Morphometric traits and iPBS based molecular characterizations of walnut (*Juglans regia* L.) genotypes

İbrahim Başak · Göksel Özer · Ferhad Muradoğlu

Received: 8 November 2021 / Accepted: 20 April 2022 / Published online: 16 May 2022 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

**Abstract** In this study, walnut genotypes that were selected during two growing seasons among thousands of seedlings were analyzed in terms of detailed morphometric, phenological, and chemical traits. A multivariate analysis was conducted with valuable traits for breeding and selection such as morphometric traits, chemical composition, and phenological characteristics. Also, genotypes were characterized by a retrotransposon-based iPBS marker system. The correlation analysis showed significant positive and negative correlations between agro-morphological characters. The principal component analysis explained 71.44% of the total variance into five main components. Principal component and hierarchical cluster analysis divided genotypes into three groups and identified subgroups based on both agro-morphological characters and iPBS marker systems. A high level of polymorphism ratio was observed for tested markers. Mantel’s test demonstrated relatively low correlations between molecular and morphological traits ($r=0.04$). The genetic similarities among all individuals ranged from 0.39 (between 018 and 015 or 045 genotypes) to 0.98 (between 090 and 094 genotypes) with a mean similarity of 0.67. Remarkable phenotypic and molecular variations were observed among the genotypes. The features of some investigated genotypes were above the acceptable thresholds for walnut selection in breeding programs, and our study indicated that iPBS markers can be beneficial in walnut breeding programs, allowing the evaluation of the genetic relationship between genotypes, helping to differentiate and select the best genotypes to improve agronomic properties.

**Keywords** Genetic resources · Agromorphological · Phenotyping · Retrotransposons · Multivariate analysis

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10722-022-01394-7.

İ. Başak
Department of Horticulture, Agriculture Faculty, Van Yüzüncü Yıl University, Van, Turkey

G. Özer
Department of Plant Protection, Agriculture Faculty, Bolu Abant Izzet Baysal University, Bolu, Turkey

F. Muradoğlu
Department of Horticulture, Agriculture Faculty, Bolu Abant Izzet Baysal University, Bolu, Turkey
e-mail: muradogluf@ibu.edu.tr

**Introduction**

The *Juglans* genus includes about 20 walnut species of deciduous trees of the family Juglandaceae. The economically and commercially cultivated species in this genus is the Persian walnut (*Juglans regia* L.), which is a monoecious and self-compatible species. Walnuts are pollinated by windy dichogamy, which provides open pollination and is an important source
of genetic diversity for suitable opportunities in walnut breeding to obtain new cultivars (Cosmulescu and Botu 2012). Because of the open pollination, Turkish walnut populations have high genetic variability consisting of millions of natural hybrids grown on their own roots, which are important sources of genetic diversity for J. regia (Akça 2016). This genetic variation in the native walnut populations presents many opportunities for walnut breeding (Karadağ and Akça 2011). Thus, studies of phenological and morphological traits in natural walnut populations have remained for some time, while morphological characteristics often do not indicate clear relationships between genotypes, as the results can be affected by changing environmental conditions (Kumar 1999). Therefore, several techniques have been developed to estimate genetic diversity in walnut genotypes (Sharma and Sharma 2001). Walnuts trees have long-term infertility in adolescents, and breeding studies take long-time periods; thus, molecular markers have been used in walnut breeding to overcome these difficulties. Molecular markers are one of the most important methods for the characterization of genotypes and determining genetic resources (Badenes and Parfitt 1998; Li and Quiros 2000). Several marker techniques have applied to reduce the duration of breeding programs, examine relationships between genotypes, and perform genotype selections with greater precision. To identify phylogenetic relationships and characterize germplasm, molecular techniques including RFLP (Restriction Fragment Length Polymorphism: Fjellstrom et al. 1994), RAPD (Randomly Amplified Polymorphic DNA: Nicese et al. 1998; Doğan et al. 2014), ISSR (Inter Simple Sequence Repeat: Christopoulos et al. 2010; İpek et al. 2019), AFLP (Amplified Fragment Length Polymorphism: Kañkas et al. 2005), and SSR (Simple Sequence Repeats: Orhan et al. 2020) have been employed in recent years.

