Scientia Horticulturae 213 (2016) 1–4

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Novel 307 polymorphic SSR markers from BAC-end sequences in walnut (*Juglans regia* L.): Effects of motif types and repeat lengths on polymorphism and genetic diversity

Adi Surya Ikhsan [a], Hayat Topçu [a], Mehmet Sütyemez [b], Salih Kafkas [a,*]

[a] Department of Horticulture, Faculty of Agriculture, University of Çukurova, Adana 01330, Turkey
[b] Department of Horticulture, Faculty of Agriculture, University of Sütçüimam, Kahramanmaraş, Turkey

ARTICLE INFO

Article history:
Received 30 July 2016
Received in revised form 4 October 2016
Accepted 8 October 2016

Keywords:
Walnut
J. regia
SSR
BES-SSR

ABSTRACT

Walnut, *Juglans regia* L., has been cultivated for its edible nut and timber since ancient times. It is one of the most valuable and widely cultivated horticultural commodities in the world. Simple sequence repeats (SSRs) have a wide range of applications in plants, and there is still a necessity to develop novel SSR markers for each plant species. Bacterial artificial chromosome-end sequences (BES) are good sources to generate novel SSR markers. In pursuit of this objective, a total of 558 BES-SSR primer pairs were designed for *J. regia*. Of them, 507 (91%) had amplifications, and 307 were polymorphic when tested in eight walnut cultivars. A total of 1097 alleles were generated from 307 polymorphic SSR loci, ranging from two to eleven, with an average of 3.6 per locus. Polymorphism information contents (PIC) varied from 0.11 to 0.88 with an average of 0.46. Di-nucleotide motifs had higher polymorphism and genetic diversity values than the other motif types. The rate of polymorphism increased with the repeat length, therefore, class I SSRs are useful source of polymorphic DNA markers in walnut. Novel SSR markers developed in this study may possess potential applications such as fingerprinting, marker-assisted breeding, germplasm characterization and genetic linkage mapping in walnut and other species of *Juglans*.

1. Introduction

The genus *Juglans* includes 20 species, and belongs to the family Juglandaceae, *Juglans regia* L. (walnut) has edible nuts and a long history of cultivation (McGranahan and Leslie, 1990). Walnuts grow as a wild or semi cultivated tree in a wide area ranging from southeastern Europe and to the Caucasus area, although the origin of *J. regia* seems to be a large area in the Mountain range of central Asia. More than three million tons of walnuts are produced every year; and China, USA, Iran, Turkey and Ukraine are the main producer countries in the world (Faostat, 2016).

Its nuts are used worldwide in human nutrition as it contains high amounts of proteins, fats, vitamins, and minerals. The nuts are also a good source of flavonoids, sterols, polyphenols, and contains high amounts of essential dietary fatty acids, omega-6 and omega-3. Therefore, the walnut is classified as a strategic species for human nutrition and included in the FAO list of priority plants (Gandev, 2007; Taha and Al-wadaan, 2011).

In walnut, numerous molecular markers have been employed for germplasm characterization and genetic diversity studies (Fjellstrom et al., 1994; Nicese et al., 1998; Woeste et al., 2002; Dangl et al., 2005; Kafkas et al., 2005; Chen et al., 2014; Doğan et al., 2014). Simple sequence repeat markers (SSRs) are very convenient DNA fingerprinting method in comparison to other molecular markers owing to its codominant inheritance, large number of alleles per locus, abundance in genomes and high repeatability. SSRs also have a wide range of usage in plants such as evaluation of genetic diversity, cultivar identification, fingerprinting and the construction of genetic linkage maps. There are several studies that attempted development of SSRs for *J. regia*: Zhang et al. (2010) and Yi et al. (2011) used the EST database in the NCBI to develop a total of 148 SSRs. Najafi et al. (2014), Chen et al. (2014) and Topçu et al. (2015) generated 12, 13 and 185 SSRs for *J. regia*. There are also several reports on SSR development in other *Juglans* species such as *J. nigra*, *J. cinerea* and in *J. mandshurica* (Woeste et al., 2002; Hoban et al., 2008; Chen et al., 2013).

Nowadays, genomics is also employed in breeding programs. It assists in many approaches such as marker assisted selection (MAS), quantitative trait loci (QTL), association mapping and genomic selection. Bacterial artificial chromosome sequences

* Corresponding author.
E-mail address: skafkas@cu.edu.tr (S. Kafkas).

http://dx.doi.org/10.1016/j.scienta.2016.10.006
0304-4238/

(BACs) are often used to sequence the whole genome of an organism in the genome projects (Poulsen and Johnsen, 2004). Sequencing of the ends of BAC clones (BES) is an effective in producing low-pass genome wide sequences, which serve as a valuable resource in molecular marker development. Lai et al. (2006), Bartos et al. (2008) and Terol et al. (2008) generated novel SSR markers developed from BESs in papaya, apple and citrus, respectively. Wu et al. (2012) recently published BAC-end sequences of walnut, and are good resources for SSR marker development in *Juglans*.

Therefore, we aimed to develop novel SSR markers for walnut using BAC-end sequences (BES) available in the NCBI database. A total of 558 SSR loci were tested for polymorphism in this study, and the effects of motif types and repeat lengths on the genetic diversity values were also reported. The new BES-SSRs will provide an opportunity to construct a saturated SSR-based map, and will also be valuable for genetic diversity, fingerprinting, and germplasm characterization studies in walnut.

