and vonHoldt 2012) program in the SSR based STRUCTURE analysis for 94 different genotypes (Pritchard et al. 2000). The optimal number of subpopulations was found for K = 3 (Fig. S2). Bayesian-based STRUCTURE analysis showed that 94 bean genotypes were distributed between 3 main groups (Fig. 3). DARwin 6.0 software was used to evaluate the distance matrix and dendrograms. The dendrogram created as a result of the SNP analysis is given in Fig. 4. When the dendrogram was examined, it was observed that the genetic diversity varied between 0.46 and 1.00. Some of the genotypes such as OT3/OT30, OB3A/OB3B had genetic diversity value of 1.0. 94 bean genotypes used in the study are divided into 2 main clusters in the dendrogram. At the same time, the two main clusters were divided into subgroups, differentiating bean varieties to a large extent and gave successful results in obtaining the expected subgroups. The degree of kinship of bean genotypes was determined with the help of Euclidean similarity index coefficients, and 3D graphics were created for SNP (Fig. 5). Clusters formed by the studied bean genotypes were observed in accordance with the results in UPGMA dendrograms. As a result of the SNP analysis, it was seen that the genotypes of Andean (63 genotypes) and Mesoamerican (31 genotypes) origin were clearly separated.

### Discussion

The molecular marker studies provide plant breeders with information about the available genetic diversity and the genetic control of significant agronomic traits related to domestication. DNA fingerprinting studies using molecular markers (especially SSR and SNP markers) have a high impact on revealing the differences between genotypes (Assefa et al. 2019). Various DNA marker systems are used in genetic diversity studies, and the comparison of the use of these systems is extremely important for molecular plant breeding studies and analysis. Interlaboratory transfer of the DNA marker systems is necessary for standardization and comparison of the data obtained in order to obtain reproducible results. Thus, the financial costs of the work done are reduced and time can be saved. Cortes et al. (2011) reported that SSR and SNP markers are ideal markers when used together in diversity studies in beans. In the study conducted by Ulukapi and Onus (2012), genetic analyses were made using SCAR and SSR markers in beans, and a UPGMA dendrogram was created based on SCAR and SSR data of 39 genotypes. In this study,

