Scientia Horticulturae 210 (2016) 277–284

Contents lists available at ScienceDirect

Scientia Horticulturae

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

Preliminary studies of genom-wide association mapping for some selected morphological characters of watermelons

Metin Yagcioglu[a], Osman Gulsen[a,*], Halit Yetisir[a], Ilknur Solmaz[b], Nebahat Sari[b]

[a] Department of Horticulture, Erciyes University, Melikgazi, 38039, Kayseri, Turkey
[b] Department of Horticulture, Çukurova University, Balcali, 01330, Adana, Turkey

ARTICLE INFO

Article history:
Received 27 February 2016
Received in revised form 30 July 2016
Accepted 2 August 2016
Available online 6 August 2016

Keywords:
Cucurbitaceae
Citrullus lanatus
Association mapping
Linkage disequilibrium
Bayesian analysis

ABSTRACT

Developing more efficient tools for watermelon (*Citrullus lanatus* (Thunb.) Matsum. & Nakai) breeding programs may greatly contribute to genetic improvement. The aims of this study were to investigate genome-wide association analysis (GWAS), and to develop regression models for some important watermelon traits such as yield, fruit weight, flesh color, total soluble solid, seed coat color, seed length and flowering time using 96 lines selected from 258 pure lines. An initial analysis was conducted to reduce sample size to 96 in order to avoid identical, highly similar and distant relatives causing low level of polymorphism and rare alleles. These 96 lines were morphologically and genetically characterized for several morphological characters under the eastern Mediterranean conditions, Adana, Turkey and PCR-based molecular markers. Associated markers were estimated by TASSEL program based on general linear model and a regression model because each trait was established through backward regression procedure nested in the SAS program. Levels of variation explained by regression models were 41% with 6 markers for flesh color, 22% with 2 markers for fruit weight, 48% with 5 markers for total soluble solids, 21% with 6 markers for seed size, 9% with 4 markers for fruit flesh firmness, 39% with 6 markers for 50% flowering time. This first report of GWAS for the selected characteristics of watermelon indicated a great potential to increase the efficiency of breeding programs.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Watermelon (*Citrullus lanatus* (Thunb.) Matsum. and Nakai) belongs to the genus *Citrullus* Schrad. Ex Eckl. et Zeyh in *Cucurbitaceae* family (Whitaker and Bemis, 1976). The genus comprises of four known diploid (n = 11) species (Dane and Liu, 2007). Of these, *Citrullus lanatus*, the most commercially important species, has two botanical varieties: *C. lanatus* var. *lanatus* Schrad. Ex Eckl. et Zeyh and var. *citroides* (L.H. Bailey). The other members of the genus *Citrullus* are *C. colocynthis* (L.) Schrad, *C. ecirrhosus* and *C. rehmii* De Winter.

Watermelon has many characters of interest for consumers and producers. Most of them are quantitatively controlled, and the genetic mechanisms of some of these traits were reviewed by Wehner (2008). The molecular tools helping in understanding the genetic mechanism of these traits might be very critical for increasing the efficiency of breeding programs. Trait loci affecting the phenotype can be detected by using linkage analysis and linkage disequilibrium mapping or a combination of both (Pikkuhookana and Sillanpaa, 2014). Both procedures use crossovers or recombinations between two loci. However, they differ in types of populations and power of detecting the number of alleles in a given population. Linkage mapping employs populations derived from controlled hybridizations such as F1, F2, backcross and recombinant inbred lines whereas linkage disequilibrium (LD) or association mapping (AM) uses populations with broad genetic background such as germplasm collections, elite lines and synthetic populations. Linkage mapping is simple and effective, and can be handled by a single software (i.e. Mapmaker, MapManager, JOINMAP). Association mapping, on the other hand, requires sophisticated statistical procedure to avoid spurious associations dealing with structured populations (Ball, 2007). More importantly, linkage mapping detects alleles only present in two parents of a specific mapping population while AM finds all associating alleles with trait interest in a given population.

Many genetic linkage maps have been reported for several species of *Cucurbitaceae* family such as melon (Dogimont et al., 2000); cucumber (Park et al., 2000); and watermelon (Hawkins et al., 2001; Hashizume et al., 2003; Levi et al., 2006). Current status of mapping efforts on QTLs associating with several important

* Corresponding author.
E-mail address: ogulsen@erciyes.edu.tr (O. Gulsen).

http://dx.doi.org/10.1016/j.scienta.2016.08.001
0304-4238/© 2016 Elsevier B.V. All rights reserved.

fruit traits such as Brix, rind thickness, fruit length, seed oil percentage and 100 seed weight by using F2 and recombinant inbred line (RIL) populations were summarized by Reddy et al. (2014). Several studies on AM were reported for several major crops such as wheat (*Triticum aestivum* L.) (Tadesse et al., 2014); perennial ryegrass (*Lolium perenne* L.) (Aleliunas et al., 2015); pearl millet (*Pennisetum glaucum* (L.) R.Br.) (Kannan et al., 2014) and rapeseed (*Brassica napus* L.) (Cai et al., 2014). Linkage disequilibrium between loci in watermelons was previously reported (Ocal et al., 2014; Reddy et al., 2014). To date, however, there has been one report of marker-trait association through AM in the *Cucurbitaceae* family. Reddy et al. (2015) indicated utilization of AM based on 96 accessions sampled throughout the world and mapped four quantitative trait loci (QTL) for fruit length, width and weight. This approach has advantages over that derived from only two parents and may detect all alleles available in a given germplasm. Aims of this study were to make survey for GWAS and develop regression models for some selected characters of watermelon.

