

A novel genetic map of pomegranate based on transcript markers enriched with QTLs for fruit quality traits

Rotem Harel-Beja^{1,3} · Amir Sherman² · Mor Rubinstein² · Ravit Eshed² · Irit Bar-Ya'akov¹ · Taly Trainin¹ · Ron Ophir² · Doron Holland¹

Received: 18 May 2015 / Revised: 2 September 2015 / Accepted: 18 September 2015 / Published online: 1 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract A novel genetic map of pomegranate (*Punica granatum* L.) enriched with quantitative trait loci (QTLs) for seven traits was constructed using an F2 population. The population was generated from a cross between two varieties of *P. granatum*: “Nana” and “Black.” Phenotyping of 76 F2 plants was conducted over two seasons. The map includes 1092 SNP markers which were newly developed from de novo transcriptome assembly and a comparison of the sequences of the two varieties or accessions. The map covers 1141 centimorgans (cM) with an average of 1.17 cM between markers over 11 linkage groups. Twenty-five QTLs were identified for fruit traits and plant size. The map includes QTLs for total soluble solids (TSS), fruit weight and perimeter, seed hardness, aril color and weight, and plant height. In an effort to explore the potential of the Agricultural Research Organization (ARO) pomegranate germplasm collection to associate traits with gene markers, an association study was

conducted using a set of 346 SNPs described in an earlier study. Of these SNPs, 233 were mapped on the genetic map and found to be distributed in the different linkage groups (LGs). The associated traits were anchored to the map by these common markers. The F2 population described here and the corresponding genetic map provide a useful resource for further genomics and genetic studies of pomegranate, as well as a reservoir of markers for fruit improvement.

Keywords Genetic map · Pomegranate (*Punica granatum* L.) · QTLs · Fruit quality traits

Introduction

Pomegranate (*Punica granatum* L.) belongs to the Lythraceae family and is grown worldwide. According to Masoud et al. (2005), pomegranate is a diploid plant ($2n=16$). The pomegranate fruit has a pleasant taste and an attractive appearance, and many parts of the plant are used for commercial purposes. Traditional and modern medicine indicate that the pomegranate tree, its flowers, and its fruit have a high content of important metabolites with health-promoting properties. The agricultural value of the pomegranate tree and its products is constantly increasing as its various benefits are being explored by modern science. Numerous scientific studies (Reviewed in Seeram et al. (2006) and Holland and Bar-Ya'akov (2014)), with human patients, animals, and cell cultures suggest that pomegranate-derived substances have antioxidant, anti-carcinogenic, anti-inflammatory, antimicrobial, and estrogenic properties. The prominent bioactive phytochemicals present in pomegranate are various polyphenolic compounds, including ellagitannins and flavonoids.

Worldwide genetic and molecular studies of pomegranate have so far been done mainly on local collections of

Communicated by V. Decroocq

This article is part of the Topical Collection on *Complex Traits*

Electronic supplementary material The online version of this article (doi:10.1007/s11295-015-0936-0) contains supplementary material, which is available to authorized users.

✉ Doron Holland
vhhollan@volcani.agri.gov.il

¹ Department of Fruit Tree Sciences, Neve Ya'ar Research Center, Agricultural Research Organization, Ramat Yishai P.O.Box 1021, 30095, Israel

² Department of Fruit Tree Sciences, Volcani Center, Agricultural Research Organization, Bet Dagan, Israel

³ The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel

accessions (Currò et al. 2010; Hasnaoui et al. 2010, 2012; Pirseyedi et al. 2010; Basaki et al. 2011; Sarkhosh et al. 2011; Soriano et al. 2011; Jian et al. 2012; Noormohammadi et al. 2012; Zhang et al. 2012; Hajiahmadi et al. 2013; Singh et al. 2013; Ferrara et al. 2014; Orhan et al. 2014). These studies involved the usage of various DNA markers in order to distinguish between accessions within local collections and in an effort to associate genetic markers with traits. The main molecular markers that were used to screen pomegranates were simple sequence repeats (SSRs) (Currò et al. 2010; Hasnaoui et al. 2010, 2012; Pirseyedi et al. 2010; Basaki et al. 2011; Soriano et al. 2011; Jian et al. 2012; Noormohammadi et al. 2012; Ferrara et al. 2014). However, studies with other molecular markers such as random amplification of polymorphic DNA (RAPD) (Noormohammadi et al. 2012; Zhang et al. 2012; Singh et al. 2013; Orhan et al. 2014), amplified fragment length polymorphisms (AFLP) (Sarkhosh et al. 2011), or internal transcribed spacers (Noormohammadi et al. 2012; Hajiahmadi et al. 2013; Singh et al. 2013) were reported in the literature. In most cases, no overlap was found between the classification by genotype and by the phenotype.

Recently, a wide molecular genetics study, which includes de novo establishment of the pomegranate transcriptome and the identification of several thousand SNP markers, was reported (Ophir et al. 2014). A subset of 346 SNPs was used to investigate the genetic structure of the Agricultural Research Organization (ARO) germplasm collection. The ARO collection includes more than 100 accessions, local accessions, and accessions introduced from foreign countries such as, India, China, Central Asia, the USA, Spain, Turkey, and the Mediterranean region (Ophir et al. 2014). The selected SNPs are highly diverse in the ARO pomegranate germplasm (half of the SNPs had PIC ≥ 0.43). Moreover, this set of SNPs was highly polymorphic, with only 10 % loci with minor allele frequencies of (MAF < 0.05). These SNPs were successfully used to classify the ARO pomegranate collection by hierarchical clustering to three statistically significant groups (bootstrap value > 90 %) (Online Resource Table 1). The STRUCTURE program was used suggesting that there are two major subpopulations.

Few genetic studies on pomegranate associated markers with traits (Basaki et al. 2011, 2013; Singh et al. 2015). Basaki et al. (2011, 2013) screened seven SSRs on 202 pomegranate accessions representing 22 provinces of Iran. Despite the limited number of markers, the associated SSRs explained up to 29.6 % of the variation for 20 individual traits, including fruit and flower traits, leaf size, and shoot production. The Indian pomegranate collection was genotyped with 44 microsatellites (Singh et al. 2015); three traits were found associated to three different markers.

Genetic maps based on segregating populations are powerful tools for linking genes and characters. A saturated linkage

map should discover possible links between genotype and phenotype and also identify efficient markers for use in marker-assisted breeding. In melon, markers were mapped close to the trait of fruit acidity (Danin-Poleg et al. 2002; Harel-Beja et al. 2010). Another example is the genetic map of *Artemisia annua* L. that identified loci affecting yield that co-located with quantitative trait loci (QTLs) for artemisinin concentration, thus representing targets for a breeding program (Graham et al. 2010). In trees such as apple, QTL analysis and the candidate gene approach led to two haplotypes that were characterized and associated with fruit flesh browning (Di Guardo et al. 2013). In apricot and peach, QTL for chilling requirements was detected with a major locus at LG1 (Olukolu et al. 2009; Fan et al. 2010).

In the present study, an enriched pomegranate map was constructed based on an F2 population developed from a cross between two varieties, “Nana” and “Black.” These varieties differ in fruit quality traits, including aril characters such as sweetness, acidity, and color, seed hardness, external peel characters, and fruit size. In addition, the Nana parental variety is a dwarf pomegranate while the Black variety has a normal pomegranate tree size. Thus, the F2 population generated from a cross between Nana and Black is an excellent tool for mapping fruit quality traits along with tree growth characteristics such as tree size. The pomegranate genetic map was established with ~ 1100 SNP markers that were derived from transcripts. QTLs for total soluble solids (TSS), fruit size, aril color, seed hardness, and plant size were mapped. Finally, association mapping was done for fruit traits in the ARO collection.

Materials and methods

Plant material

A segregating F2 population ($n=76$) was constructed from a cross between Black (P.G.127-28, according to the nomenclature of the ARO collection at Newe Ya’ar cultivar and Nana seedling selection (P.G.232-243, *P. granatum* var. Nana). Nana seedlings are characterized as dwarf, bushy plants that have very small and sour fruits with hard seeds and a green to red peel. Black is a domesticated cultivar characterized as a deciduous normal-sized tree with a very distinct deep-purple peel. The Black accession has a fruit of medium size, with sweet taste and soft seeds. The parents are part of the ARO collection (Holland et al. 2009). The F1 plant, which was the progenitor of the F2 population, was planted in 2008 and was self-pollinated at 2009 (Nana pomegranate flower and fruit few months after germination). The F2 population was planted in 2010 at the Newe Ya’ar Research Center in northern Israel.

The ARO’s pomegranate germplasm collection is located at the Newe Ya’ar Research Center (<http://igb.agri.gov.il/>

[main/index.pl?page=22](#)). It consists of more than 150 accessions originating from different geographical locations all over the world and in particular from Israel (Holland et al. 2009). For the association study, 94 accessions were analyzed (Online Resource Table 1).

Phenotype evaluation of a segregating F2 population

The two parents, the F1 plant and the F2 population, were grown in the orchard at Newe Ya'ar, Israel. Three mature fruits from each plant of the F2 population were harvested, from different parts of the tree in two successive fruiting seasons (2 and 3 years after planting). Maturity was determined by the combination of aril and peel color and by tasting for astringency and sourness. Each fruit was examined for the

following traits: fruit perimeter, fruit and aril weight, TSS, peel and aril color, aril taste, and seed hardness. Peel and aril color, aril taste, and seed hardness were evaluated by visual screening and by tasting the fruit. These traits were graded from 1 to 5, where 1 is the least favored and 5 is the most favored. Fruit perimeter was measured by measuring tape; fruit weight was measured by scale (Merav 2000, Shekel Scales Ltd. Israel) and aril weight was measured by a digital scale (BBA-1200, M.R.C. Israel). TSS was measured by a hand refractometer (ATC-1 Atago Japan). Plant height was measured in the second, third, and fifth year after planting by measuring tape. Plant growth was calculated by subtracting the height in the second year from the height in the third year.