| SNP Code | Allele Groups | PIC | SNP Code | Allele Groups | PIC |
|----------|---------------|-----|----------|---------------|-----|
| SNP1     | 3             | 0.187 | SNP38    | 4             | 0.450 |
| SNP2     | 3             | 0.376 | SNP39    | 3             | 0.412 |
| SNP3     | 3             | 0.371 | SNP40    | 4             | 0.462 |
| SNP4     | 3             | 0.403 | SNP41    | 3             | 0.399 |
| SNP5     | 2             | 0.331 | SNP42    | 3             | 0.268 |
| SNP6     | 3             | 0.394 | SNP43    | 3             | 0.388 |
| SNP7     | 3             | 0.393 | SNP44    | 3             | 0.411 |
| SNP8     | 3             | 0.377 | SNP45    | 4             | 0.375 |
| SNP9     | 3             | 0.405 | SNP46    | 3             | 0.332 |
| SNP10    | 3             | 0.375 | SNP47    | 3             | 0.386 |
| SNP11    | 3             | 0.391 | SNP48    | 3             | 0.291 |
| SNP12    | 3             | 0.349 | SNP49    | 3             | 0.295 |
| SNP13    | 2             | 0.362 | SNP50    | 3             | 0.149 |
| SNP14    | 3             | 0.375 | SNP51    | 5             | 0.436 |
| SNP15    | 3             | 0.314 | SNP52    | 3             | 0.118 |
| SNP16    | 4             | 0.319 | SNP53    | 3             | 0.152 |
| SNP17    | 3             | 0.352 | SNP54    | 3             | 0.318 |
| SNP18    | 3             | 0.329 | SNP55    | 3             | 0.261 |
| SNP19    | 3             | 0.384 | SNP56    | 3             | 0.385 |
| SNP20    | 3             | 0.291 | SNP57    | 2             | 0.270 |
| SNP21    | 3             | 0.386 | SNP58    | 3             | 0.390 |
| SNP22    | 2             | 0.326 | SNP59    | 4             | 0.523 |
| SNP23    | 3             | 0.387 | SNP60    | 3             | 0.388 |
| SNP24    | 3             | 0.378 | SNP61    | 2             | 0.042 |
| SNP25    | 3             | 0.393 | SNP62    | 2             | 0.116 |
| SNP26    | 3             | 0.425 | SNP63    | 3             | 0.119 |
| SNP27    | 3             | 0.388 | SNP64    | 3             | 0.288 |
| SNP28    | 2             | 0.358 | SNP65    | 5             | 0.458 |
| SNP29    | 3             | 0.390 | SNP66    | 3             | 0.298 |
| SNP30    | 3             | 0.349 | SNP67    | 3             | 0.363 |
| SNP31    | 3             | 0.100 | SNP68    | 3             | 0.391 |
| SNP32    | 3             | 0.414 | SNP69    | 2             | 0.042 |
| SNP33    | 3             | 0.163 | SNP70    | 3             | 0.389 |
| SNP34    | 3             | 0.384 | SNP71    | 3             | 0.466 |
| SNP35    | 3             | 0.415 | SNP72    | 2             | 0.367 |
| SNP36    | 4             | 0.407 | SNP73    | 3             | 0.377 |
| SNP37    | 3             | 0.288 |
significant information about the genetic structure of 94 genotypes from commercial bean varieties and breeding genotypes cultivated in the world and different regions of Turkey has been obtained via SSR and SNP markers. 89 polymorphic bands were obtained by using 10 SSR primers. The mean number of polymorphic bands per primer is 8.9. Blair et al. (2006) performed SSR analysis (129 SSRs) in order to determine the genetic structure of 43 *P. vulgaris* genotypes and 1 *P. acutifolius* obtained from different regions of America, and the polymorphism rate obtained with genomic microsatellites was determined as (0.446). In the study of Kwak and Gepts (2009), the average number of alleles was determined as 16 with 26 SSRs in bean genotypes collected from different geographical regions. Burle et al. (2010) were analysed 67 SSRs in 279 bean genotypes collected from Brazil and the average number of alleles was calculated as 6. Since Brazil is one of the gene centers of the bean, high genetic diversity has been reported among the genotypes there. In Cabral et al. (2011), 16 SSR markers were used to determine genetic diversity in 57 bean genotypes collected from the Brazilian region, 13 SSRs were found to be polymorphic, and the number of alleles obtained from these markers was calculated as 29 and the average number of alleles per locus was calculated as 2.2. Khaidizar et al. (2012) reported 72 alleles at 30 SSR loci in bean genotypes sampled from North Anatolia. Bilir et al. (2019), a total of 192 alleles were identified in 13 SSR markers, and the average number of alleles per locus was reported as 14.8. In the study of Ekbic and Hasancaoglu (2019), it was stated that 63 alleles (polymorphism rate 73%) belonging to 18 SSR loci in bean genotypes and the average number of alleles per locus was 2.55. The average number of polymorphic bands obtained by the researchers is close to the values we obtained from this study.

Shannon information index (I), which is one of the genetic diversity parameters, was calculated from 0.579 to 2.211. Hence, this high I value (mean 1.468) indicates high variation within genotypes. The mean observed heterozygosity (Ho) was calculated as 0.023 and the mean expected heterozygosity (He) was calculated as 0.654. Obtained gene diversity level was notably lower when compared to similar studies. Since the genotypes used in this study belongs to commercial varieties and breeding materials, most of the samples have homozygous genotypes and therefore Ho was low. In Bilir et al. (2019), the observed heterozygosity was calculated as 0.452 and the expected heterozygosity was 0.724. Valentini et al. (2018) reported that genetic diversity (h) on 18 SSR loci in 109 bean genotypes sampled from Brazil was 0.44. In Valentini et al. (2018), 4 groups were observed in which two Andean and two Mesopotamian genotypes clustered within themselves at K=4. Pereira et al. (2019) reported He as 0.55 and
Ho as 0.05 in 17 bean varieties. Pereira et al. (2019) using the toucher method and evaluating 27 SSR loci, 17 bean genotypes formed 4 different groups in their clustering analysis. In the study of Carucci et al. (2017), the genetic structure of Italian local bean cultivars was analysed with 12 SSR loci, the Ho value was calculated as 0.24. The genotypes examined in the different studies show variation such as local genotypes, breeding material, commercial varieties etc., and it can be argued that interventions with breeding narrow the genetic basis of breeding material compared to local genotypes. Therefore, differences in genetic diversity level of studied bean genotypes were expected.