2. Materials and methods

2.1. Plant materials and DNA isolation

A total of 258 accessions used in this study were described by Ocal et al. (2014). The samples mainly included selfed (4–6 times) lines of the local genotypes. Seeds of each accession were sown in plastic multi-pots (4 × 4 × 4 cm) filled with peat:perlite (2:1) for germination and bulk of fresh true leaves from each accession were harvested into freezer bags, and kept at −20 °C until DNA isolation. Total DNA was extracted from 30 mg tissue by using a modified CTAB DNA extraction procedure (Gulsen et al., 2010). DNA pellets were diluted with 300 μl of TE (10 mM Tris, 0.1 mM EDTA, pH 7.4). For PCR reactions, 10 ng/μl DNA templates were made using double-distilled water and checked on%2 agarose gels.

2.2. Molecular marker analysis

Sequence related amplified polymorphism (SRAP), random amplified polymorphic DNA (RAPD) and inter-simple-sequence repeats (ISSR) analyses were performed according to standard protocols. Two-hundred and eigth SRAP, 40 RAPD and 4 ISSR primers were tested for producing clear, polymorphic and repeatable bands. Selected primers were applied to the 258 lines. Each of 15 μl PCR components consisted of 0.66 mM of each of primers, 200 μM of each dNTPs, 1.5 μl of 10× PCR buffer, 2–2.5 mM of $MgCl_2$, ddH_2O, one unit of Taq polymerase and 20 ng of template DNA. PCR products of three marker systems were separated on 2–3% agarose gel at 110 V for 3–6 h and visualized under UV light. Peroxidase gene based markers were kindly provided by Ocal et al. (2014).

2.3. Selection of subpopulation for association analysis

The 258 accessions were subjected to initial screening to select individuals suitable for association mapping that requires non-identical, unique, but with low ratio of rare alleles. In order to achieve this, all accessions were analyzed for the similarity. Each marker was visually scored as present (1) or absent (0) and data were analyzed with the Numerical Taxonomy Multivariate Analysis System (NTSYS-pc) version 2.1 software package (Exeter Software, Setauket, New York, USA) (Rohlf, 2000). For estimating relationships among the accessions, all polymorphic markers were used to produce simple matching coefficients. Then the similarity matrix was used to construct a dendrogram using the unweighted pair group method arithmetic average (UPGMA) to determine genetic relationships among the accessions studied. Finally, the number of plant materials was reduced to 96 by eliminating highly similar and distinct ones that likely cause low level of polymorphism or rare alleles.

2.4. Morphological characterization

Morphological characterization of the 96 accessions selected in previous step was performed in the conditions of Adana. Experimental and treatment designs were randomized block and unstructured design, respectively. Seeds were sown to 2:1 in perlite and when seedlings had come to the level with 2 true leaves, they have been planted under plastic tunnels with 2.0 m × 0.5 m spacing and 7 plants in each replication. The morphological traits investigated are as follows: fruit weight by averaging at least 3 fruits for each line, flesh color (orange, dark pink, yellow, light pink, pink, red, white, light red), fruit flesh firmness by averaging three fruits measured in 3 different points for fruit with 11 mm diameter penetrometer), the percentage of total soluble solid (TSS) after one slice from each three fruits in every lines have been squeezed and filtered, 3–5 drops were read in hand refractometer, the first flowering (beginning from sowing to that at least one female flower occurs in 50% of the plants), flowering time (beginning from sowing to when 50% of flowers in the plant were female), seed coat color (spotted black, black, cream, cream with black hilum, brown, milky brown, creamy pink, cream black edge, spotted cream, reddish brown, dark brown) and seed length (mm) (measured with digital caliper), rind thickness, fruit diameter and height, presence of lines in fruit, fruit line width, ovary diameter and height measured by ruler or caliper. Ovary hairiness and existence of lines were visually detected. Distributions of phenotypes were drawn with Microsoft Excel version 2007.

2.5. Data analysis

The UPGMA dendrogram was produced as described above. Principal component analysis (PCA) was used to estimate efficiency of selection of subpopulation in the previous step and easily visualize dispersion among the individuals on 2-D plot. The EIGEN module was used to calculate Eigen values and two-dimensional plots based on the variance-covariance matrix calculated between each two pairs of the 96 watermelons. Then PROJ module was used to construct 2-D graph of 258 accessions.

Population structure was analyzed using a model-based approach, Bayesian method by STRUCTURE software version 2.3.2 (Pritchard et al., 2000). Output of this analysis was used as covariate to avoid spurious linkages. Since this software was developed for codominant markers such as simple sequence repeat (SSR), the data file was modified based on suggestions by the manual. We designated if the band was scored as present, and we used (1,−9), unlike (2,−9), where −9 is the value used for missing data. Model-based cluster analysis was used to test whether $K = 1–10$, where K is the number of sub-populations. Several statistical procedures to allow a user to decide what value of K best fits the data were discussed by Barkley et al. (2006). In addition to these, there is informal pointer to detect the best *K*, in which values of log *Pr* (*X*/*K*) reach more or less plateaus after major decrease. In this sort of situation, where several values of *K*s give similar estimates of log *Pr* (*X*/*K*), it seems that the smallest is often correct according to the manual. We used this approach to estimate *K*. The authors state that while it may not be possible to know the true value of *K*, one should try to pick the smallest value of *K* that captures the major structure of the data (Pritchard et al., 2000). Admixture and independent allele frequencies models were used. For each population (*K*), 5000 iteration and 5000 burn-in period options were used. For each number of *K* from 1 to 10, five independent calculations were performed, and likelihood values obtained from these calculations were averaged for each *K*. Finally a graph was drawn by using Microsoft Office Excel

Download English Version:

https://daneshyari.com/en/article/4565939

Download Persian Version:

https://daneshyari.com/article/4565939

Daneshyari.com