Fruits from the parental lines were harvested and evaluated like the fruits of the segregating population in the 2011 and

Table 1 Phenotypic data of the parental accessions, the F1 hybrid parent and the F2 population, for seven traits

Trait	Years of trial	Black, mean±SD	Nana, mean±SD	F1, mean±SD	F2 population, mean±SD	F2 population, max.	F2 population, min.
Fruit perimeter (cm)	2011	33.17±0.74	29.77±4.24	29.47±1.17	NA	NA	NA
	2012	NA	NA	20.23±0.50	20.64±2.94	28.40	10.50
	2013	NA	NA	NA	22.414±3.14	28.00	11.74
	2014	24.83±1.63	14.22±1.25	20.07±1.76	NA	NA	NA
Fruit weight (g)	2011	194.37±31.53	69.47±19.07	123.9±12.26	NA	NA	NA
	2012	NA	NA	137.67±9.05	147.04±52.3	336.37	50.98
	2013	NA	NA	NA	181.58±64.64	338.40	26.36
	2014	237.23±41.1	47.3±12.23	135.1±36.12	NA	NA	NA
Aril weight (g)	2011	0.36±0.03	0.23±0.01	0.31±0.01	NA	NA	NA
	2012	NA	NA	0.37±0.02	0.23±0.06	0.35	0.07
	2013	NA	NA	NA	0.25±0.05	0.38	0.12
	2014	0.37±0.09	0.19±0.03	0.39±0.03	NA	NA	NA
Peel color (1-5)	2011	5±0	1.43±0.12	4±0	NA	NA	NA
	2012	NA	NA	3.5±0	3.63±1.19	5.00	1.00
	2013	NA	NA	NA	4.24±0.94	5.00	1.10
	2014	5.0±0	1.25±0.05	4.53±0.12	NA	NA	NA
Aril color (1-5)	2011	4.87±0.06	2.73±0.46	4±0	NA	NA	NA
	2012	NA	NA	3±0.87	2.67±0.86	4.50	1.00
	2013	NA	NA	NA	2.42±0.9	4.20	1.00
	2014	4.33±0.12	3.17±0.29	4.17±0.15	NA	NA	NA
Seed hardness (1-5)	2011	4.5±0	1±0	1±0	NA	NA	NA
	2012	NA	NA	1±0	1.97±1.24	5.00	0.37
	2013	NA	NA	NA	1.83±1.3	4.90	1.00
	2014	4.3±0	1±0	1±0	NA	NA	NA
TSS	2011	18±0	15.07±0.12	18.13±0.92	NA	NA	NA
	2012	NA	NA	16.13±0.50	14.43±1.58	17.73	9.13
	2013	NA	NA	NA	15.78±1.31	18.27	9.67
	2014	17.67±0.31	14.33±0.82	17.33±0.42	NA	NA	NA

The traits were measured and evaluated over two seasons for the parents (2011 and 2014) and for the F2 population (2012 and 2013); the F1 hybrid parent was evaluated over 3 years (2011, 2012, and 2014). The data is summarized as the means, standard deviation (SD), and maximum and minimum values of the segregating population

2014 fruit seasons; in addition, they were characterized as part of the ARO collection for at least 5 years. The F1 plant was examined during 2011, 2012, and 2014. The heights of mature plants of the maternal parent and of F1 plant were measured by measuring tape. The height of mature plants of the parental variety was evaluated.

Phenotypic evaluation of the pomegranate variety collection

The pomegranate collection was established in 1978, and since then, new accessions have been added to the orchard. Once each of the accessions reached the fruiting stage, the tree and fruit were characterized for at least 5 years. Characterization included, among other traits, date of maturity (when the fruit reached edible quality), fruit size (weight and diameter), peel and aril colors (visual evaluation), taste (organoleptic description), and TSS (by refractometer (ATC-1 Atago, Japan)). Total acid content was calculated as percentage of citric acid by titrating the pomegranate juice with a solution of NaOH 0.1 N. Five mature fruits were harvested from two different trees of each accession, from different parts of each tree. The characterizations of the different traits were averaged over 5 years (Online resource Table 1).

DNA extraction

The DNA extraction protocol was based on Porebski et al. (1997) with few modifications. Young leaves from the parental lines, the F1 parent of the F2 population, and the population plants were used for DNA preparation, 0.5 g resuspended in a 6-ml extraction buffer. The chloroform–octanol solution was replaced with chloroform–isoamyl alcohol. DNA was precipitated with sodium acetate instead of sodium chloride.

Statistical analysis

Means, standard deviations, trait distribution, pairwise correlation, ANOVA and Welch ANOVA, Shapiro–Wilk W test, and Wilcoxon/Kruskal–Wallis test (rank sums) analyses were conducted with the JMP program, v. 7.0 (SAS Institute Inc., Cary, NC).

Marker analysis and map construction

Array and probe design

SNP markers were discovered by whole-transcriptome sequencing of two mRNA samples that were extracted from the accessions Nana and Black (Ophir et al. 2014) and from the P.G.160-61 and P.G.100-1 accessions (unpublished data) on 454-GS FLX Titanium. Total RNA was extracted from a pool of tissues including leaves, roots, flowers (petals, anthers,

ovaries), and fruit (stage 3 (Ben-Simhon et al. 2011)) of Nana and Black. The same pool of tissues, excluding the roots, was used for RNA extractions from the P.G.160-61 and P.G.100-1 accessions. De novo assembly was performed separately on pairs of accessions' sequencing with MIRA (Chevreux et al. 2004), and polymorphic loci of low quality ($Q < 30$; PHRED scale; (Ewing et al. 1998)) were filtered out. The repertoire of SNPs for array design was composed of three SNP subsets: homozygous but polymorphic between Nana and Black (6001), homozygous but polymorphic between P.G.160-61 and P.G.101-2 (4311), and heterozygous at least in one of the parental accessions, i.e., Nana or Black (4800). Altogether the array was composed of 15,494 SNPs. Four probes were designed for each SNP to fill an Agilent 60K SurePrint array. Two replicate probes were designed for each allele, one to detect the forward DNA strand and the other to detect the reverse DNA strand. The probes were T_m optimized, as described by Rubinstein et al. (2015).

Array hybridization and signal preprocessing

Labeling of DNA pools and hybridization were performed at the DNA array unit of the Weizmann Institute (Rehovot, Israel) following the Agilent CGH protocol for comparative genomic hybridization (<http://www.agilent.com>); hybridization temperature was modified to 55 °C. Emission intensities (signal values) were preprocessed in three steps: quantile normalization across arrays and dyes, removal of spatial slide defects, e.g., scratches, bubbles, etc. by the method described in Chai et al. (2010), and subtracting each probe by its median.

Genotype calling

An SNP genotype call was performed by k -means partition clustering. Differential signals (M =allele X–allele Y) were calculated for forward and reverse probes. All F2 hybridization signals per SNP were used as two vectors of forward- M and reverse- M , and a k -means clustering, with $k=3$, was performed (see an example in Liu et al. (2003)). The genotype call of a cluster was called by projecting the medoid of the clusters on the diagonal of the forward- M and reverse- M axes. The cluster with the maximal medoid was typed as XX, the cluster with the median medoid was typed as XY, and the cluster with minimal medoid was typed as YY. Only SNPs with three distinct clusters, i.e., clusters with a minimal Bhattacharyya distance of 0.5 were selected for downstream analysis. In addition, SNPs were selected only if the parent F1 plant was heterozygous.

Map construction and QTL analysis

Mapping was performed using JoinMap[®] 3.0 software (Van Ooijen and Voorrips 2001). Markers were grouped at a minimum LOD score of 4.0 and a recombination frequency value of 0.4. JoinMap[®] 3.0 software uses Kosambi mapping functions to translate recombination frequency into map distance.

Using the MapQTL[®] 5 software (Van Ooijen 2004), QTLs and their significance were calculated using interval mapping (IM), multiple QTL model (MQM), and permutation analysis. A QTL was determined significant when its LOD score was higher than the calculated threshold (1000 permutation at $p=0.05$).

Map association of traits within the germplasm collection

Association between traits and SNPs within the germplasm collection was analyzed using the TASSEL (version 3.0.164) program (Bradbury et al. 2007). The general linear model (GLM) function was used for association analysis. For each marker–trait combination, GLM finds the ordinary least squares solution. The p value was specified as less than 10^{-4} . Associations within the germplasm collection were excluded if one of the allele compositions was presented in more than 80 % of the accessions within the collection. Significant associations were analyzed by the JMP program, v. 7.0 (SAS Institute Inc., Cary, NC) using an ANOVA test if the trait was normally distributed (fruit weight and TSS); otherwise, they were analyzed by Wilcoxon/Kruskal–Wallis tests (rank sums). In addition, mixed linear model (MLM) was analyzed using

the TASSEL (version 3.0.164) program (Bradbury et al. 2007) that implement the kinship and structure analysis of the germplasm collection.

Results

Development and characterization of an F2 population

An F2 population, designated “Nana×Black,” was developed from a cross between Nana (P.G.232-243, *P. granatum* var. Nana) and Black (P.G.127-28) cultivars. The population was polymorphic in many characteristics (Fig. 1), including growth rate and fruit traits.

The population was genotyped using the SNP CHIP. All the SNPs were newly developed markers from the pomegranate transcriptome (Ophir et al. 2014). Markers of 1835 were selected as described in genotype calling in the “Materials and methods” section. Two hundreds and fifteen of the SNPs were not polymorphic.

Although the size of the F2 population is not high, 3102 recombination events occurred in the 76 F2 plants. The number of recombination events per plant ranged from 25 to 73, with a mean of 45.6.