According to the results of the Bayesian-based STRUCTURE analysis based on the SSR data conducted within the scope of the study, 94 bean genotypes were genetically divided into 3 main groups. Phylogenetic tree was created using UPGMA grouping also gave 3 main groups. When the results of STRUCTURE analysis and UPGMA dendrogram were compared, it has been observed that obtained SSR data could not able to differentiate bean genotypes especially based on Mesoamerican or andean origin exactly. Such a result may be due to the type and number of markers selected and used in this study. Therefore, it was planned to select and use a different marker type and SNP analysis was performed. Burle et al. (2010) reported that 279 bean genotypes found in Brazil were divided into two groups (K=2), andean and Mesoamerican, as a result of STRUCTURE analysis by using microsatellite diversity. Blair et al. (2012) reported that 108 bean genotypes analysed with with 36 microsatellites...
were divided into 5 groups as Andean, Colombian, Ecuadorian, Northern Peruvian, Guatemalan, and Mesoamerican. Savic et al. (2021) studied Serbia germplasm with 27 SSRs, and clear separation of studied landraces and cultivars as Mesoamerican and Andean was reported.

For SNP analysis, 73 of 128 SNP primers were selected from different regions of 11 chromosomes of *P. vulgaris*, and they were determined as polymorphic. The mean PIC value was calculated as 0.337. In Cortes et al. (2011), the SNP diversity was performed in beans, the mean PIC value of 94 SNP primers in 70 bean genotypes (28 Andean and 42 Mesoamerican) was reported as 0.437. In the study, 2 main clusters with Andean and Mesoamerican gene pools were observed in 70 bean genotypes cultured and it was reported that SNP analysis differentiated these groups as expected. Blair et al. (2013), in a study conducted on *P. vulgaris* to screen for parental polymorphism and to determine genetic diversity, 736 SNPs were primarily scored in 236 different bean genotypes and the mean PIC value was calculated as 0.328. The mean PIC value of the SNP primers we used in our study was determined between the average PIC values obtained by Cortes et al. (2011) and Blair et al. (2013). SNP markers are highly sensitive markers based on single nucleotide polymorphism, and also the SNP primers have a very high power to discriminate between genotypes/populations.