Retrotransposons are repetitive and mobile DNA fragments that can copy themselves to another region of the genome where they are located. In this way, the genome size can be increased, and variations can be created in the genome, making retrotransposons the best tools for molecular markers (Kalander et al. 2010). For this reason, it is considered an excellent source of markers since it causes mutations, increased genome size, and ultimately genetic variations (Schulman et al. 2004). Retrotransposon is abundant in the entire genome of eukaryotic cells, particularly plant species (Finnegan 1989). Plant genomes vary in their retrotransposon content and often makeup 50–90% of the plant genome (San Miguel et al. 1996). In recent years, various retrotransposon marker systems have been used extensively in many studies on evolutionary and genetic diversity due to their general applicability, ease of use, and providing high genotype resolution among eukaryotic organisms (Schulman et al. 2004; Kalander and Schulman 2006). Because of the high potential for marker generation, retrotransposon markers such as inter-retrotransposon amplified polymorphism (IRAP) and retrotransposon-microsatellite amplified polymorphism (REMAP) have been used for characterization and gene mapping of many eukaryotes as well as plants (Castro et al. 2012; Kaya and Yılmaz-Gökdoğan 2016; El Zayat et al. 2021). These marker systems, however, have some bottlenecks, such as the size of various PCR products and the sequence information requirements for designing primers that match neighboring genomic DNA at any given location, preventing widespread use of them (Özer et al. 2016).

The inter Primer Binding Site (iPBS) retrotransposons-based amplification technology was declared by Kalander et al. (2010) as a universal DNA labeling method that can be used in plants and animals based on the primer binding site for the reverse transcription enzyme of long terminal repeats (LTR retrotransposon). The iPBS marker system has proven to be a powerful DNA fingerprinting technique that does not require sequence information and is the preferred universal marker system for the genetic differentiation of several eukaryotic organisms at both an intraspecific and interspecific level (Özer et al. 2016; Milovanov et al. 2019; Aydın et al. 2020; Pérez-Vargas et al. 2020; Erper et al. 2021; Ouyang et al. 2021). Molecular markers are one of the most important methods for the characterization of genotypes and determining genetic resources (Badenes and Parfitt 1998; Li and Quiros 2000).

Molecular markers including SSR, ISSR, AFLP, and RAPD have been used for generating genetic polymorphisms among walnut genotypes; however, no record of determination of genetic differences in walnut genotypes using iPBS markers exists. This study is the first characterization study of walnut genotypes using the iPBS retrotransposon marker system. Furthermore, relationships between nut agromorphological traits and genotypes were determined.
by multivariate analysis for the selection of desired features that are essential for breeding programs.

Materials and methods

Plant material and fruit measurements

The study was conducted on seedling walnut trees from the Şanlıurfa region in two consecutive growing seasons, 2015–2016 and 2016–2017, in Turkey (Fig. 1). Among thousands of seedling walnut genotypes in the investigated population, according to UPOV criteria, 121 genotypes which the mature, healthy, and full crops were labeled at optimal ripening time.

The walnuts were harvested when 50% of their husks were opened. After the husk was removed, walnuts were dried to reduce their moisture content to 10–12% immediately. The nut characteristics, including nut width (mm), nut length (mm), nut height (mm), nut weight (g), kernel weight (g), and shell thickness (mm) were determined according to Union for the Protection of New Varieties of Plants (UPOV) walnut descriptors (UPOV 2017). A digital caliper was used to measure the dimensions of nuts, and an electronic scale with an accuracy of 0.01 g was used for nut weights with 30 nuts in the replications for each measurement. Kernel percentage was calculated using the formula ‘kernel weight/nut weight × 100’.

The shell color was classified into three categories “light”, “amber”, and “dark” and the kernel color was evaluated using the design for assembly (DFA) guidelines (United States Standards for Grades of Shelled Walnuts, 2017). The DFA is based on a chart for color evaluation that classified kernels into one of four categories “extra light”, “light”, “light amber”, and “amber”.

After the first evaluations, 20 genotypes with the highest value according to the UPOV descriptors from 121 genotypes were determined as “promising”. The further nut and phenological characters (first leafing, the receptive period in male and female flowers, dichogamy, % of fruitful laterals (FL), harvest date, and chemical compositions were estimated for the promising genotypes for two consecutive years, 2015 and 2016.

Chemical compositions

The kernel of the walnut was milled and moisture (MS) determinations were made (AOAC 1995). The protein (PRT) amount was determined by the modified Kjeldahl method (AOAC 1990). The total fat (OIL) content was obtained by extracting 10 g of ground crushed walnut kernel petroleum ether at 45–50 °C for 8–9 h on a soxhlet device and weighing the pre-and post-extraction weights (AOAC 1995).

Molecular characterizations

In promising genotypes, the genetic variation was determined using the iPBS molecular markers. The leave samples from promising genotypes were taken into sterile 50 ml tubes, labeled, and stored in an icebox to transfer to the laboratory. A leaf of each sample was frozen with liquid nitrogen and homogenized with a mortar and a pestle under liquid nitrogen. To extract genomic DNA, approximately 100 mg grounded tissue was put into a 1.5 ml microcentrifuge tube, and treated with preheated (65 °C) extraction/lysis buffer according to “Plant DNA Extraction Protocol for DArT” (https://www.diversityarrays.com), followed by application of twice chloroform: isoamyl alcohol (24:1) mixture, and a precipitation stage was conducted with ice-cold isopropanol. The final quality and quantity
of the resultant DNA were estimated spectrophotometrically by the A260/A280 ratio determined by the DS-11 FX+ nano spectrophotometer (Denovix Inc., Wilmington, DE, USA) and diluted to 10 ng/μl with sterile ultra-pure water. Template DNA samples were stored at −20 °C till used.