The heterozygosity in the population was 47 %, calculated as the average of the percent of heterozygous markers for each plant of the population. The percent of heterozygosity was lower than 30 % in only one progeny that had 28 % heterozygosity. Six plants had over 60 % heterozygosity; the highest percentage was 69 %. The alleles of the

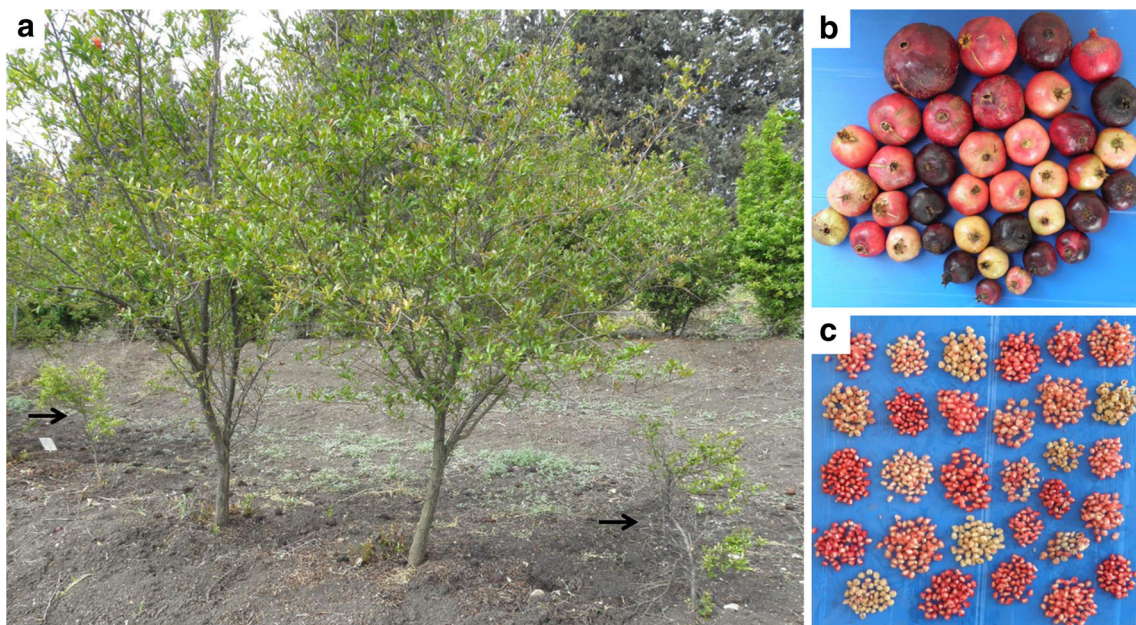


Fig. 1 Polymorphism within the Nana×Black F2 population. The segregating population is polymorphic regarding plant height (a), fruit size (b), and aril color and size (c). Arrows indicate two low trees in a row with high trees

maternal parent were 1.3 times more abundant in the F2 population than the alleles of the paternal parent. Five progenies had more than twice the maternal alleles than the paternal alleles.

Development of an enriched linkage map

Total of 1625 SNP markers were analyzed by the JoinMap® 3.0 program. The segregation of 298 markers was not Mendelian segregation as was tested by chi-square test ($\chi^2 > 10.4$, $df=2$) (Van Ooijen and Voorrips 2001). In addition, 235 SNPs were removed since they were identically segregated. Overall, a total of 1092 SNP markers were used to construct the linkage map that was named Nana×Black (Fig. 2 and Online Resource Table 1).

The Nana×Black F2 genetic map covers 960 centimorgans (cM) and is divided into seven major linkage groups (LGs), with an average of 1.17 cM per marker. The largest linkage group is LG1, which covers 155 cM (Online Resource Table 2). Interestingly, when 176 of the deviant distributed markers were included in the map calculations, four additional linkage groups were created; therefore, overall, the map covers 1141 cM. Two of the additional groups (LG8 and LG9) included 67 and 58 markers, respectively, and covered more than 60 cM each (Fig. 2). The additional deviant markers were located to LG2 (2), LG3 (5), LG6 (12), LG10 (27), and LG11 (5) (marked in Online Resource Table 3). The deviation from the Mendelian segregation was mainly due to a maternal homozygote genotype which was more frequent than the paternal genotype. In addition, there were 281 markers for which one of the parents was a heterozygote; they were spread all over the linkage groups.

Of the mapped SNPs, 233 were previously genotyped within the ARO collection (Ophir et al. 2014). The names of these markers are bolded and end with “_fl”. These markers were distributed all over the map (Fig. 2 and Online Resource Table 2 and 3).

Fruit traits

The F2 population was used to phenotypically characterize and map variation in fruit qualities such as size, taste, seed hardness, and color. The F2 plants as well as the F1 plant (parent of the mapping population) and its parental accessions were phenotyped during at least two growing seasons: the F1 in 2011, 2012, and 2014 and the F2 population in 2012 and 2013. The parent trees (Nana and Black) were characterized as part of the ARO collection for at least 5 years. In addition, they were characterized in 2011 and 2014. Mean and minimum and maximum values of the fruit traits for the parents and the F1 and the F2 population are presented in Table 1. Fruit perimeter, fruit

weight, and aril weight were distributed as normal distribution in 2013 (Fig. 3) according to the Shapiro–Wilk goodness-of-fit test. The other three traits, aril color, seed hardness, and TSS, were not normally distributed ($p \leq 0.01$) in 2013; however, in 2012, only seed hardness was not normally distributed (data not shown).

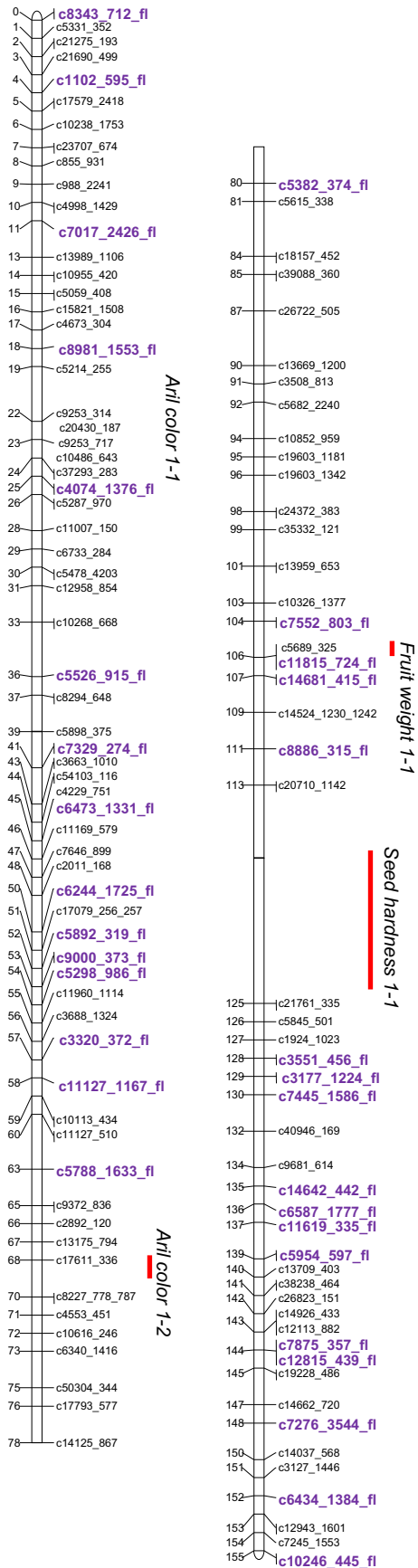
The correlation between fruit perimeter and fruit weight was high, 0.86 and 0.94 for 2012 and 2013, respectively (Table 2). In addition, aril weight was also highly correlated with fruit weight, 0.53 and 0.68 for 2012 and 2013, respectively. The other fruit traits were correlated at lower levels ($r \leq 0.56$).

The year effect was not significant according to Wilcoxon/Kruskal–Wallis Tests ($p \geq 0.01$) for three traits: aril weight, color, and seed hardness. However, the year effect was significant ($p \leq 0.01$) for three fruit traits: fruit perimeter, weight, and TSS, possibly since the trees were young. Young pomegranate trees tend to yield slightly different fruits than fully mature trees. Yield load can also affect these traits. The parental lines are mature trees; the year effect within the parents was not significant ($p \geq 0.01$) for the various traits, except of fruit perimeter. Therefore, it may be assumed that fruit weight and TSS were affected by the age of the tree over the F2 population. Nevertheless, it may indicate a genotype×environment interaction.

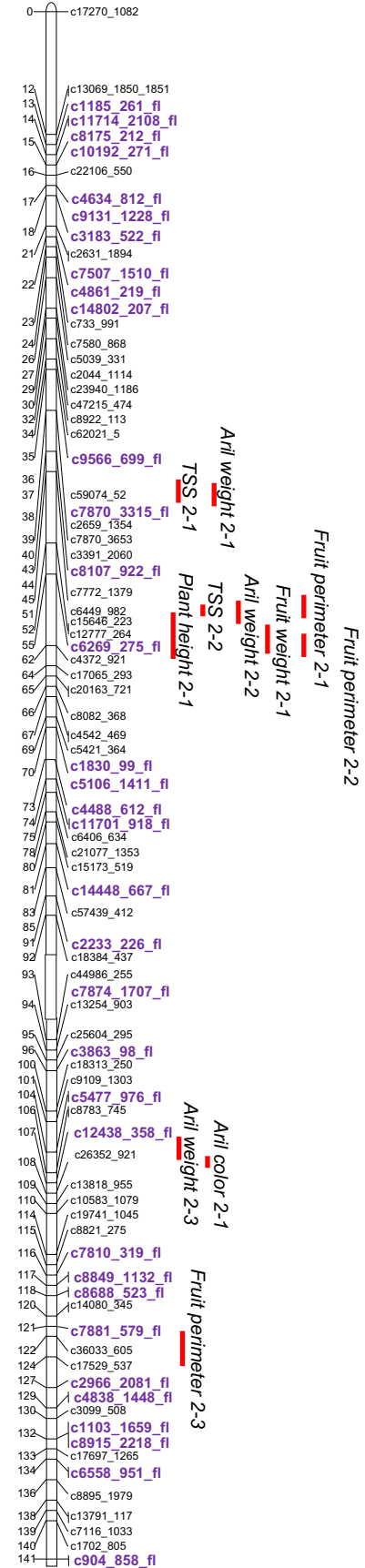
Transgressive segregation was detected for five fruit traits (Fig. 3): fruit and aril weight, fruit perimeter, aril color, and TSS. Regarding the fruit weight trait, 16 plants had heavier fruits than the fruits of the Black variety, and only two plants with fruits lighter than Nana, were present in the population. With respect to aril weight, on the other hand, nine plants had fruits with smaller arils than Nana and only one had heavier arils than Black. Twenty-one plants had fruits with larger perimeters than Black and only one that was smaller than the fruits of Nana. The aril color of both parents was evaluated with high values, Nana arils are red and Black arils are dark pink. On the other hand, most of the plants in the population (61) were evaluated with lower scores for their aril color as compared to Nana. The TSS levels of Black were high (about 17.7) and the TSS levels of Nana were low (about 14.3). Ten of the plants had much lower TSS levels and three had higher TSS levels than Black.