Fig. 5 3D graphic obtained using the Euclidean similarity index as a result of SNP analyses
According to the results of the Bayesian-based STRUCTURE analysis based on the SNP data, 94 bean genotypes were genetically divided into 3 main groups. UPGMA dendrogram gave 2 main groups. Based on the dendrogram created as a result of SNP analyses, it was observed that the genetic diversity ranged between 0.46 and 1.00. Distance matrix based dendrogram constructed by using SNPs created Cluster I and Cluster II (Fig. 4). Cluster I included 31 bean genotypes (%33 of studied genotypes) with Mesoamerican origin. The genotypes of *P. acutifolius*-1, *P. acutifolius*-2 and *P. coccineus* species used as standard (control) varieties appeared in the Cluster I. Cluster II included 63 common bean genotypes (%67 of studied genotypes) with Andean origin. The genetic diversity was higher in the Andean gene pool than within the Mesoamerican gene pool. When the dendrogram and STRUCTURE results were examined in detail, it has been observed that obtained SNP data could differentiate bean genotypes based on Mesoamerican or Andean origin successfully as expected. Thus, SNP is an efficient, more accessible and effective approach for genotyping the bean genotypes and analysing genetic relatedness for large-scale screening. Therefore, bean breeding programs should be able to utilize marker-assisted selection (MAS) to introgressive genes of economic importance into own bean breeding lines. This information allows the breeders to make a good decision in the creation of new cross combinations and estimate how diversity will be created in the advanced breeding stages. In the following breeding planning, it may be the subject of study to reveal the relationship between the obtained SNP data and some desired traits or linked candidate genes, and the selection of genotypes with the desired trait can be done with MAS. In Cortes et al. (2011), KASPar technology was used to develop SNP markers in 70 bean genotypes belonging to the Andean and Mesoamerican gene pools. In that study, 84 genomic and 10 EST-SNP markers were developed and it was reported that the Mesoamerican and Andean gene pools were successfully separated using these primers. In Cortes et al. (2011), compared to the Mesoamerican gene pool, more diversity was observed in individuals belonging to the Andean gene pool. In addition, it has been reported that SSR and SNP markers are ideal markers when used together in bean diversity studies. Genotypes that are in the same cluster and have the desired genetic distance between groups and sub-group can help the breeders in order to create the new population in terms of breeding targets and in providing expected diversity.

There are lots of tools to obtain desired genetic diversity in plant breeding. However, it is very important to know the genetic distance of the unknown genotypes within the germplasm in the plant breeding programs. Although the classical breeding studies in many agricultural plant species have reached the desired rate, the use of molecular markers in the development of new genotypes and varieties has made significant contributions to breeding programs. Classical and molecular breeding programs are created by adding DNA markers that provide valuable data to existing breeding programs that are developing very rapidly. It is important to reveal the genetic relationship between bean species, breeding lines under development in bean breeding programs, and existing commercial bean varieties in detail, as well as in family selections, genetic analyses for various purposes, and in the planning of breeding programs. Genetic diversity analyses of owned gene resources also allow the use of data at molecular, geographic, functional, and morphological levels (Lu et al. 2009). Genetic distance and proximity studies provide the emergence of differences between the studied genotypes and contribute to increasing the genetic diversity in the gene pool in breeding programs. The more genetic distance the genotypes have from each other, the greater the variation seen. These openings seen in breeding genotypes shape the selection and the more variation is obtained, higher the chance of success of the breeding program, which makes it easier for the breeder to reach the goal. Performing genetic analyses in genetic diversity studies, determining distance and proximity conditions, contributes to the creation of new populations and to obtain high yielding combinations with heterosis. Plant breeders use the evaluation of genetic diversity using various methods as an alternative selection method, the genetic diversity data obtained helps to organize the studied genotypes into groups. Thus, the creation of the most promising hybrid combinations among genotypes with known morphological, agronomic, and genetic features allows the creation of combinations that can be cost and time effective (Souza et al. 2008). In this study, it was aimed to determine
the genetic closeness-distances by obtaining the bean species, breeding lines and commercial varieties, and phylogenetic tree from the data obtained by using molecular methods and to use these data in the breeding program. The importance of the study is clearly revealed, as it reveals the possibility of increasing the chance of success both in the selections to create new strong populations and in the selections in the formation of productive hybrids, by revealing the kinship relations between the lines via molecular markers. The genetic structure of the studied genotypes corresponded to the expected differentiation of common bean genotypes according to gene pool origin, Mesoamerican or Andean, by SNP analysis. This information allows the breeders to make a good decision in the creation of new cross combinations and estimate how diversity will be created in the advanced breeding stages. Some genotypes from the dwarf slicing bean (OT group) can be divided into the groups within itself and genotypes that are in the desired genetic distance can be used to create a new initial population in order to gather targeted traits in one linevariety. Some candidate genotypes (such as; OT22 and OT29) can be selected as putative parents for bean breeding. The same is true for the other bean groups too. In this study, we had important knowledge with molecular characterization of studied bean genotypes to our practices for mating design, and the selection of the genotypes that will be used to construct the examination population in-field phenotyping trials. Our study was also shown that chosen SNP markers, proved their usefulness and efficiency in discriminating bean genotypes according to the gene pool origin.

