Scientia Horticulturae 176 (2014) 151–161

Contents lists available at ScienceDirect

Scientia Horticulturae

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

Genetic diversity, population structure and linkage disequilibrium among watermelons based on peroxidase gene markers

Necip Ocal[a], Mikail Akbulut[a], Osman Gulsen[b,*], Halit Yetisir[b], Ilknur Solmaz[c], Nebahat Sari[c]

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

ARTICLE INFO

Article history:
Received 21 February 2014
Received in revised form 24 June 2014
Accepted 1 July 2014
Available online 26 July 2014

Keywords:
Watermelon
AMOVA
Linkage disequilibrium
Turkish watermelons
Diversity
Population structure

ABSTRACT

Plant peroxidases are a family of stress-related proteins possessing highly conserved domains, which previously allowed to develop peroxidase gene polymorphism (POGP) markers. Diversity, population structure and linkage disequilibrium (LD) based on POGP markers among 256 watermelons [*Citrullus lanatus* (Thunb.) Matsum. and Nakai] genotypes along with two outgroups were evaluated by using 14 peroxidase gene-specific primers and several statistical procedures. Ratio of polymorphism of POGP markers was high (98%). The unweighted pair-group method arithmetic average (UPGMA), neighbor-joining (NJ) and principle component analysis (PCA) indicated low level of genetic diversity and few geography-based clustering among the watermelon accessions. Substructuring Bayesian analysis revealed that there were three subpopulations among 258 watermelons. Analysis of molecular variance (AMOVA) indicated that 54% of the molecular variation resided within subpopulations, 39% within geographic origins and 7% among subpopulations. Linkage disequilibrium (LD) analysis indicated presence of LD among POGP loci. Overall, this study concluded that LD existed despite that it was variable among subpopulations, association studies were possible among watermelons, and substructuring analysis was efficient in analyzing population structure among watermelons. This study also indicated that the POGP markers would be useful for clarifying genetic parameters such as population structure and relationships among watermelons.

1. Introduction

Members of the genus *Citrullus* are important for feeding humans and animals worldwide. It has four species of watermelons (Whitaker and Davis, 1962; Jeffrey, 1975; Levi et al., 2001a). Within this genus, *Citrullus lanatus* is the most common and includes two forms: most commercially cultivated *C. lanatus* var. *lanatus* and *C. lanatus* var. *citroides* [(L. H. Bailey) Mansf.], also known as citron. The other three species are *Citrullus colocynthis* (L.) Schrad., *Citrullus eccirrhosus* Cogn. and *Citrullus rehmii* De Winter (De Winter, 1990). *Praecitrullus fistulosus* (Stocks) Pangalo, a related species, is found in India and Pakistan. In spite of common morphological characters it slightly differ in chromosome number ($2n = 2\times = 24$) (Robinson and Decker-Walters, 1997). Perennial *C. colocynthis* is grown in North Africa, southwest Asia and Mediterranean basin. Perennial *C. eccirrhosus* (Meeuse, 1962) and annual *C. rehmii* (De Winter, 1990) are naturally found in Namibia Desert (Levi et al., 2005). Africa is the center of origin for all *Citrullus* species and has considerable level of diversity. The other diversity center of watermelons is Far East. Being in the middle of these two diversity centers, Mediterranean Basin and the Middle East may have potential in hosting diverse range of *Citrullus* species (Robinson and Decker-Walters, 1997; Wehner, 2008).

Turkey is the second leading producer of watermelon after China, with about 4.044 million tons of production (FAO, 2012) and almost all regions of Turkey from Mediterranean to Black Sea region and from Hakkari bordering Iran to Edirne neighboring Bulgaria and Greece produce watermelon. These regions have diverse environmental conditions such as arid and cooler temperatures in central Anatolia and very humid in Black Sea region. As a result, considerable number of local landraces have emerged which may have important in breeding programs (Huh et al., 2008; Solmaz and Sari, 2009). Genetic studies on watermelons representing Anatolia and the Middle East are scarce.

* Corresponding author at: Erciyes University, Department of Horticulture, Melikgazi 38039, Kayseri, Turkey. Tel.: +90 352 207 6666; fax: +90 352 437 6902.
E-mail address: o_gulsen@yahoo.com (O. Gulsen).

http://dx.doi.org/10.1016/j.scienta.2014.07.001
0304-4238/

Various techniques indicated low level of diversity within *C. lanatus*, but inter-species variation within this genus: EST-SSR (Hwang et al., 2011), RAPD and (Solmaz et al., 2010), SRAP (Solmaz, 2010; Uluturk et al., 2011). Probably this was due to relative recent evolution of watermelon or bottleneck in their history. Estimating genetic diversity is useful for organizing genetic resources, identifying heterotic groups or duplicates. It is also useful for discovering the function of genes through the mapping of qualitative and quantitative trait loci (QTL) in structured populations. For example, discovering genes underlying agronomically important characters based on natural populations requires deliberate statistical analyses. Without considering population structure among a set accessions, association analysis between marker and trait are likely to be negatively affected causing false positives in marker-trait association analyses. However population structure is common in plant species and that of watermelon accessions are largely unknown (Mujaju et al., 2012).