Fig. 2 Genetic map of pomegranate based on the Nana×Black F2 population including QTLs for fruit traits and plant height and data from map association of fruit quality traits in the ARO collection of varieties. LG numbers are marked above each group. Distances in centimorgans are written on the *left side* and marker names are on the *right side* of each linkage group. Markers that were genotyped within the variety collection are written in *purple* and end with *_fl*. Significant QTLs (by permutation test) analyzed by MQM analysis are represented by *red lines*. *Rectangles* and *blue arrows* indicate SNPs that are associated with fruit traits within the variety collection by MLM analysis

LG1



LG2



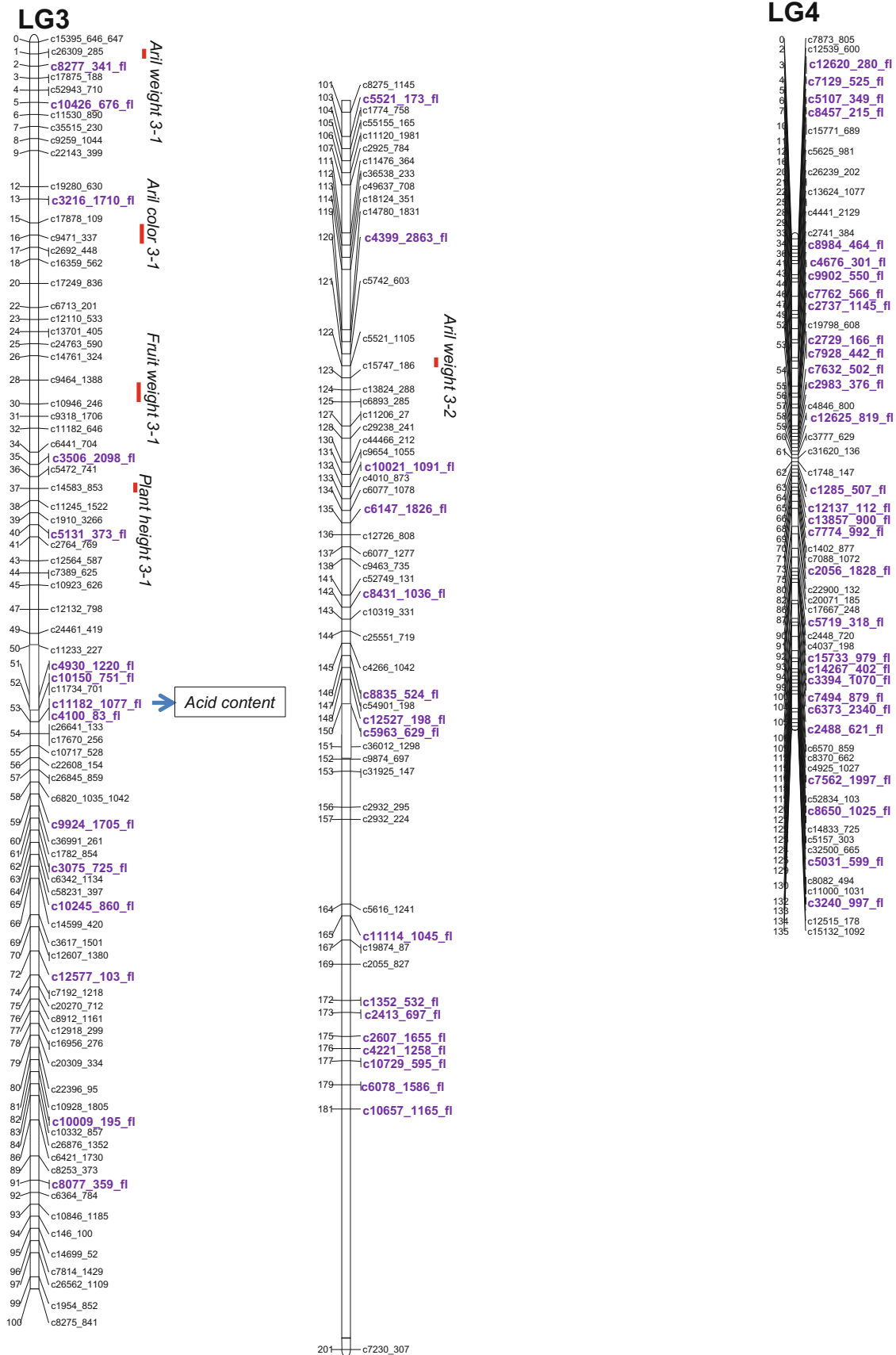


Fig. 2 (continued)

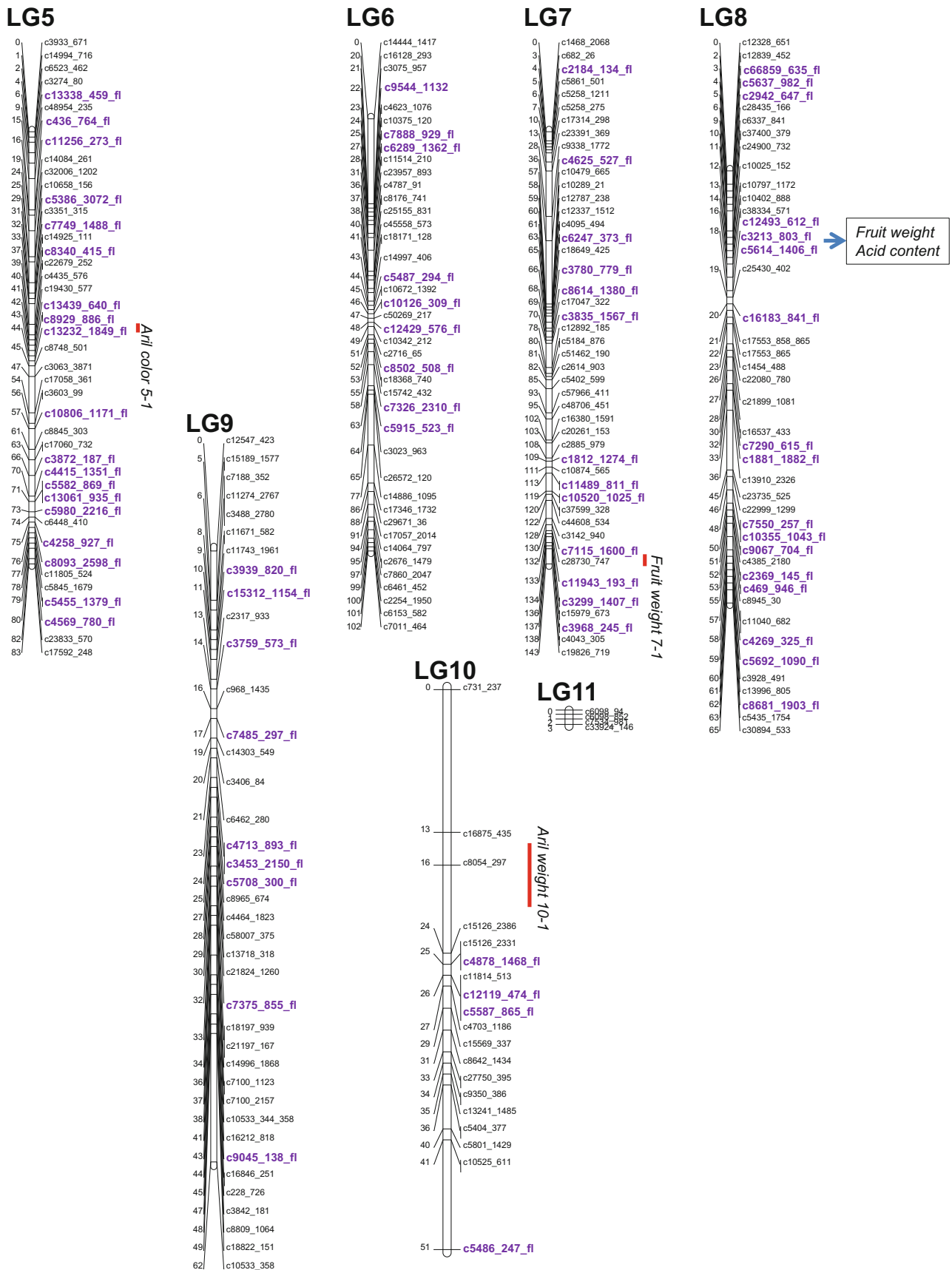


Fig. 2 (continued)

Table 2 Correlation between fruit traits within the segregating F2 population

	Fruit perimeter	Fruit weight	Aril weigh	Peel color	Aril color	Seed hardness
2012						
Fruit weight	0.86					
Aril weight	0.57	0.53				
Peel color	0.17	0.13	0.42			
Aril color	-0.30	-0.29	-0.02	-0.01		
Seed hardness	0.17	0.24	0.10	-0.10	-0.02	
TSS	0.25	0.22	0.56	0.43	0.32	-0.06
2013						
Fruit weight	0.94					
Aril weight	0.67	0.68				
Peel color	0.22	0.19	0.35			
Aril color	-0.20	-0.19	-0.22	-0.24		
Seed hardness	-0.03	0.02	-0.10	-0.30	0.11	
TSS	0.25	0.20	0.34	0.41	0.28	-0.17

Plant growth

The maternal parent of the segregating population was Nana (P.G.232-243, *P. granatum* var. Nana), which is a dwarf pomegranate. The mature plant height of this cultivar is 70 cm. The other parent, Black, has normal pomegranate plant height, about 3 m. The F1 plant was medium sized, about 2.7 m at 6 years after planting. The F2 population plant height was highly polymorphic (Fig. 1a), it was measured in the second and third year after planting. Plant height segregated from 0.6 to 3 m when measured 3 years from planting. During the third year, the growth rate distribution was from 2 to 187 cm/year. The distribution of the two traits was normal according to the Shapiro–Wilk goodness-of-fit test ($p \geq 0.01$) (Fig. 3h, i). There was a high correlation ($r=0.75$) between plant height in the fifth year of growth and in the third year of growth.

The correlations between plant height and fruit traits were low ($r \leq 0.42$), as shown in Table 3.