Acknowledgements We gratefully acknowledge Dr. Hasan O. SIGVA for scientific and technical support for SNP analysis. We thank to Assist. Prof. Dr. Ibrahim Celik (Pamukkale University, Çal Vocational High School) for the technical support in using the Darwin program.

Author contributions OA & BBB contributed to the design of the study, laboratory analysis, data processing, writing manuscript.

Funding This study was funded by Tekirdağ Namık Kemal University, Scientific Research Projects Unit (Project No: NKUBAP.03.YL.18.171).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Human participants or animals This article does not contain any studies with human participants or animals performed by any of the authors.

References

Akcin A (1973) Erzurum şartlarında yetiştirilen kuru fasulye çeşitlerinde gübreleme. ekim zamanı ve sıra aralığının tane verimine etkisi ile bu çeşitlerin bazı fenolojik, morfolojik ve teknolojik karakterleri üzerine bir araştırma. Atatürk Üniversitesi Ziraat Fakültesi Dergisi 4(2):65–76 (in Turkish)

Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G, Negri V, Spagnoletti Zeuli PL, Papa R (2010) Beans in Europe: origin and structure of the European landraces of Phaseolus vulgaris L. Theor Appl Genet 121:829–843

Assefa T, Mahama AA, Brown AV, Cannon EKS, Rubyogo JC, Rao IM, Blair MW, Cannon SB (2019) A review of breeding objectives, genomic resources and marker-assisted methods in common bean (Phaseolus vulgaris L.). Mol Breed 39:20

Aydin MF, Baloch FS (2019) Exploring the genetic diversity and population structure of Turkish common bean germplasm by the iPBS-retrotransposons markers. Legum Res 42(1):18–24

Aydin-Kandemir F, Demir A (2021) Endangered species in Turkey: a specific review of endangered Fabaceae species with IUCN Red List data (Türkiye’de tehlike altındaki türler: IUCN Kırmızı Liste verileri ile tehlike altındaki Fabaceae türlerine yönelik özel bir inceleme). Turk J Biod 4:53–65 (in Turkish)

Benchimol LL, Campos T, Carbonell SAM, Colombo CA, Chioratto AF, Formighieri EF, Souza AP (2007) Structure of genetic diversity among common bean (Phaseolus vulgaris L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. Genet Resour Crop Evol 54:1747–1762

Bilir O, Yuksel Ozmen C, Ozcan S, Kibar U (2019) Genetic analysis of Turkey common bean (Phaseolus vulgaris L.) genotypes by simple sequence repeats markers. Russ J Genet 55:61–70

Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Gioia T, Santo D, Nanni L, Attene G, Papa R (2017) Beans (Phaseolus spp.) as a model for understanding crop evolution. Front Plant Sci 8:722

Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, Beebe SE (2006) Microsatellite marker diversity in common bean (Phaseolus vulgaris L.). Theor Appl Genet 113:100–109

Blair MW, Diaz LM, Buendia HF, Duque MC (2009) Genetic diversity, seed size associations and population structure of a core collection of common beans (Phaseolus vulgaris L.). Theor Appl Genet 119:955–972
Blair MW, Soler A, Cortes AJ (2012) Diversification and population STRUCTURE in common beans (Phaseolus vulgaris L.). PLoS ONE 7(11):e49488

Blair MW (2013) Mineral biofortification strategies for food staples: the example of common bean. J Agric Food Chem 61:8287–8294

Blair MW, Cortes AJ, Pennetisa RV, Farmer A, Carraquilla-Garcia N, Cook DR (2013) A high-throughput SNP marker system for parental polymorphism screening and diversity analysis in common bean (Phaseolus vulgaris L.). Theor Appl Genet 126(2):535–548

Bozoglu H, Sozen O (2007) Some agronomic properties of the local population of common bean (Phaseolus vulgaris L.) of Artvin province. Turk J Agric For 31:327–334