Seven iPBS markers (Table 3) designated by Kalendar et al. (2010) were selected for molecular characterization of walnut genotypes collected in the Şanlıurfa region. PCR amplifications were conducted in a 25 μL reaction mix containing 1× Dream Taq Buffer, 0.2 mM of dNTPs, 1 μM of primer, 1.2-unit of Dream Taq DNA polymerase (Thermo Fischer Scientific, Waltham, MA, USA), 0.04-unit of Pfu DNA polymerase (Thermo Fischer Scientific, Waltham, MA, USA) and 20 ng template DNA. PCR reactions were carried out in a T100 thermocycler (Bio-Rad, Hercules, CA, USA) as follows: 35 cycles of 30 s at 94 °C, 30 s at 52–60 °C (depending on primers), and 2 min at 72 °C. All reactions included an initial denaturation step for 3 min at 94 °C and an ending step for 10 min at 72 °C. The amplified fragments were separated on a 1.6% (w/v) agarose gel with 1× TAE buffer for over 2 h and stained with ethidium bromide before visualizing via a UV transilluminator (G: BOX F3, Syngene, UK).

All the PCR amplifications were repeated at least two times, and only reproducible bands were evaluated. All bands from iPBS analysis were scored as present (1) or absent (0) at positions to construct a binary data matrix. The data matrix from marker systems was combined and converted into a genetic similarity matrix using Jaccard’s similarity coefficient and the Numerical Taxonomy System, NTSYS-pc version 2.10 program (Rohlf 2000) was used to conduct an unweighted pair group method using arithmetic average (UPGMA). To calculate resolving power (RP), the band informativeness (Ib) was calculated first, as follows: 
$I_b = 1 - (2 \times (0.5 - p))$, where p represents the proportion of promising genotypes. Then RP value was calculated as the sum of Ib for each band obtained for each marker (Prevost and Wilkinson 1999). The the polymorphic information content (PIC) value for each marker was calculated as $PIC = 2f (1 - f)$; where f is the frequency of the amplified band (Roldán-Ruiz et al. 2000).

### Statistical analysis

Data were subject to analysis of variance (ANOVA) using R software (Allaire 2012; R Core Team 2020). The minimum and maximum values with standard deviation and the coefficient of variation (CV%; SD/mean×100) were calculated for the measured characters. The coefficient of variation was used as a variability index. Pearson’s correlation was performed to determine relationships among agro-morphological characters using R Studio software by the package of ‘corrplot’ (Wei and Simko 2017). The Mantel test between the agro-morphological trait distance matrix and Nei’s genetic distance matrix was conducted with GenAlEx (Peakall and Smouse 2012). The relationships of the studied characters and genotypes with each other were determined by principal component analysis (PCA) with the package of ‘ggplot2’ of R software (Wickham 2016).

### Results and discussion

#### Morphometric and phenological characters

The genotypes showed significant differences in agro-morphological features (ANOVA, $P<0.05$). The minimum, maximum, mean, and coefficient variations (CV) values of the studied characters are presented in Table 1. The highest CV was observed in the shell color (51.29%), followed by kernel color (45.48%) and leafing date (45.16%), while the lowest CV was observed in the kernel ratio (6.47%) and nut length (6.77%). In total, the CV in 8 out of 17 measured characters was smaller than 10% that indicating low variations among the genotypes; thus, these characters may be considered more stable characters. Similarly, Khadivi et al. (2019) reported that CV for nut diameter, nut length, nut weight, and kernel weight were 11.03, 12.89, 23.34, and 25.99%, respectively. Moreover, Poggetti et al. (2017), reported 32.0% CV for nut weight, 28.7% for kernel weight, 25.0% for shell thickness and 48.0% for kernel skin color.