Fruit traits within the pomegranate collection

The ARO pomegranate germplasm collection includes more than a hundred accessions from all over the world (Ophir et al. 2014). The fruits of each variety were evaluated during at least 3 years. Six fruit traits were analyzed within the collection (Table 4). Fruit weight, acid content, TSS to acid ratio, and seed hardness were highly polymorphic within the collection, as can be concluded from the traits' high standard deviations (Table 4). However, TSS and aril weight were less variable within the collection. Fruit weight, TSS, and aril weight were normally distributed as was examined by Shapiro–Wilk W test ($p \geq 0.01$).

Previous work divided the ARO pomegranate germplasm collection to two main groups by STRUCTURE analysis

(Ophir et al. 2014); however, this division was not significant ($p \geq 0.05$) for the fruit traits, except for seed hardness.

Identification of QTLs in the Nana × Black F2 population

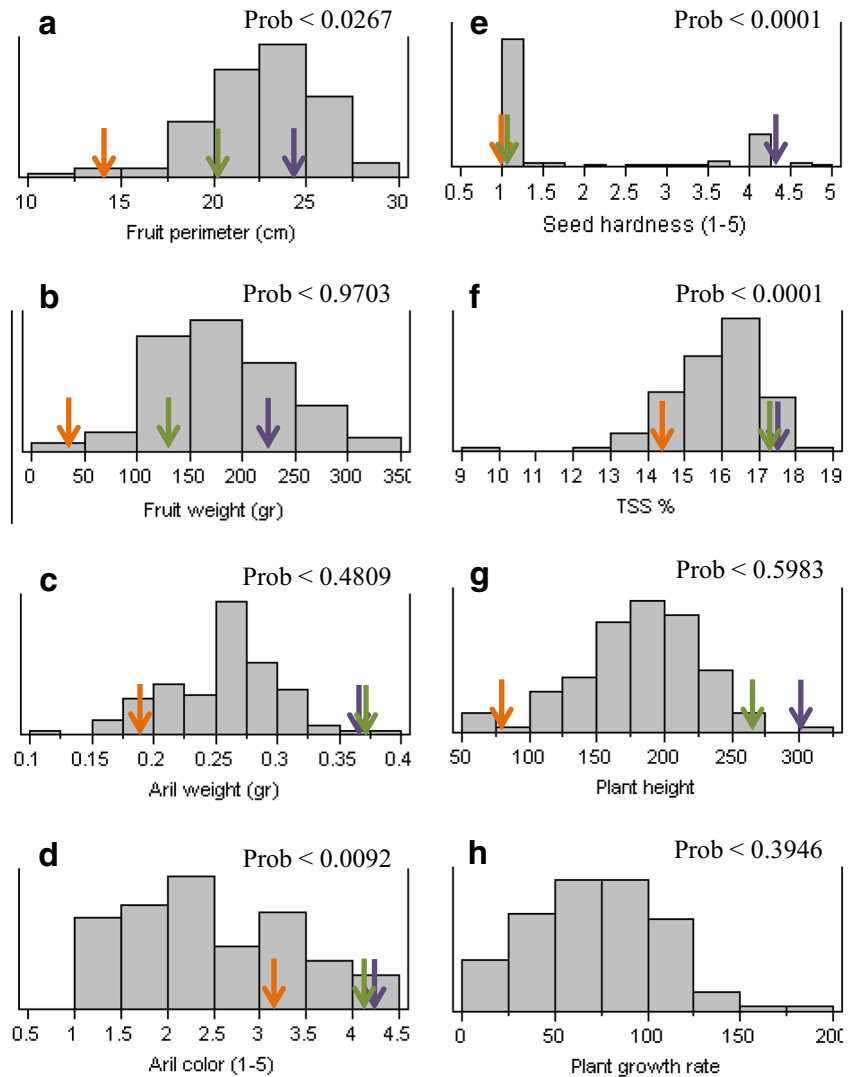
QTLs were analyzed by MQM analysis, which narrows down the QTL around the most significant markers (Van Ooijen 2004). The QTL threshold base was calculated by the permutation analysis (1000 permutation at $p=0.05$), ranging from 4.0 for aril color, aril weight, and fruit weight to 26.4 for seed hardness in the 2013 fruit season. Twenty-five QTLs were analyzed for seven different traits (Table 5 and Fig. 2), including plant growth rate, fruit weight and perimeter, aril weight and color, seed hardness, and TSS of the juice. LOD values varied from 39.45 to 3.97. The highest LOD value was 39.45 for seed hardness at LG1 (114–125 cM). The QTLs were analyzed at six linkage groups out of ten. Only two QTLs, one for seed hardness and one for fruit weight, were statistically significant for both 2012 and 2013 fruit seasons; all the other QTLs were significant for one fruit season only suggesting age effect on the traits within the population.

The highest significant QTL was for seed hardness (*Seed hardness 1-1*) at LG1 (114–125 cM) between the markers

Table 3 Correlation between plant height and fruit traits of the F2 population at 2013

	Plant height, third year
Fruit perimeter	0.30
Fruit weight	0.24
Aril weight	0.42
Peel color	0.41
Aril color	-0.3
Seed hardness	-0.08
TSS	0.34

Fig. 3 Distribution of nine traits within the F2 population. Seven fruit traits were evaluated in the segregating population in 2013: fruit perimeter (a), fruit weight (b), aril weight (c), aril color (d), seed hardness (e), and TSS (f). The results are averages of three fruits per plant. Plant height (g) and plant growth rate (h) were measured in the third year of growth. The distribution is accepted as normal when the prob >0.001 as was calculated by Shapiro–Wilk *W* test. Orange, purple, and green arrows point to the Nana, Black, and F1 parents values in 2014, respectively



c22061_572 and c21761_335. It had a very high LOD value of 18.33 or 39.45 for the 2012 and 2013 fruit seasons, respectively, and can explain 65 or 94 % of the trait for the 2012 or 2013 fruit season, respectively. The hard seed trait was contributed by the Nana parent. Plants with Nana alleles at this locus are with low evaluation score, which means harder seeds.

Plant height was measured in the third and fifth year of growth. The same two QTLs were identified for both measurements. The more significant QTL (*plant height 2-1*) is at LG2 (51–80 cM), where the markers c6449_982, c15646_223, c12777_264, and c6269_275_fl are mapped. The LOD values of this QTL were 11.44 for the third year height, which explains 38.6 % of the trait, and 6.3, which

Table 4 Phenotypic data of the variety collection for six traits

Traits	Average±SD	Maximum	Minimum	Shapiro–Wilk <i>W</i> test
Fruit weight (g)	355.9±117.02	617.6	40	0.1452
TSS (%)	14.92±1.14	19.9	11.8	0.0936
Acid (%)	0.84±0.8	3.9	0.1	<0.0001
TSS/acid	28.46±19.28	96.8	3.9	<0.0001
Aril weight (g)	0.37±.08	0.7	0.2	0.0117
Seed hardness (1-5)	3.05±1.25	5	1	<0.0001

The traits were described over three seasons. The data is summarized as the means, SD, and maximum and minimum values. The probability of normal distribution was analyzed by Shapiro–Wilk *W* test

Table 5 Significant QTLs (LOD >3.0) for fruit traits and plant height identified by interval mapping analysis

QTL ^a	Trait	Year of trial	Group	Length of QTL (cM)	Locus position (cM) ^b	Marker close to the highest LOD	Max. LOD ^c	Max. mu_A ^d	Max. mu_H ^e	Max. mu_B ^f	Max. variance ^g	Max. % expl. ^h	Additive ⁱ	Max. dominance ^j
Seed hardness 1-1	Seed hardness	2012	LG1	10	114-124		18.33	1.24	1.17	3.55	0.29	65	-1.16	-1.23
Seed hardness 1-1	Seed hardness	2013	LG1	11	114-125		39.45	1.03	1.07	3.85	0.1	94.1	-1.41	-1.38
Fruit weight 1-1	Fruit weight	2013	LG1	1	106	c5689_325	7.64	128.52	193.73	174.89	1425.45	18.7	-25.39	44.23
Fruit weight 2-1	Fruit weight	2012	LG2	3	52-55	c12777_264	5.97	105.99	171.79	137.73	1960.67	32.1	-18.55	54.88
Fruit weight 2-1	Fruit weight	2013	LG2	3	52-55	c12777_264	6.72	133.78	196.55	173.47	1394.18	16.8	-19.84	48.36
Fruit weight 3-1	Fruit weight	2013	LG3	2	28-30	c10946_246	7.82	113.93	125.67	185.82	1363.9	19.2	-36.32	-25.95
Fruit weight 7-1	Fruit weight	2013	LG7	1	132	c28730_747	3.97	178.53	141.78	120.46	1288.67	8.6	29.03	-7.71
Fruit perimeter 2-1	Fruit perimeter	2013	LG2	6	45-51	c6449_982	6.49	18.9	21.99	21.55	4.369	19.4	-1.47	2.22
Fruit perimeter 2-2	Fruit perimeter	2012	LG2	9	52-61	c6269_275_fl	8.62	17.59	21.47	20.26	5.74	42.3	-1.71	3.18
Fruit perimeter 2-3	Fruit perimeter	2013	LG2	3	121-124	c36033_605	3.92	19.94	21.96	19.89	3.94	10.8	-0.07	2.21
Aril weight 2-1	Aril weight	2012	LG2	2	37-38	c7870_3315_fl	5.58	0.24	0.28	0.25	0.0007	8	-0.01	0.04
Aril weight 2-2	Aril weight	2013	LG2	6	45-51	c6449_982	7.35	0.22	0.27	0.27	0.001	36.5	-0.03	0.04
Aril weight 2-3	Aril weight	2012	LG2	2	107-108	c26352_921	5.49	0.22	0.24	0.27	0.0007	8.2	-0.02	-0.01
Aril weight 3-1	Aril weight	2012	LG3	1	1	c26309_285	3.79	0.25	0.27	0.23	0.0006	5.3	0.01	0.03
Aril weight 3-2	Aril weight	2012	LG3	1	122	c15747_186	8.38	0.21	0.24	0.27	0.0007	13.7	-0.03	0.002
Aril weight 10-1	Aril weight	2012	LG10	8	14-22	c8054_297	8.11	0.18	0.17	0.31	0.0006	13.9	-0.07	-0.08
Aril color 1-1	Aril color	2013	LG1	1	22	c20430_187	6.38	2.99	2.02	2.15	0.35	19.5	0.42	-0.58
Aril color 1-2	Aril color	2012	LG1	2	68-69	c17611_336	5.23	2.19	1.94	1.27	0.38	22.7	0.51	0.26
Aril color 2-1	Aril color	2013	LG2	1	108	c26352_921	4.02	2.59	3.19	2.55	0.33	11.4	0.02	0.63
Aril color 3-1	Aril color	2013	LG3	1	16	c9471_337	4.78	2.68	1.91	2.45	0.33	13.9	0.12	-0.66
Aril color 5-1	Aril color	2013	LG5	1	44	c13232_1849_fl	4.38	2.06	2.66	3.07	0.33	12.6	-0.51	0.09
TSS 2-1	TSS	2012	LG2	2	36-38	c7870_3315_fl	6.03	12.7	14.55	13.96	1.19	21.8	-0.64	1.25
TSS 2-2	TSS	2013	LG2	1	52	c15646_223	6.76	14.55	15.77	15.88	0.51	16.3	-0.67	0.56
Plant height 2-1	Plant height	2013	LG2	1	51	c15646_223	11.44	147.93	208.56	214.25	854.92	38.6	36.06	29.39
Plant height 2-1	Plant height	2015	LG2	8	52-60	c6269_275_fl	6.26	164.86	216.6	225.5	1364.21	25.2	32.01	24.45