Plant peroxidases play important roles in many stress-related interactions such as pathogen infection, insect and salt tolerance, auxin degradation, cell wall lignification, tissue suberization and plant senescence (Hinman and Lang, 1965; Espelie et al., 1986; Amaya et al., 1999; Chittoor et al., 1999; Passardi et al., 2005; Gulsen et al., 2010a) and belong to a multigene family. They contain heme group that can oxidize compounds in the presence of peroxide (H_2O_2) or oxygen (O_2) (Duroux and Welinder, 2003). Therefore, they possess three highly conserved motifs, the distal heme-binding domain, the central domain with unknown function and the proximal heme binding domain (Hiraga et al., 2001). The level of sequence identity among peroxidase genes was the second lowest among 25 multigene families, possibly indicating above average evolutionary force on this gene family (Zhang et al., 2001). This feature of peroxidases allowed designation of POGP primers to study peroxidase-based diversity of buffalo grasses (*Buchloe dactyloides*), wheat (*Triticum* spp.) and apple (*Malus communis*) (Gulsen et al., 2007, 2010a) and to sequence some of peroxidase in these plant species (http://www.ncbi.nlm.nih.gov). All indicated considerable level of diversity and significant associations between the POGP markers and ploidy levels, insect tolerance and species (Gulsen et al., 2007; Ceylan, 2010). Watermelons are exposed to various biotic or abiotic stresses that could enhance diversity in peroxidase gene family. Therefore, peroxidase markers may be useful in estimating genetic parameters interest and likely reflect environmental variations. Peroxidase genes are likely to be monophyletic origin due to common features and distributed in clusters or single over the genome as in the *Arabidopsis thaliana* (http://www.ncbi.nlm.nih.gov/). Gulsen et al. (2010b) mapped few of the POGP markers and found that few of them were closely linked while others were dispersed among the *Citrus* chromosomes. There is no report on organization of peroxidase genes among watermelons. The POGP markers are PCR-based and multilocus. They produce dominant and co-dominant markers (Gulsen et al., 2010b). Thus the POGP markers may provide alternative approach to elucidate genetic structure of watermelons.

Linkage disequilibrium (LD) was reviewed by Oraguzie et al. (2007) and it is a population genetics term used to describe non-random association of alleles at different loci in a population. This happens due to physical linkage and unequal recombination events between loci. LD is very useful in estimating association between marker locus and agronomically important traits in plants and LD among watermelons is unknown. The objectives of this study were to estimate: (1) peroxidase gene-based genetic diversity among watermelons, (2) distribution of molecular variation among and within their subpopulations, and among geographic origins, (3) population structure among watermelons, and (4) linkage disequilibrium among POGP markers among watermelons.

2. Materials and methods

2.1. Plant materials

A total of 258 accessions consisted of 249 cultivated watermelons (*C. lanatus* var. *lanatus*), seven wild forms (*C. lanatus* var. *citroides* (Bailey) Mansf.) known as citron or preserving melon and two accessions of a related species (*P. fistulosus* (Stocks) Pangalo) as out-group were used (electronic supplementary Table A.1). Four commonly cultivated cultivars (36, Sugar Baby; 39, Crimson Sweet; 233, Calhoun Gray; 235, Charleston Gray) were also used as controls. Of the *C. lanatus* accessions, 232 were collected from Turkey. All accessions were selfed 4–6 times and are maintained in the cold room at the University of Cukurova, Adana, Turkey. Seven seeds of each accession were sown in May 2010 in peat-filled plastic multipots (4 cm × 4 cm × 4 cm) for germination and bulk of fresh true leaves from each accession were harvested into freezer bags, and kept at −20 °C until DNA isolation.

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2014.07.001.

2.2. DNA isolation and marker production

Total DNA was extracted from 30 mg fresh leaf tissue according to a modified CTAB DNA extraction procedure (Gulsen et al., 2010a). DNA pellets were diluted with 300 μL of TE (10 mM Tris, 0.1 mM EDTA, pH 7.4), and their concentrations were measured with a semi-automated microplate reader, PowerWave HT (Bio-Tek Instruments, Inc., Winooski, VT, USA), and 10 ng/μL DNA templates for PCR amplifications were made using double distilled water.

Primers were previously designed based on 34 rice peroxidase cDNA sequences and their sequence information was provided by Gulsen et al. (2007). Oligonucleotides were synthesized by Iontek Inc., Merter, Istanbul (www.iontek.com.tr). Fourteen peroxidase primer pairs were used to amplify peroxidase genes of watermelons. Each of 15 μL reaction consisted of 450 μM of the primer pairs, 200 μM of each of dNTPs, 1.5 μL of 10× PCR buffer, 2.5 mM of $MgCl_2$, 0.8 μg μL^{-1} bovine serum albumin and 1 unit of Taq polymerase and 25 ng of template DNA. Cycling parameters included one cycle of 3 min at 94 °C, 34 cycles of 1 min at 94 °C, 1 min at 40–53 °C, 2 min at 72 °C, and for final extension one cycle 5 min at 72 °C. PCR products were separated on 2.5% agarose gel at 110 volt for 4 or 5 h. Bands of amplified DNA were visualized with ethidium bromide staining.

2.3. Data analyses

For studying POGP based genetic analysis in watermelons, each band was scored as present (1) or absent (0). Frequency of rare present alleles (<5 and <10% level) for each marker system was calculated as: $Q = a/b$, where a is the number of rare (5 or 10%) present alleles and b is the total number of bands scored for a specific locus. The number of bands was detected based on the observed total number of bands in all genotypes.

For unweighted pair group method arithmetic average (UPGMA) analysis, binary data were analyzed with the Numerical Taxonomy Multivariate Analysis System (NTSYS-pc) version 2.1 software package (Exeter Software, Setauket, NY, USA) (Rohlf, 1993). Then, a similarity matrix was constructed based on Dice' similarity coefficient (Dice, 1945), which considers only one to one matches between two taxa for similarity. The similarity matrix was used to construct the UPGMA dendrogram to determine genetic relationships among the germplasm studied. To provide a "goodness-of-fit" test for the GS matrix to cluster analysis, first, COPH module was used to transform the tree matrix to a matrix of ultrametric