As shown in Table 2, fruitful laterals, leafing date, kernel color, shell color, the receptive period in male and female flowers, and harvest date showed more than 15.0% CV, indicating a high level of variation, and these vegetative growth parameters affected by genetical and ecological conditions.
In walnut breeding, fruit characteristics are an important parameter for domestic and foreign markets. Nut weight, kernel weight, shell thickness, and kernel ratio are important quality parameters. These parameters in our study varied between 7.90–15.52 g, 4.15–7.55 g, 1.04–1.59 mm, and 45.25–56.12%, respectively (Table 1; Supp. Data 1). In walnut breeding, the acceptable value should be 6–8 g for nut weight, 50–70% for kernel ratio, and 0.7–1.5 mm for shell thickness (Zhadan and Strokov 1977; Akça 2016). All of our samples were above these critical thresholds. Nut width ranged from 26.78 to 35.99 mm, while nut length varied from 28.38 to 35.73 mm. Fruit height varied from 34.35 to 45.14 mm. These values are consistent with Mahmoodi et al. (2019) but are lower than Khadivi et al. (2019). Nut width is worthy for growers who accept a minimum grade of 32.0 mm, and a bigger nut diameter means more income for growers. More than half of our samples yielded above the base limit. Skin and kernel color are very important for walnut breeding, while commercially acceptable walnut

| Table 1  | Descriptive statistics for used characters utilized in the promising walnut genotypes |
|----------|-----------------------------------------------|
| Characters | Abbreviation | Unit/code | Minimum | Maximum | Mean | SD | CV |
| Nut width | NW | mm | 26.78 | 35.99 | 31.42 | 0.55 | 7.80 |
| Nut length | NL | mm | 28.38 | 35.73 | 32.69 | 0.49 | 6.77 |
| Nut height | NH | mm | 34.35 | 45.14 | 39.59 | 0.71 | 8.05 |
| Nut weight | NWT | g | 7.90 | 15.52 | 12.17 | 0.43 | 15.93 |
| Kernel weight | KWT | g | 4.15 | 7.55 | 6.05 | 0.20 | 14.67 |
| Shell thickness | ST | mm | 1.04 | 1.59 | 1.36 | 0.03 | 9.58 |
| Kernel ration | KR | % | 45.25 | 56.12 | 50.53 | 0.73 | 6.47 |
| Moisture | MS | % | 2.41 | 3.20 | 2.80 | 0.06 | 8.77 |
| Total oil | OIL | % | 50.49 | 62.50 | 55.72 | 1.04 | 8.36 |
| Protein | PRT | % | 15.40 | 20.74 | 17.96 | 0.34 | 8.41 |
| Fruitful laterals | FL | % | 40.00 | 75.00 | 57.75 | 1.76 | 13.61 |
| Leafing date | LD | Date | 1.00 | 4.00 | 1.80 | 0.20 | 49.69 |
| Kernel color | KC | Code | 1.00 | 3.00 | 1.95 | 0.19 | 45.48 |
| Shell color | SC | Code | 1.00 | 3.00 | 1.20 | 0.13 | 51.29 |
| Receptive period in male flowers | RPMF | Date | 1.00 | 4.00 | 2.95 | 0.16 | 25.73 |
| Receptive period in female flowers | RPFF | Date | 1.00 | 4.00 | 3.05 | 0.24 | 36.03 |
| Altitude | Alt | m | 527 | 600 | 550.9 | 6.52 | 5.29 |
| Harvest date | HT | Date | 1.00 | 2.00 | 1.90 | 0.06 | 16.19 |

| Table 2  | Frequency distribution for agronomic characters utilized in the promising walnut genotypes |
|----------|-----------------------------------------------|
| Characters | Abbreviation | Unit/code | Frequency % |
| Leafing date | FFL | Date | Very early (45.00) Early (35.00) Medium (15.00) Late (5.00) |
| Kernel color | KC | Code | Extra light (40.00) Light (25.00) Light amber (35.00) |
| Shell color | SC | Code | Light (90.00) Dark (10.00) |
| Receptive period in male flowers | RPMF | Date | Very early (5.00) Early (15.00) Medium (60.00) Late (20.00) |
| Receptive period in female flowers | RPFF | Date | Very early (10.00) Early (15.00) Medium (30.00) Late (45.00) |
| Dichogamy | DC | Code | Protandry (100.00) |
| Harvest date | HT | Date | Early (10.00) Medium (90.00) |
skin and kernel color should be ‘light’ for skin and ‘extra light’ or ‘light’ for the kernel. A total of 18 genotypes (90%) had ‘extra light’ or ‘light’ kernel skin color, and only 2 genotypes were ‘dark’. Kernel color of 8 genotypes (40%) was observed as ‘extra light’, 5 (25%) as ‘light’, and 7 (35%) as ‘light amber’. The 90% and 60% of genotypes were acceptable for commercial cultivation in terms of skin color and kernel color, respectively.

Significant differences ($P < 0.05$) were observed among genotypes in terms of the chemical compositions of kernels. Total oil, the predominant component, varied from 50.49 to 62.50%, followed by protein ranging from 15.40 to 20.74%, and moisture from 2.41 to 3.20%. Oil and protein consisted of approximately 70–80% of kernel weight (Table 1; Supp. Data 2). These results are by previously reported values in several commercial varieties studied by researchers who found walnut kernel as compensating 40–60% of walnut weight, while about 60–70% of the kernel weight was oil and 24% was protein (Muradoğlu et al. 2010; Kabiri et al. 2019; Verma et al. 2020). Genotypes with a kernel that has 60% or above oil and 12% or above protein are more desirable in walnut breeding programs. The genotypes that yielded higher oil and protein ratios were considered precious materials for future breeding programs. Furthermore, the walnut chemical composition varies depending on many factors such as genetic characteristics, maturity, environmental factors, harvest date, and soil characteristics.