Table 5 (continued)

QTL ^a	Trait	Year of trial	Group	Length of QTL (cM)	Locus position (cM) ^b	Marker close to the highest LOD	Max. LOD ^c	Max. mu_A ^d	Max. mu_H ^e	Max. mu_B ^f	Max. variance ^g	Max. % expl. ^h	Additive ⁱ	Max. dominance ^j
Plant height 3-1		2013	LG3	1	37	c14583_853	4.63	158.96	176.61	203.18	768.63	12.3	22.11	-4.46
Plant height 3-1		2015	LG3	1	36.733	c14583_853	4.94	163.44	183.24	220.63	1281.53	18.7	28.59	-8.8

^a QTLs are defined by the trait abbreviation, linkage group number, and QTL number

^b Position of the QTL in centimorgans on the linkage group

^c Maximum LOD score for each QTL

^d Maximum score in trait values for Nana (A) alleles for each QTL

^e Maximum score in trait values for Black (B) alleles for each QTL

^f Maximum score in trait values for heterozygous for each QTL

^g Maximum variance for each QTL

^h Maximum percent of explanation for each QTL

ⁱ Additive effect is positive when Nana alleles increase the trait score and negative when Black alleles increase the trait score

^j Maximum dominance for each QTL

explains 25 % of the trait for the fifth year height. The additional QTL (*plant height 3-1*) is at LG3 (37 cM) with a LOD value higher than 4.6, which explains about 15 % of the trait. Two QTLs for fruit size traits were identified in close proximity to the position of plant height 2-1 QTL. A QTL for fruit weight (*fruit weight 2-1*) was localized at LG2 (52–55 cM) with a LOD value of 5.97; the marker c12777_264 had the highest LOD value at this QTL. This QTL was analyzed for two successive fruit seasons. In addition, a QTL for fruit perimeter (*fruit perimeter 2-1*) was detected at LG2 (45–51 cM) with a LOD value of 6.49; the highest LOD value was by the marker c6449_982. At the same location, a QTL for aril weight (*aril weight 2-2*) was localized with a LOD value of 7.35. The juice TSS trait, as measured in the fruit season of 2013, also had a QTL (*TSS 2-2*) at LG2 (52 cM) with a LOD value of 6.76. It seems that the LG2 (50–55 cM) region contained sites for several fruit quality traits and for the plant height trait. However, all these fruit traits have more than one QTL.

There are two QTLs for fruit weight with almost the same LOD values, one (*fruit weight 1-1*) at LG1 (106 cM) with a LOD value of 7.64, the other one (*fruit weight 3-1*) at LG3 (28–30 cM) with a LOD value of 7.82. An additional smaller QTL (*fruit weight 7-1*) was located at LG7 (132 cM) with a LOD value of 3.97. Surprisingly, at this locus, the Nana alleles contributed to the existence of heavier fruits. The four QTLs for fruit weight were analyzed for the 2013 fruit season, each of the main three, at LG1, LG2, and LG3, which explains about 20 % of the trait. For the 2012 fruit season, only one QTL was analyzed at LG2 (52–55 cM), at the same location as in the 2013 season, which explains 32.1 % of the trait.

Fruit perimeter in the 2013 fruit season had two QTLs, one at LG2, fruit perimeter 2-1, as described, in the region (45–51 cM), and it explains 19.4 % of the trait. An additional QTL (*fruit perimeter 2-3*) with a lower LOD value was detected at another part of LG2 (121–124 cM). A single QTL was analyzed for the 2012 season (*fruit perimeter 2-2*) at LG2 (61 cM). This QTL was more significant than the QTLs identified for the 2013 fruit season, with a LOD value of 8.62, and it explains 42.3 % of the trait.

Six QTLs were identified for the aril weight trait. Only one of these was found for the 2013 season (*aril weight 2-2*), which explains the highest percent of the trait (36.5 %). The other five QTLs were analyzed in the 2012 fruit season, the most significant (LOD=8.38) was *aril weight 3-2* at LG3 (122 cM), close to SNP c15747_186. There is an additional QTL (*aril weight 3-1*) at LG3 with a lower LOD value (LOD=3.79). Another highly significant QTL (LOD=8.11) is positioned at LG10 (*aril weight 10-1*). Two separate QTLs (*aril weight 2-1* and *aril weight 2-3*) were at LG2 (37 and 107 cM, respectively), both with a

Table 6 Significant associations by GLM ($p \leq 10^{-4}$) for fruit traits and SNPs within the variety collection

Trait	Marker	LG	Locus position	Association p value	R^2
Seed hardness	c4861_219	LG2	22.141	7.27E-08	0.3
Seed hardness	c3506_2098	LG3	35.29	6.13E-04	0.15
Seed hardness	c4399_2863	LG3	120.334	8.32E-04	0.15
Seed hardness	c10126_309	LG6	45.554	2.92E-04	0.17
Aril weight	c3506_2098	LG3	35.29	5.69E-04	0.15
Aril weight	c7562_1997	LG4	116.17	1.41E-05	0.23
Fruit weight ^a	c7552_803	LG1	103.806	2.48E-05	0.21
Fruit weight	c3177_1224	LG1	129.38	7.43E-04	0.14
Fruit weight	c12815_439	LG1	144.136	7.77E-04	0.15
Fruit weight ^a	c8107_922	LG2	42.631	5.53E-04	0.15
Fruit weight	c7874_1707	LG2	93.72	5.82E-07	0.27
Fruit weight	c11182_1077	LG3	53.207	5.98E-04	0.16
Fruit weight	c3075_725	LG3	62.12	1.97E-07	0.29
Fruit weight	c13857_900	LG4	66.03	1.57E-04	0.17
Fruit weight	c7749_1488	LG5	32.48	7.28E-04	0.15
Fruit weight	c9544_1132	LG6	22.103	4.28E-05	0.2
Fruit weight	c7888_929	LG6	25.363	3.85E-04	0.16
Fruit weight	c6289_1362	LG6	26.623	1.62E-05	0.22
Fruit weight	c8681_1903	LG8	62.184	2.19E-05	0.21
TSS	c7552_803	LG1	103.806	6.13E-07	0.27

^aSignificantly associated within the variety collection and positioned at QTL to fruit weight

LOD value of about 5.5, these QTLs are not at the same position as aril weight 2-2.

Aril color was evaluated and scored from 1 to 5. The evaluation outcome was analyzed as a continuous trait. Five QTLs were analyzed for this trait. One of them (*aril color 1-2*) for the 2012 fruit season is at LG1 (67–69 cM). The other four aril color QTLs were analyzed for the 2013 fruit season, including an additional one at LG1 (*aril color 1-1*), but at a different position (22 cM), one at LG2 (*aril color 2-1*), one at LG3 (*aril color 3-1*), and one at LG5 (*aril color 5-1*). In QTL aril color 5-1, the Nana alleles reduced the evaluation score, while in all the other QTLs, the Nana alleles increased the evaluation scores, which meant increasing the aril red color.

TSS of the aril juice was assessed for the fruits of the population in the 2012 and 2013 fruit seasons. Two QTLs were analyzed at LG2, one for the 2012 fruit season (*TSS 2-1*) and *TSS 2-2* in the 2013 fruit season. The two QTLs had a LOD value of about 6. Although the two QTLs are located on LG2, they are separated by 20 cM. The Black alleles contributed to higher TSS values.

Association study

The ARO germplasm collection was genotyped with 346 SNPs by Fluidigm-EP platform (Ophir et al. 2014). Association between fruit traits and polymorphic SNPs was analyzed

by the TASSEL program using initially the GLM analysis (Table 6). SNPs found to be significantly associated by the GLM ($p \leq 10^{-4}$) were verified for normal allele distribution ($p \geq 0.01$), and the significance of the association was confirmed by ANOVA or Wilcoxon/Kruskal–Wallis tests (Online Resource Table 4). The associations were among seed hardness, aril weight, fruit weight, and TSS. The associated traits were anchored to seven linkage groups on the genetic map by common markers (Fig. 2). A marker's R^2 is the portion of total variation explained by that marker. The maximum R^2 value (0.3) was found between SNP c4861_219 at LG2 (22.141 cM) and seed hardness. Fruit weight trait was the only trait that overlapped between the markers that were associated by association mapping and linked by QTL mapping, at two different locations. The QTL fruit weight 1-1 at LG1 (106 cM) with a high LOD value of 7.64 is close to SNP c7552_803_fl positioned at LG1 (104 cM). The other common position was at LG2, where the QTL fruit weight 2-1 at LG2 (52–55 cM), which was significant for both the 2012 and the 2013 seasons, is close to the associated SNP c8107_922_fl at LG2 (43 cM). When TASSEL was used to take in account the kinship and Q value from structure analysis within the germplasm collection (MLM), the corrected results yielded the data presented in Table 7. It can be seen that only acid content and fruit weight were significantly associated ($p \leq 10^{-3}$).