Buah S, Buruchara R, Okori P (2017) Molecular characterisation of common bean (Phaseolus vulgaris L.) accessions from southwestern Uganda reveal high levels of genetic diversity. Genet Resour Crop Evol 64:1985–1998

Burle ML, Fonseca JR, Kami JA, Gepts P (2010) Microsatellite diversity and genetic STRUCTURE among common bean (Phaseolus vulgaris L.) landraces in Brazil, a secondary center of diversity. Theor Appl Genet 121:801–813

Cabra PDS, Soares TCB, Lima ABP, De Miranda FD, Souza FB, Gonçalves LSA (2011) Genetic diversity in local and commercial dry bean (Phaseolus vulgaris) accessions based on microsatellite markers. Genet Mol Res 10(1):140–149

Carucci F, Garramone R, Aversano R, Carputo D (2017) SSR markers distinguish traditional Italian bean (Phaseolus vulgaris L.) landraces from Lamon. Czech J Genet Plant Breed 53(4):168–171

Castro-Guerrero NA, Isidra-Arellano MC, Mendoza-Cozatl DG, Valdés-López O (2016) Common bean: a legume model on the rise for unraveling responses and adaptations to iron, zinc, and phosphate deficiencies. Front Plant Sci 7:600

Chandrakanth E, Hall TC (2008) Phaseolin: structure and evolution. The Open Evol J 2:66–74

Cortes AJ, Blair MW (2017) Lessons from common bean on how wild relatives and landraces can make tropical crops more resistant to climate change. In: Grillo O (ed.) Rediscovery of landraces as a resource for the future. InTech ISBN 978–953–51–5806–6

Cortes AJ, Chavarro MC, Blair MW (2011) SNP marker diversity in common bean (Phaseolus vulgaris L.). Theor Appl Genet 123:827–845

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Ekbc E, Hasancaoglu EM (2019) Morphological and molecular characterization of local common bean (Phaseolus vulgaris L.) genotypes. Appl Ecol Environ Res 17(1):841–853

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

FAO (2021) FAOSTAT Statistical Database. Food and Agriculture Organization of the United Nations, Rome, Italy. [Access date: 07 February 2021]

Gaitan-Solis E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (Phaseolus vulgaris): isolation, characterization, and cross-species amplification in Phaseolus spp. Crop Sci 42:2128–2136

Galvan MZ, Menendez-Sevillano MC, De Ron AM, Santalla M, Balatti PA (2006) Genetic diversity among wild common beans from northwestern Argentina based on morpho-agronomic and RAPD data. Genet Resour Crop Evol 53:891–900

Khaidizir MA, Haliloglu K, Elkoça E, Aydin M, Kantar F (2012) Genetic diversity of common bean (Phaseolus vulgaris L.) landraces grown in Northeast Anatolia of Turkey assessed with simple sequence repeat markers. Turk J Field Crops 17(2):145–150

Kwak M, Gepts P (2009) Structure of genetic diversity in the two major gene pools of common bean (Phaseolus vulgaris L., Fabaceae). Theor Appl Genet 118(5):979–992

Lu Y, Yan J, Guimaraes CT, Tabu S, Hao Z, Gao S, Chen S, Li J, Zhang S, Vivek BS, Magorokosho C, Mugo S, Makumbi D, Parenton SN, Shah T, Rong T, Crouch JH, Xu Y (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theor Appl Genet 120:93–115

Marotti I, Bonetti A, Minelli M, Catizzone P, Dinelli G (2007) Characterization of some Italian common bean (Phaseolus vulgaris L.) landraces by RAPD, Semi-Random and ISSR molecular markers. Genet Resour Crop Evol 54:175–188

Métais I, Hamon B, Jalouzot R, Peltier D (2002) Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. Theor Appl Genet 104(8):1346–1352