The late spring frosts are an important factor limiting walnut cultivation. For this reason, the flowering date of secondary branches is the main requisition of walnut breeding programs. The including leafing dates are evaluated relatively very early, early, medium, and late within this period. In our study, leafing date of one genotype was ‘late’, and three genotypes were ‘medium’ (Table 2; Supp. Data 3). The receptive periods in males were from 20–22 to 28–30 March and from 25–27 March to 1–3 April in females that are relative to very early, early, late, and very late flowering. The flowering and female flower formation are important factors affecting the productivity of walnut. In our study, the receptive period of male flowers was determined ‘medium’ in eight genotypes and ‘late’ in twelve genotypes. In addition, the female flowers were ‘medium’ in six genotypes and ‘late’ in nine genotypes in terms of flowering dates.

Walnut cultivars or genotypes are generally tending dichogamy, and especially protandry is seen more than protogyny and homogamy, which provides an advantage in terms of pollination and productivity. The blooming types of all promising genotypes were protandry (Table 2; Supp. Data 3). Generally, *J. regia* has dominant characteristics of protandry in dichogamy, and it is affected by climatic conditions at the flowering time. The development of catkins (male flowers) is highly affected by the temperature compared to female flowers (Akça 2016; Kumar and Sharma 2013). Persian walnut has two bearing habits. Fruit set can occur in two ways both only from the terminal buds of new branches or from both terminal and lateral buds, crucial characteristics in the breeding programs that are affected by phytohormones such as IAA, and zeatin, and plant nutrition (Muradoğlu

| iPBS ID | Primer sequences (5′–3′) | Tm (°C) | GC (%) | TB | PB | PPB (%) | PIC | RP |
|--------|--------------------------|---------|--------|----|----|---------|-----|----|
| 2081   | GCAACGGCGCACA            | 56      | 75.0   | 13 | 10 | 76.92   | 0.22| 3.90|
| 2387   | GCGCAATACCA              | 52      | 58.3   | 14 | 10 | 71.43   | 0.20| 3.80|
| 2219   | GAACCTTATGCCGATACCA      | 52      | 44.4   | 11 | 8  | 72.73   | 0.24| 3.70|
| 2078   | GCGGAGTGCACA             | 54      | 75.0   | 12 | 8  | 66.67   | 0.21| 3.80|
| 2395   | TCCCCACGGAGTCGCA         | 60      | 72.2   | 14 | 11 | 78.57   | 0.23| 4.40|
| 2077   | CTGACGATGCA              | 50      | 58.3   | 10 | 6  | 60.00   | 0.15| 2.20|
| 2242   | GCCCCATGGTTGGCGCCA       | 58      | 77.8   | 12 | 9  | 75.00   | 0.16| 2.54|
| Total  |                         |         |        |    |    |         |     |    |
| Avg./primer |                   | 12.29 | 8.86  | 72.09 | 0.20 | 3.48 |

$Tm$ (°C), annealing temperature; GC (%), percentage of guanine-cytosine content; TB, total band; PB, polymorphic band; PPB (%), percentage of the polymorphic band; PIC, polymorphism information content; RP, resolving power.
et al. 2009; Amiri et al. 2010). In our study, the rate of fruiting in the sub-branches was between 40 and 75%. The harvest dates of the genotypes took place from 1 to 7 October (Table 2; Supp. Data 3).

Molecular characters

The seven iPBS primers yielded 86 scorable and reproducible fragments to evaluate the extent of genetic variation among 20 promising genotypes. Sixty-two of those fragments (72.09%) were polymorphic. The number of fragments generated with the primers ranged from 10 (2077) to 14 (2387/2395), with a mean of 12.29 fragments per iPBS primer. The values of the PIC and RP index calculated for these primers are shown in Table 3. The range of PIC was 0.15 (2077)-0.24 (2219), averaging 0.20, which was consistent with the PIC value (0.28) for ISSR markers among 56 walnut accessions (Christopoulos et al. 2010). The mean of the RP values, a parameter that indicates the discriminatory potential of the primers selected, was on average 3.48 for all primers. The highest RP value of 4.40 was recorded for 2395, while the lowest RP value was observed at 2.20 for 2077. Similarly, Christopoulos et al. (2010) found a 4.52 RP value for Greek walnut populations. The band profiles obtained with iPBS markers (Fig. 2) were validated by the gel-to-gel normalization of migrations using 100 bp DNA ladder (Solis BioDyne, Tartu, Estonia) from each end of the gel was used. The sizes of reproducible and scorable bands ranged from 200 to 4000 bp.