Table 7 Significant associations by MLM ($p \leq 10^{-3}$) for fruit traits and SNPs within the variety collection

Trait	Marker	LG	Locus position	Association p value	R^2
Acid	c11182_1077	LG3	53.207	3.93E-08	0.45
	c5614_1406	LG8	18.39	1.59E-04	0.21
	c7721_1222	Not mapped		4.17E-10	0.61
	c8554_297	Not mapped		2.46E-09	0.54
	c6885_901	Not mapped		2.17E-09	0.55
Fruit weight	c5614_1406	LG8	18.39	8.62E-04	0.16
	c3213_803	LG8	18.39	9.88E-04	0.16

Discussion

Based on novel SNP markers and an F2 population, we were able for the first time to assemble a genetic map for pomegranate. Although several types of genetic markers were used to genotype some of the world pomegranate collections (Currò et al. 2010; Hasnaoui et al. 2010, 2012; Pirseyedi et al. 2010; Basaki et al. 2011; Sarkhosh et al. 2011; Soriano et al. 2011; Jian et al. 2012; Noormohammadi et al. 2012; Zhang et al. 2012; Hajiahmadi et al. 2013; Singh et al. 2013; Ferrara et al. 2014; Orhan et al. 2014), these sets of markers were separately used for each collection and no general reference for relative position of the various markers was available. The Nana×Black genetic map provides for the first time an important reference tool of the pomegranate genome for positioning markers and traits. This genetic linkage map may be useful as a basic pomegranate map. Since this map includes more than 230 SNPs that are polymorphic within the variety collection, it can be of relevance for all pomegranate varieties. The current pomegranate genetic map is spread over 1141 cM. The markers are spread at an average of 1.17-cM intervals and the 1200 SNP markers are all positioned within transcribed genes. Because these genetic markers are within genes, they specify genes as anchors for genetic markers even in populations where the original SNPs themselves are not polymorphic.

The Nana×Black F2 population was chosen for the construction of the genetic map because the two parent accessions Nana and Black differ in many characteristics, and the resulting F2 population was expected to be genetically diverse. Nana and Black differ in important characteristics including fruit quality parameters, tree size and growth habit. Previous genetic classification divides the pomegranate germplasm collection into two statistically significant distinct genetic groups, G1 and G2; each of the parents of the population belongs to a different group. Nana belongs to the G1 group. The paternal parent, the Black accession, is part of the G2 group (Ophir et al. 2014).

The heterozygosity in the mapped F2 population was found to be 47 %, which is expected of an F2 population (Van Ooijen and Voorrips 2001). The maternal (Nana) alleles were 30 % more abundant within the F2 population. SNPs of 176 were

abnormally distributed, mainly in four linkage groups. Deviating distortion is known in other mapped population of different plants. In maize, two F2 populations were compared before and after five generations of intermating; deviant segregation ratios were observed in both F2 populations; the overall allele frequencies were 51 % and 49 % for the different parents' alleles (Lee et al. 2002). However, ratios of genotypic classes at 38 loci distributed among all chromosomes deviated from the expectations in F2. In a recombinant inbred line population of tomato, a significant deviation from the expected 1:1 ratio between the two homozygous classes was found in 73 % of the markers. In 98 % of the deviating markers, *Lycopersicon esculentum* alleles were present at a greater frequency than that of the *Lycopersicon cheesmanii* alleles; the skewed loci were scattered throughout the genome (Paran et al. 1995). Genotypic analysis of individual progenies within the pomegranate F2 population revealed that five plants contained more than twice the alleles of Nana. Phenotypic characterization of these individual progenies revealed that they were not dwarfed like Nana, nor did they resemble Nana in other phenotypic aspects. The possibility of obtaining such a high enrichment of parental alleles in F2 progenies is of high importance for breeding, as it provides a useful tool to hasten breeding towards a similarity of genotypes to one of the parents in advanced generations. Identification of shifts in allele frequencies within the genome can be important information for monitoring specific alleles or haplotypes and can be used to design appropriate breeding strategies (Collard and Mackill 2008).

QTL mapping allowed us to map several agriculturally important traits including TSS, fruit size, seed hardness, and plant size. While multiple genes were found to be responsible for some phenotypic traits in pomegranate (for example, fruit weight and aril weight), only one QTL was detected for fruit weight in peach (Dirlewanger et al. 1999) and two major QTLs were detected in tomato (Causse et al. 2002). For other traits such as TSS, seed hardness, and plant height, only two sites on the genetic map were found to be involved. Major genes were found for each of these three traits, suggesting the potential of the map to facilitate breeding for these traits. QTLs for TSS, seed hardness, and plant height were characterized in other plants as well. In peach, two major QTLs for

TSS were also found (Dirlewanger et al. 1999). The seed hardness trait was mapped also in soybean (Zhang et al. 2008). Interestingly, seed hardness in soybean was also found to be mapped to two loci in the genetic map. QTL mapping of plant height was done in populus. In this plant, only one QTL was identified for plant height in the second year of growth (Bradshaw and Stettler 1995). In addition, in peach, growth habit was not mapped as one gene trait (Sajer et al. 2012).

A few QTLs were mapped in close proximity at LG2, including QTLs for plant height, fruit weight, fruit perimeter, aril weight, and TSS. One explanation for this phenomenon is that plant size has a pleiotropic effect on fruit traits. In rice, plant size was mapped to the same locus as grain yield, heading date (Zhang et al. 2006; Yan et al. 2011). Indeed, map-based cloning reveals that this locus encodes a putative HAP3 subunit of the CCAAT-box-binding transcription factor (Wei et al. 2010). In a complementary experiment, this gene increased plant height and the number of grains per panicle significantly. Another explanation is that the proximity of the QTLs is due to closely linked multiple QTLs. In tomato, it was shown that a cluster of QTLs of multiple characters such as fruit weight, fruit elasticity, color, sourness, aroma intensity, candy aroma, mealiness, dry matter weight, soluble solids, sugars, and eugenol content could be dissected by further genetic studies, such as fine mapping of this region (Causse et al. 2002).

Association studies for pomegranate were conducted within local collections, Basaki et al. (2011, 2013) who associated seven SSRs with 20 traits within the Iranian collection. Singh et al. (2015) analyzed the association of three traits, by GLM, with 44 microsatellites within the Indian collection. However, the map position of these SSRs was not determined. The present study used more than 200 mapped markers to analyze the association with fruit traits within a worldwide collection. When the structure of the collection was considered in the analysis, two traits (fruit weight and acid content) were found to be associated. The associated markers were mapped to two linkage groups.

The linkage map provides basic data about the order of markers and genes along the genome and detects QTLs for interesting traits. However, the QTLs might be significant only for a specific population due to exclusive genetic differences between the two parents. The association study within the variety collection took advantage of historic recombination events. In maize, for example, the nested association mapping population showed that there is a variation for recombination frequencies (McMullen et al. 2009); therefore, there could be different recombination events between two varieties that are not presented in the linkage map. Common mapped markers that were associated with traits within the collection of varieties can give information about the position of the associated markers even if no similar QTL was detected within the mapping population. Thus, population mapping and

studies of associations within collections can yield a broader view of the number of loci involved in the control of complex traits. Moreover, QTLs that are co-localized with associated markers to the same trait, as was found for the two fruit weight QTLs, strongly suggest that these QTLs could be relevant in several populations and highly informative for breeding. Up till now, no draft of the pomegranate genome has been reported. Genome sequence is today a crucial tool for functional analysis of genes in many plant species and its absence severely inhibits the ability to associate gene function with important traits. A dense linkage map may be a platform to facilitate a new assembly of a genome (Mascher and Stein 2014). The pomegranate genetic map described in this study together with the de novo assembly of the pomegranate transcriptome published earlier could contribute significantly to the establishment of the pomegranate genome.

Acknowledgments We thank Kamel Hatib for growing and maintaining the orchards. This research was supported by funding from the Chief Scientist Office, Ministry of Agriculture & Rural Development, <http://www.science.moag.gov.il/>, Agricultural Biotechnology 203-0753-09.

Authors' contribution statement R. H. B. conducted the phenotyping of the mapping population and the genetic and mapping analysis and wrote the paper; A. S. and R. E. conducted the sequencing, marker development, and the genotyping of the mapping population; M. R. and R. O. conducted the marker development, array and probe design, array hybridization and signal preprocessing, and genotype calling; I. B. Y. conducted the phenotyping of the ARO collection and the mapping population; T. T. conducted the total RNA extraction and developed the method for DNA extraction from pomegranate; and D. H. wrote the paper and designed and supervised the study. All authors discussed the results and commented on the manuscript.

Data Archiving Statement Data base on which this study was based was published before by Ophir et al. (2014) and deposited in GenBank ([http://www.ncbi.nlm.nih.gov/sra/SRX395468\[accn\]](http://www.ncbi.nlm.nih.gov/sra/SRX395468[accn]) and [http://www.ncbi.nlm.nih.gov/sra/SRX395465\[accn\]](http://www.ncbi.nlm.nih.gov/sra/SRX395465[accn])).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Basaki T, Choukan R, Nekouei SMK, Mardi M, Majidi E, Faraji S, Zeinolabedini M (2011) Association analysis for morphological traits in pomegranate (*Punica granatum* L.) using microsatellite markers. *Middle-East J Sci Res* 9:410–417
- Basaki T, Nejat MA, Nejad RJ, Faraji S, Keykhaei F (2013) Identification of simple sequences repeat (SSR) informative markers associated with important traits in pomegranate (*Punica granatum* L.). *Int J Agron Plant Prod* 4:575–583
- Ben-Simhon Z, Judeinstein S, Nadler-Hassar T, Trainin T, Bar-Ya'akov I, Borochov-Neori H, Holland D (2011) A pomegranate (*Punica granatum* L.) WD40-repeat gene is a functional homologue of *Arabidopsis TTG1* and is involved in the regulation of anthocyanin