Nadeem MA, Habyarimana E, Ciftci V, Nawaz MA, Karakoy T, Comertpay G, Shahid MQ, Shahid MQ, Yeken MZ, Fawad A, Ercisli S, Chung G (2018) Characterization of genetic diversity in Turkish common bean gene pool using phenotypic and whole-genome DArTseq-generated silicoDArT marker information. PLoS ONE 13(10):e0205363

Nadeem MA, Yeken MZ, Shahid MQ, Habyarimana E, Yilmaz H, Alsaleh A, Hatipoglu R, Cilesiz Y, Khawar KM, Luddi N, Ercisli S, Aasim M, Karakoy T, Baloch FS (2021) Common bean as a potential crop for future food security: an overview of past, current and future contributions in genomics, transcriptomics, transgenics and proteomics. Biotechnol Biotecnol Equip 35:758–786

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, NY, p 512

Nogueira AF, Moda-Cirino V, Delfini J, Brandao LA, Mian H, Alsaleh A, Hatipoglu R, Cilesiz Y, Khawar KM, Ludidi N, Ercisli S, Aasim M, Karakoy T, Baloch FS (2021) Common bean as a potential crop for future food security: an overview of past, current and future contributions in genomics, transcriptomics, transgenics and proteomics. Biotechnol Biotecnol Equip 35:758–786

Özhatay FN, Kültür Ş, Gürdal MB (2011) Check-list of additional taxa to the supplement flora of Turkey V. Turk J Bot 35:589–624
Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295
Pereira HS, Mota APS, Rodrigues LA, de Souza TLPO, Melo LC (2019) Genetic diversity among common bean cultivars based on agronomic traits and molecular markers and application to recommendation of parent lines. Euphytica 215:38
Perrier X, Jacquemoud-Collet JP (2006) DARwin software v.06. http://darwin.cirad.fr/darwin
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
Rodino AP, Drevon JJ (2004) Migration of a grain legume Phaseolus vulgaris in Europe. In: Werner D. (eds.) Biological Resources and Migration. Springer, Berlin, Heidelberg
Sadohara R, Izquierdo P, Couto Alves F, Porch T, Beaver J, Urrea CA, Cichy K (2022) The Phaseolus vulgaris L. yellow bean collection: genetic diversity and characterization for cooking time. Genet Resour Crop Evol. https://doi.org/10.1007/s10722-021-01323-0
Savic A, Pipan B, Vasic M, Meglic V (2021) Genetic diversity of common bean (Phaseolus vulgaris L.) germplasm from Serbia, as revealed by single sequence repeats (SSR). Sci Hortic 288:110405
Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234
Sicard D, Nanni L, Porfiri O, Bulfon D, Papa R (2005) Genetic diversity of Phaseolus vulgaris L. and P. coccineus L. landraces in central Italy. Plant Breed 124:464–472
Souza SGH, Pipolo VC, Ruas CF (2008) Comparative analysis of genetic diversity among the maize inbred lines (Zea mays L.) obtained by RAPD and SSR markers. Braz Archiv Biol Technol 51(1):183–192
Toksoy S, Ozturk M, Sagiroglu M (2015) Phylogenetic and cladistic analyses of the enigmatic genera Bituminaria and Cullen (Fabaceae) in Turkey. Turk J Bot 39:60–69
TUIK 2021 Bitkisel üretim istatistikleri (Crop Production Statistics). http://www.tuik.gov.tr [Access date: 07 February 2021]
Ulukapi K, Onus AN (2012) Molecular characterization of some selected landrace green bean (Phaseolus vulgaris L.) Genotypes. J Agric Sci 18:277–286
Valentini G, Gonçalves-Vidigal MC, Elias JCF, Moiana LD, Mindo NNA (2018) Population structure and genetic diversity of common bean accessions from Brazil. Plant Mol Biol Report 36:897–906
Yu K, Park SJ, Puya V, Gepts P (2000) Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (Phaseolus vulgaris L.). The Am Genet Assoc 91:429–434
Yuvaraj, M, Pandiyan, M, Gayathri, P (2020) Role of legumes in improving soil fertility status. In (Ed.), Legume Crops - Prospects, Production and Uses. IntechOpen.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.