2. Material and methods

2.1. Plant materials and DNA extraction

In this study, eight *J. regia* cultivars from different origins were used to test new SSRs for polymorphism. Three cultivars were from Turkey (Kaplan-86, Maras-12 and Kaman-1), two were from France (Franquette and Fernette), and three were from California (Chandler, Serr and Midland). DNAs were extracted from young leaf tissues by CTAB method (Doyle and Doyle, 1987) with minor modifications (Kafkas et al., 2005). DNA concentrations were measured using a Qubit Fluorimeter (Invitrogen) or were estimated by comparing the band intensity with λ DNA of known concentrations following 0.8% agarose gel electrophoresis and ethidium bromide staining. DNA samples were subsequently diluted to a concentration of 10 ng/μL for SSR-polymerase chain reaction (PCR).

2.2. SSR primer design

The bacterial artificial chromosome (BAC) end sequences (BESs) constructed from *J. regia* cv. Chandler were retrieved from the National Center for Biotechnology Information (NCBI) database (Wu et al., 2012). SSR primer pairs were designed using the online software BatchPrimer3 v1.0 web-based software (You et al., 2008). The standard settings of the program were kept except for these parameters: max mispriming: 8; pair max mispriming: 16; min GC%: 40; max self-complementarity: 6; max 3'self-complementarity: 2; and max poly X: 4. The loci were given acronyms comprising 'JR' for *J. regia*, 'HR' and consequent number are the accession numbers in the NCBI, and consequent letter (a or b) indicates different SSR loci in the same clone. In this way, a total of 558 SSR primer pairs originally from BESs of *J. regia* were designed and tested in eight walnut cultivars for polymorphism.

2.3. SSR-PCR amplification

Gradient PCR was firstly conducted to determine the optimum annealing temperatures of the SSR primer pairs, and also to find if the primer pairs can generate amplification products. The gradient PCR products were run on 3% agarose gels and visualized under UV after ethidium bromide staining. If a primer pair had amplification, then PCR and capillary electrophoresis were performed to determine its allelic profile at the optimum annealing temperature. SSR-PCR reactions were carried out using a three primer strategy according to the method described by Scheulke (2000) with some modifications (Zaloğlu et al., 2015). PCR reactions, cycling conditions and capillary electrophoresis were performed as described by Topçu et al. (2015).

2.4. Data analysis

After capillary electrophoresis of the SSR loci, genetic diversity values such as number of alleles (*Na*), effective number of alleles (*Ne*), expected heterozygosity (*He*), and observed heterozygosity (*Ho*) were calculated using GenAlex v6.5 (Peakall and Smouse, 2012). The polymorphism information content (PIC) of each locus was calculated using PowerMarker software version 3.25 (Liu and Muse, 2005). The SSR loci were classified on the basis of their motif types and repeat numbers. SSRs that span ≥20 bp were classified as Class I and those with <20 bp were classified as Class II according to Temnykh et al. (2001). The genetic diversity values of Class I and Class II SSRs were separately calculated to determine the effect of motif types and repeat numbers on polymorphism.

3. Results

3.1. Development of BES-SSRs in walnut

We designed a total of 558 BES-SSR primer pairs from *J. regia*. The repetitive architecture of all SSR loci was perfect. Based on motif types, 51.1% were di-nucleotides, 9.5% were tri-nucleotides, 8.8% were tetra-nucleotides, 17.4% were penta-nucleotides and 13.2% were hexa-nucleotides.

Eight *J. regia* cultivars of different origins (Kaplan-86, Maras-12, Kaman-1, Franquette, Fernette, Chandler, Serr and Midland) were used to test polymorphism level of the SSR loci. Of tested 558 primer pairs for the ability in PCR amplification, 507 primer pairs (91%) had amplification patterns, while 51 (9%) failed in gradient PCR. Then, the amplified 507 SSR loci (Supplementary file 1) were tested for polymorphism in the eight walnut cultivars, and 307 were polymorphic (60.6%). A total of 1097 alleles were generated, ranging from 2 to 11, with an average of 3.6 per locus. The average observed heterozygosity (Ho) was 0.42, and the average expected heterozygosity (He) was 0.51 ranging from 0.12 to 0.89. Polymorphism information contents (PIC) of SSR loci varied from 0.11 to 0.88 with an average of 0.46. The highest PIC values in this study were obtained from JRHR223139, JRHR230700, JRHR226534 and 223389a loci. The highest Na (11), Ne (9.1), He (0.89) and PIC (0.88) values were obtained from JRHR223139 locus (Supplementary file 2).

3.2. Effect of motif type and repeat length on polymorphism and genetic diversity values

The polymorphic and monomorphic SSR markers for each motif type are illustrated in Fig. 1. Of the five motif types analyzed, di-nucleotides produced distinctly the highest percentage of polymorphism (82.3%) followed by tetra- (51.1%), hexa- (45.7%), and tri-nucleotides (42.9%) with approximately similar rates, while penta-nucleotides showed the lowest polymorphism rate (24.7%). Among the polymorphic 307 SSR loci, di-nucleotides had the highest genetic diversity values, followed by hexa- and tetra nucleotide motifs, while the penta-nucleotide had the lowest diversity values. For example, Na, Ne, Ho, He and PIC values were 3.89, 2.64, 0.46, 0.55 and 0.50 in dinucleotide SSRs, while they were 2.41, 1.75, 0.26, 0.39 and 0.33 in the penta-nucleotide SSRs, respectively (Table 1).

Of the 208 polymorphic di-nucleotides, 68 (32.6%) had Na = 5–9, and 41 (19.7%) amplified Na = 4 alleles, while 43 (20.6%) and 56 (26.9%) of them amplified Na = 2 and Na = 3 alleles, respectively. Only four SSR loci in tri-nucleotides had 5–6 alleles, while the rest of them had either two or three alleles. Similar results were also