Correlations and multivariate analysis among the characters

Pearson’s pairwise correlation revealed the relationships among characteristics (Fig. 3). The correlations with highly significant values were observed among most of the studied characters. Significant positive correlations were detected among nut variables such as nut width, nut length, nut weight, and kernel weight. The kernel ratio was negatively correlated with nut width, nut height, nut weight, and shell thickness. The results were in line with previous reports by Cosmulescu and Botu (2012), Poggetti et al. (2017), and Khadivi et al. (2019). Knowing the relationships among phenotypic characters can lead breeders in breeding studies. In addition, a linear correlation among different properties indicates that an improvement of a property may evolve the other related traits (Yucel et al. 2009). Leafing date was positively correlated with shell color and was not significantly correlated with lateral branch fruitfulness, harvest date, and altitude. This situation proves that the nut characteristics and the phenology change parallelly in similar altitude and ecological conditions.

PCA is used as a descriptive method to evaluate prominent features in the data. PCA provides the opportunity to minimization of effective factors according to individuals. These superior features
made it an essential part of breeding and population genetics research and were frequently used in recent years (Hashemi and Khadivi 2020). The PCA analysis was performed based on nut characterizations with selected walnut genotypes. As a result, PCA indicated that 71.44% of the observed variability was described by the first five components. The first two components described nearly 44% of the total variability. The relationships among genotypes are presented in Fig. 4. Nut weight, kernel weight, nut width, nut height, and nut length had the highest effects on PC1, and they comprised 28.6% of the total variability. PC2 explained 14.99% of total variation and was strongly related to nut and agronomic traits, including FL, SC, RPFF, and RPMF. PC3 showed crucial positive variability with oil and negative variability with LD and HD (Supp. Data 4). In our study, the first two PCs were significant at the level of 99%, while the third

**Fig. 3** Correlation coefficients between major nut characteristics in walnut genotypes. * and ** indicates significance at $P \leq 0.05$ and $P \leq 0.01$, respectively. NW: nut width, NL: nut length, NH: nut height, NWT: nut weight, KWT: kernel weight, ST: shell thickness, KR: kernel ratio, FL: fruitful laterals, MS: moisture, OIL: total oil, PRT: protein, FD: first leafing, KC: kernel color, SC: shell color, RPMF: receptive period in male flowers, RPFF: receptive period in female flowers, ALT: altitude, HD: harvest date,
Principal components with a statistical significance, the first three PCs, explained 54.77% of the total variance and indicated a high phenotypic diversity in agro-morphological characters for studied genotypes. Similar to our study, previous studies revealed a high phenotypic diversity in pomological characteristics of walnut genotypes (Khadivi et al. 2019; Farrokhi Toolir and Mozaffari 2020; Skender et al. 2020).

Hierarchical cluster analysis based on agro-morphological characters classified the genotypes into three major groups as supported by the biplot (Figs. 4, 5B). The first group, BI bifurcated into two subgroups: BI-I included 008, 018, and 026.
genotypes, while BI-II consisted of 015, 031, 29, 48, and 103 genotypes. The second group, BII, was bifurcated into two sub-groups, and the primary group (BII-I) was also partitioned into two sub-groups consisting of 030, 036, and 076 genotypes one group and 061, 094, and 090 genotypes respectively. The second subgroup (BIII-II) formed of two genotypes, 045 and 073 genotypes. The third group (BIII) bifurcated into two sub-groups, and both two groups (BIII-I) and (BIII-II) formed 027, 118 genotypes and 075, 089 genotypes, respectively. The highest similarity was between 061 and 094 genotypes, while the farthest distance was observed between 008 and 027 genotypes. (Fig. 4). The distribution of the genotypes depends on agro-morphological characters on the biplot supported by the cluster analysis (Figs. 4, 5B).

The UPGMA dendrogram produced using the Jaccard’s similarity coefficient for iPBS profiling methods clustered distinctly 20 walnut genotypes into three major groups (Fig. 5A). From selected walnut genotypes, two genotypes (030 and 075) unite in the first group and the first branch (AI-I) and five genotypes (094, 090, 045, 015, and 020) combine the second branching (AI-II). Within this group, the highest genetic similarity rate was determined as 0.98% between 090 and 094 genotypes. In the second group, one genotype (018) was brought to the first branching. In the third group, the genotypes 026 and 027 formed the first branch (AIII-I), and the second (AIII-II) branch consisted of two sub-groups having 008, 031, 073, 118, 048, 061, 036, and 076 genotypes, and 089 and 103 genotypes in other groups. The highest similarity rate in this group was determined between

Fig. 6 Mantel test correlation matrixes between walnut characters and combined molecular markers

In the study, the rate of polymorphism obtained from iPBS (72.9%) markers were higher than previously reported with the RAPD (69.1%) molecular marker technique by Doğan et al. (2014), with the ISSR (61.1%) technique by İpek et al. (2019), and with the ISSR (71.1%) by Doğan et al. (2014). It is known that iPBS-retrotransposon markers are more informative than other molecular marker systems such as IRAP, RAPD, ISSR, and SRAP (Boronnikova and Kalendar 2010). These results showed that the iPBS-retrotransposon technique could be used to determine genetic variation in walnut breeding.