- biosynthesis during pomegranate fruit development. *Planta* 234: 865–881
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635
- Bradshaw HD, Stettler RF (1995) Molecular genetics of growth and development in populus. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. *Genetics* 139: 963–973
- Causse M, Saliba-Colombani V, Lecomte L, Duffe P, Rousselle P, Buret M (2002) QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. *J Exp Bot* 53:2089–2098
- Chai HS, Therneau TM, Bailey KR, Kocher JP (2010) Spatial normalization improves the quality of genotype calling for Affymetrix SNP 6.0 arrays. *BMC Bioinforma* 11:356
- Chevreur B, Pfisterer T, Drescher B, Driesel AJ, Muller WE, Wetter T, Suhai S (2004) Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 14:1147–1159
- Collard BCY, Mackill DJ (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos Trans R Soc B Biol Sci* 363:557–572
- Curro S, Caruso M, Distefano G, Gentile A, La Malfa S (2010) New microsatellite loci for pomegranate, *Punica granatum* (Lythraceae). *Am J Bot* 97:e58–e60
- Danin-Poleg Y, Tadmor Y, Tzuri G, Reis N, Hirschberg J, Katzir N (2002) Construction of a genetic map of melon with molecular markers and horticultural traits, and localization of genes associated with *ZYMV* resistance. *Euphytica* 125:373–384
- Di Guardo M, Tadiello A, Farneti B, Lorenz G, Masuero D, Vrhovsek U, Costa G, Velasco R, Costa F (2013) A multidisciplinary approach providing new insight into fruit flesh browning physiology in apple (*Malus x domestica* Borkh.). *PLoS One* 8:e78004
- Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R (1999) Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). *Theor Appl Genet* 98:18–31
- Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175–185
- Fan S, Bielenberg DG, Zhebentyayeva TN, Reighard GL, Okie WR, Holland D, Abbott AG (2010) Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytol* 185:917–930
- Ferrara G, Giancaspro A, Mazzeo A, Giove SL, Matarrese AMS, Pacucci C, Punzi R, Trani A, Gambacorta G, Blanco A, Gadaleta A (2014) Characterization of pomegranate (*Punica granatum* L.) genotypes collected in Puglia region, Southeastern Italy. *Sci Hortic* 178:70–78
- Graham IA, Besser K, Blumer S, Branigan CA, Czechowski T, Elias L, Guterman I, Harvey D, Isaac PG, Khan AM, Larson TR, Li Y, Pawson T, Penfield T, Rae AM, Rathbone DA, Reid S, Ross J, Smallwood MF, Segura V, Townsend T, Vyas D, Winzer T, Bowles D (2010) The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science* 327: 328–331
- Hajiahmadi Z, Talebi M, Sayed-Tabatabaei BE (2013) Studying genetic variability of pomegranate (*Punica granatum* L.) based on chloroplast DNA and barcode genes. *Mol Biotechnol*
- Harel-Beja R, Tzuri G, Portnoy V, Lotan-Pompan M, Lev S, Cohen S, Dai N, Yesselson L, Meir A, Libhaber SE, Avisar E, Melame T, van Koert P, Verbakel H, Hofstede R, Volpin H, Oliver M, Fougere A, Stalh C, Fauve J, Copes B, Fei Z, Giovannoni J, Ori N, Lewinsohn E, Sherman A, Burger J, Tadmor Y, Schaffer AA, Katzir N (2010) A genetic map of melon highly enriched with fruit quality QTLs and EST markers, including sugar and carotenoid metabolism genes. *Theor Appl Genet* 121:511–533
- Hasnaoui N, Buonamici A, Sebastiani F, Mars M, Trifi M, Vendramin G (2010) Development and characterization of SSR markers for pomegranate (*Punica granatum* L.) using an enriched library. *Conserv Genet Resour* 2:283–285
- Hasnaoui N, Buonamici A, Sebastiani F, Mars M, Zhang D, Vendramin GG (2012) Molecular genetic diversity of *Punica granatum* L. (pomegranate) as revealed by microsatellite DNA markers (SSR). *Gene* 493:105–112
- Holland D, Bar-Ya'akov I (2014) Pomegranate: aspects concerning dynamics of health beneficial phytochemicals and therapeutic properties with respect to the tree cultivar and the environment. In: Yaniv Z, Dudai N (eds) *Medicinal and aromatic plants of the Middle-East*. Springer, Netherlands, pp 225–239
- Holland D, Hatib K, Bar-Ya'akov I (2009) Pomegranate: botany, horticulture, breeding. *Hortic Rev* 35:127–191
- Jian Z-H, Liu X-S, Hu J-B, Chen Y-H, Feng J-C (2012) Mining microsatellite markers from public expressed sequence tag sequences for genetic diversity analysis in pomegranate. *J Genet* 91:353–358
- Lee M, Sharopova N, Beavis W, Grant D, Katt M, Blair D, Hallauer A (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. *Plant Mol Biol* 48:453–461
- Liu W, Di X, Yang G, Matsuzaki H, Huang J, Mei R, Ryder TB, Webster TA, Dong S, Liu G, Jones KW, Kennedy GC, Kulp D (2003) Algorithms for large-scale genotyping microarrays. *Bioinformatics* 19:2397–2403
- Mascher M, Stein N (2014) Genetic anchoring of whole-genome shotgun assemblies. *Front Genet* 5
- Masoud S, Khandan M, Nasre-Esfahani S (2005) Cytogenetical study of some Iranian pomegranate (*Punica granatum* L.) cultivars. *Caryologia* 58:132–139
- McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, Brown P, Browne C, Eller M, Guill K, Harjes C, Kroon D, Lepak N, Mitchell SE, Peterson B, Pressoir G, Romero S, Rosas MO, Salvo S, Yates H, Hanson M, Jones E, Smith S, Glaubitz JC, Goodman M, Ware D, Holland JB, Buckler ES (2009) Genetic properties of the maize nested association mapping population. *Science* 325:737–740
- Noormohammadi Z, Fasihee A, Homaei-Rashidpoor S, Sheidai M, Baraki SG, Mazooji A, Tabatabaee-Ardakani SZ (2012) Genetic variation among Iranian pomegranates (*Punica granatum* L.) using RAPD, ISSR and SSR markers. *Aust J Crop Sci* 6:268–275
- Olukolu BA, Trainin T, Fan SH, Kole C, Bielenberg DG, Reighard GL, Abbott AG, Holland D (2009) Genetic linkage mapping for molecular dissection of chilling requirement and budbreak in apricot (*Prunus armeniaca* L.). *Genome* 52:819–828
- Ophir R, Sherman A, Rubinstein M, Eshed R, Sharabi Schwager M, Harel-Beja R, Bar-Ya'akov I, Holland D (2014) Single-nucleotide polymorphism markers from de-novo assembly of the pomegranate transcriptome reveal germplasm genetic diversity. *PLoS One* 9: e88998
- Orhan E, Ersicli S, Esitken A, Sengul M (2014) Molecular and morphological characterization of pomegranate (*Punica granatum* L.) genotypes sampled from Coruh Valley in Turkey. *Genet Mol Res* 13: 6375–6382
- Paran I, Goldman I, Tanksley SD, Zamir D (1995) Recombinant inbred lines for genetic mapping in tomato. *Theor Appl Genet* 90:542–548
- Pirseyedi SM, Valizadehghan S, Mardi M, Ghaffari MR, Mahmoodi P, Zahravi M, Zeinalabedini M, Nekoui SMK (2010) Isolation and characterization of novel microsatellite markers in pomegranate (*Punica granatum* L.). *Int J Mol Sci* 11:2010–2016
- Porebski S, Bailey LG, Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Report* 15:8–15
- Rubinstein M, Katzenellenbogen M, Eshed R, Rozen A, Katzir N, Colle M, Yang L, Grumet R, Weng Y, Sherman A, Ophir R (2015) Ultra high-density linkage map for cultivated cucumber (*Cucumis sativus*

- L.) using a single nucleotide polymorphism genotyping array. *PLoS one* In press
- Sajer O, Scorza R, Dardick C, Zhebentyayeva T, Abbott AG, Hom R (2012) Development of sequence-tagged site markers linked to the pillar growth type in peach (*Prunus persica*). *Plant Breed* 131:186–192
- Sarkhosh A, Zamani Z, Fatahi R, Hassani M, Wiedow C, Buck E, Gardiner S (2011) Genetic diversity of Iranian soft-seed pomegranate genotypes as revealed by fluorescent-AFLP markers. *Physiol Mol Biol Plants* 17:305–311
- Seeram NP, Schulman RN, Heber D (2006) Pomegranates: ancient roots to modern medicine. CRC Press Taylor & Francis Group, Boca Raton
- Singh SK, Meghwal PR, Pathak R, Gautam R, Kumar S (2013) Genetic diversity in *Punica granatum* revealed by nuclear rRNA, internal transcribed spacer and RAPD polymorphism. *Natl Acad Sci Lett* 36: 115–124
- Singh N, Abburi V, Ramajayam D, Kumar R, Chandra R, Sharma K, Sharma J, Babu KD, Pal R, Mundewadikar D, Saminathan T, Cantrell R, Nimmakayala P, Reddy U (2015) Genetic diversity and association mapping of bacterial blight and other horticulturally important traits with microsatellite markers in pomegranate from India. *Mol Genet Genomics* 290:1393–1402
- Soriano J, Zuriaga E, Rubio P, Ll acer G, Infante R, Badenes M (2011) Development and characterization of microsatellite markers in pomegranate (*Punica granatum* L.). *Mol Breed* 27:119–128
- Van Ooijen JW (2004) MapQTL[®] 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V, Wageningen
- Van Ooijen JW, Voorrips RE (2001) JoinMap[®] 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen
- Wei XJ, Xu JF, Guo HN, Jiang L, Chen SH, Yu CY, Zhou ZL, Hu PS, Zhai HQ, Wan JM (2010) *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol* 153:1747–1758
- Yan W-H, Wang P, Chen H-X, Zhou H-J, Li Q-P, Wang C-R, Ding Z-H, Zhang Y-S, Yu S-B, Xing Y-Z, Zhang Q-F (2011) A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol Plant* 4:319–330
- Zhang Y, Luo L, Xu C, Zhang Q, Xing Y (2006) Quantitative trait loci for panicle size, heading date and plant height co-segregating in trait-performance derived near-isogenic lines of rice (*Oryza sativa*). *Theor Appl Genet* 113:361–368
- Zhang B, Chen PY, Chen CY, Wang DC, Shi AN, Hou AF, Ishibashi T (2008) Quantitative trait loci mapping of seed hardness in soybean. *Crop Sci* 48:1341–1349
- Zhang YP, Tan HH, Cao SY, Wang XC, Yang G, Fang JG (2012) A novel strategy for identification of 47 pomegranate (*Punica granatum*) cultivars using RAPD markers. *Genet Mol Res* 11: 3032–3041