The Mantel test was performed to determine correlations between genetic and morphological distances of genotypes. As shown in Fig. 6, the mantel test resulted in a very low correlation ($r=0.04$). The low correlation between morphology and genetics indicates that molecular markers do not separate genotypes only with the morphological locus but also with a whole genome. We previously stated that iPBS retrotransposon markers are universal markers and use the whole genome. Similar results were observed between morphological and genetic (SRAP) distances ($r=0.03$) by Doğan et al. (2014).

The genetic variation between genotypes was determined for the first time using iPBS markers based on retrotransposons which is a newer DNA marker system than the ISSR and RAPD methods. The phenological and morphological characters are rather affected by environmental conditions. So, during recent decades many studies have focused on morphological traits with molecular techniques, such
as DNA-based markers SSR, ISSR, AFLP, RAPD, and SNP (Single-nucleotide polymorphism) to reveal genetic diversity, relationships among cultivars or genotypes, explain phylogenetically and origin in walnut (Fjellstrom et al. 1994; Nicese et al. 1998; Kafkas et al. 2005; Christopoulos et al. 2010; Orhan et al. 2020; Houmanata et al. 2021). The percentage polymorphism and the average polymorphic band per primer detected in our present results were higher than previously found by İpek et al. (2019) and Doğan et al. (2014). İpek et al. (2019) reported that 61.1% of DNA polymorphism was generated by 17 primers; 51 of these markers were polymorphic in walnut cultivars and genotypes. The report deals with the results of Doğan et al. (2014) who used three molecular markers for characterizations of 59 walnut genotypes and detected 181 well-resolved and clear bands were generated by 25 primers and 129 total polymorphic bands with 71.1% polymorphism percentage in RAPD, 117 total polymorphic bands from 118 with 69.1% polymorphism percentage in ISSR, and by 16 primer and 129 total polymorphic bands with 99.1% polymorphism percentage by 16 primers in SSR marker system. In our study using the iPBS primers, 72.09% polymorphism percentages were observed in walnut genotypes. Previous studies also confirmed that the retrotransposon marker system could be successfully used for molecular characterizations of different fruits such as grapevine (Castro et al. 2012), apple (Kuras et al. 2013), hawthorn (Rahmani et al. 2015), Pictasia species (Kirdök and Çiftçi 2016), olive (Kaya and Yılmaz-Gökdoğan 2016), Citrus species (El Zayat et al. 2021), peach, and nectarine (Naeem et al. 2021).

Conclusions

In this study, the iPBS retrotransposon markers were first time used in walnut and resulted in high polymorphism rates that make it an excellent tool to use in future breeding programs. Moreover, our results clearly proved that morphological variability and relationships are related to agro-ecological conditions. According to morphological properties, 008, 029, 061, 075, and 076 genotypes were superior, having greater values than the base limit of walnut breeding criteria. These genotypes stood out as an important variety candidate with their lateral branch productivity, fruit weight, kernel weight, kernel ratio, protein, and oil content.

Acknowledgements Data of the manuscript were received from the Ph.D. thesis of İbrahim Başak

Author contributions İB performed field experiments. İB and FM carried out the laboratory studies. FM performed statistical analysis and data visualization. GÖ performed molecular studies. FM, GÖ, and İB wrote and reviewed the original draft.

Declarations

Conflict of interest The authors declare there is not any known financial or personal competing interest.

References

Akça Y (2016) Walnut production. Anıt Matbaası
Allaire J (2012) R Studio: integrated development environment for R. Boston, MA 770(394):165–171
Amiri R, Vahdati K, Monseripoor S, Mozaffari MR, Leslie C (2010) Correlations between some horticultural traits in walnut. HortScience 45:1690–1694
AOAC (1990) Official Methods of Analysis, 15th edn. AOAC, Washington, DC
AOAC (1995) Official methods of analysis, 16th edn. Association of Official Analytical Chemists, Washington DC
Aydn F, Özer G, Alkan M, Çakır I (2020) The utility of iPBS retrotransposons markers to analyze genetic variation in yeast. Int J Food Microbiol 325:108647. https://doi.org/10.1016/j.ijfoodmicro.2020.108647
Badenes ML, Parfitt DE (1998) Phylogeny of the genus pistacia as determined from analysis of the chloroplast genome. Proc Natl Acad Sci USA 7:25–26
Boronnikova SV, Kalendar RN (2010) Using IRAP markers for analysis of genetic variability in populations of resource and rare species of plants. Russ J Genet 46:36–42
Castro I, D’Onofrio C, Martín JP, Mari’a Ortiz J, De Lorenzis G, Ferreira V, Pinto-Carnide O (2012) Effectiveness of AFLPs and retrotransposon-based markers for the Identification of portuguese grapevine cultivars and clones. Mol Biotechnol 52:26–39
Christopoulos MV, Roukas D, Tsantili E, Bebeli PJ (2010) Germplasm diversity and genetic relationships among walnut (Juglans regia L.) cultivars and Greek local selections revealed by Inter Simple Sequence Repeat (ISSR) marker. Sci Hortic 125(4):584–592
Cosmulescu S, Botu M (2012) Walnut biodiversity in South-Western Romania-resource for perspective cultivars. Pak J Bot 44(1):307–311
Doğan Y, Kafkas S, Süttemez M, Akça Y, Türemiş N (2014) Assessment and characterization of genetic relationships of walnut (Juglans regia L.) genotypes by three types of molecular markers. Sci Hortic 168:81–87
El Zayat MAS, Hassan AH, Nishawy E, Ali M, Amar MH (2021) Patterns of genetic structure and evidence of
Kumar LS (1999) DNA markers in plant improvement: an overview. Biotechnol Adv 17:143–182
Kumar A, Sharma N (2013) Protandrous protogynous dimorphism in indigenous selections from North Western India and some exotic cultivars of Persian walnut (Juglans regia L.). Adv Hortic Sci 61–66
Kuras A, Antonius K, Kalendar R, Kruczy´nska D, Korbín M (2013) Application of five DNA marker techniques to distinguish between five apple (Malus × domestica Borkh.) cultivars and their sports. J Hortic Sci Biotechnol 88(6):790–794
Li G, Quiros CF (2000) Use of amplified fragment length polymorphism markers for celery cultivar identification. HortScience 35(4):726–728
Mahmoodi R, Reza Dadpour M, Hassani D, Zeinalabedini M, Vendrarn E, Micali S, Zaare NF (2019) Development of a core collection in Iranian walnut (Juglans regia L.) germplasm using the phenotypic diversity. Sci Hortic 249:439–448
Milovanov A, Zvyagin A, Daniyarov A, Kalendar R, Troshlin (2019) Genetic analysis of the grapevine genotypes of the Russian Vitis ampelographic collection using iPBS markers. Genetica 147(1):91–101
Muradoğlu F, Balta F, Battal P (2009) Endogenous hormone levels in bearing and non-bearing shoots of walnut (Juglans regia L.) and their mutual relationships. Acta Physiol Plant 32(1):53–57
Muradoğlu F, Oğuz Hİ, Yildiz K, Yilmaz H (2010) Some chemical composition of walnut (Juglans regia L.) selections from eastern Turkey. Afr J Agric Res 5(17):2379–2385
Naeem H, Awan FS, Dracatos PM, Sa’jid MW, Saleem S, Yousafi Q, Khan MS, Mehmood A, Zulﬁqar B (2021) Population structure and phylgenic relationship of Peach [Prunus persica (L.) Batsch] and Nectarine [Prunus persica var. nucipersica (L.) C.K. Schneider] based on retrotransposon markers. Genet Resour Crop Evol 68:3011–3023
Niecee FP, Hormaza JI, McG Granahan GH (1998) Molecular characterization and genetic relatedness among walnut (Juglans regia L.) genotypes based on RAPD markers. Euphytica 101:199–206
Orhan O, Ey杜兰an SP, Poluja C, Akin M, Weber T, Erci B (2020) Genetic diversity detection of seed-propagated walnut (Juglans regia L.) germplasm from Eastern Anatolia using SSR markers. Folia Hortic 32(1):37–46
Ouyang Z, Wang Y, Ma T, Kanzana G, Wu F, Zhang J (2021) Genome-wide identification and development of LTR retrotransposon-based molecular markers for the Melilotus genus. Plants 10(5):890
Özer G, Bayraktar H, Baloch FS (2016) iPBS retrotransposons ‘A Universal Retrotransposons’ now in molecular phylogeny of fungal pathogens. Biochem Syst Ecol 68:142–147
Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinform 28:2537–2539
Pérez-Vargas I, Portero Alvarez AM, Pérez de Paz PL, Pérez JA (2020) Retrotransposon-based molecular markers as a tool in delimiting species in section Rynocholotus, a recent radiation group of Macaronesian Lotus. Syst
