Dissecting the genetic basis of melon (*Cucumis melo* L.) domestication

Aurora Díaz [1,4], Ana Montserrat Martín-Hernández [2], Ramón Dolcet[2], Ana Garcés-Claver[1], Jose María Álvarez[1], Jordi Garcia-Mas[2], Belén Picó[3], Antonio José Monforte[4,*]

[1]Unidad de Hortofruticultura, Instituto Agroalimentario de Aragón (IA2) (CITA-Universidad de Zaragoza), Avenida de Montañana 930. 50059 Zaragoza, Spain

[2]IRTA, Centre for Research in Agricultural Genomics CSIC-IRTA-UAB-UB, Parc de Recerca de la UAB, C/ Vall de Moronta, Edifici CRAG, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain

[3]Institute for the Conservation and Breeding of Agricultural Biodiversity (COMAV-UPV), Universitat Politècnica de València, Camino de Vera s/n. 46022 Valencia, Spain

[4]Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universitat Politècnica de València (UPV)-Consejo Superior de Investigaciones Científicas (CSIC), Ciudad Politécnica de la Innovación (CPI), Ed. 8E, C/Ingeniero Fausto Elio s/n. 46022 Valencia, Spain

[*]Corresponding author: amonforte@ibmcp.upv.es

Tel: +34 963877783

Abstract

Domestication of crop plants has had a huge impact on the human lifestyle. An analysis of Quantitative Trait Loci (QTL) controlling fruit morphology domestication-related traits, was carried out using an F_2 population from the cross between the Indian wild melon "Trigonus" and the western elite cultivar 'Piel de Sapo' (PS). Twenty-seven QTL were identified at minimum in two out of the three trials, six of them also being detected in BC1 and BC3 populations. Ten of them were related to fruit morphological traits, 12 to fruit size characters, and 5 to pulp content. The Trigonus alleles decreased the value of the character, except for the QTL at andromonoecious gene at linkage group (LG) II, and the QTL for pulp content at LGV. QTL genotypes accounted for a considerable degree of the total phenotypic variation, reaching up to 46%. Around 66% of the QTL showed additive gene action, 19% exhibited dominance, and 25% consisted of overdominance. The regions on LGIV, VI and VIII were identified as the most likely important on the domestication process. QTLs on those regions were validated in BC2S1, BC2S2 and BC3 families, with "Trigonus" allele decreasing the fruit morphological traits in all cases. Introgression lines of these three QTL clusters are being generated, what will predictably contribute to their fine mapping and ultimately to the identification of the causal genes.

Keywords: QTL, fruit size, shape, pulp content, fruit flesh

Author Contribution Statement

AD was involved in molecular marker design, genotyping, population development, phenotyping, data analysis and drafted the manuscript- AMMH and JGM were involved in marker design and genotyping. RD maintained and acclimated the F_2 population *in vitro*. AGC, JMA and BP were in charge of agronomic trials and phenotyping. AJM was involved in the design of the work, data acquisition, analysis, and drafted the manuscript. All authors critically reviewed the manuscript.

Conflict of Interest

Authors declare no conflict of interest

Key Message

The regions on LGIV, VI and VIII involved on fruit morphology traits have been found to be likely associated to melon domestication.

Introduction

Melon (*Cucumis melo* L.) is an economically important fruit crop, with an average production during the last decade higher than 29 million tons per year (FAO, 2016). It is mainly cultivated in temperate and tropical countries, with all the major producer regions located in Asia, with China leading the list. The main producers outside Asia are the United States and Spain, which occupy the sixth and seventh positions, respectively, in world rankings (FAO, 2016).

The origin of *C. melo* is still controversial as wild, non-cultivated melons can be found in both continents of Africa and Asia. Until recently, the most widespread hypothesis placed the origin of the specie in Africa (Kirkbride 1993) whereas Asia was considered an important center of diversification. This was supported by the fact that numerous wild species belonging to the genus *Cucumis*, with the same chromosomal number (2n=2x=24) as the cultivated melon, can still be found in Africa, whereas the Asian species *C. sativus* (cucumber) has a chromosome number of 2n=2x=14. However, recent phylogenetic studies based in nuclear (5.8S ribosomal internal transcribed spacers) and chloroplast sequences revealed that *C. melo* is phylogenetically closer to a cluster of Asian/Australian wild species than to the African species (Sebastian et al. 2010). Furthermore, the sequences of 5.8S ribosomal internal transcribed spacers of two *C. melo* wild forms present in India (usually named as Trigonus (TRI) and Callosus) were nearly identical to those of cultivated melons (Garcia-Mas et al. 2004). Sebastian et al. (2010) concluded that *C. melo* has an Asian/Australian origin and it has a wild Australian sister species (*C. picrocarpus*). The consensus has not been reached on this subject yet as other authors argue that the high chloroplast sequence variability found in African melons supports the origin of cultivated melon in this continent (Tanaka et al. 2013).

Nevertheless, typical wild melons show small leaves and flowers, and thin stems with numerous ramifications. Flowers are monoecious and fruits are oval to round shaped, small (20-50g) with a very thin and inedible flesh surrounding small seeds (Pitrat, 2013). Recent studies have explored the genetic relationships within the species, genotyping up to 175 accessions and cultivars from as much as 50 different countries, which covered all the botanical groups of melon as well as wild African and Asian accessions, with SNP markers (Esteras et al. 2013; Leida et al. 2015). African and Asian wild accessions are similar phenotypically, showing the typical wild melon characteristics but they are clearly genetically different, and they are more similar to cultivars from their respective continent than between them. These results suggest the possibility that at least two domestication events occurred, one in each continent. In fact, there are strong evidences showing that European and Asian cultivars diverged from the Indian subcontinent (Dhillon et al. 2007), that means that their origin was in the wild Indian germplasm. High genetic diversity has been observed in Indian wild melons at both, phenotypic and genetic level (Roy et al. 2012), in part due to frequent gene transfer between cultivated and wild plants, as numerous intermediate forms, as well as returns from domesticated to wild status ("feral"), are relatively frequent in this species (Pitrat, 2012). Thus, the melon domestication process was probably very complex, with genetic interchanges still occurring nowadays in local communities (Roy et al. 2012).

The identification of key traits involved in the domestication helps to focus on the most relevant processes. These traits are generally classified as "domestication syndrome", including increase of vigor, loss of dormancy, reduction of branching, increase seed production, higher fruit/seed size, and removal of toxic compounds (Doebley et al. 2006; Meyer and Purugganan 2013). A typical domestication syndrome trait is expected to be shared among cultivated genotypes and absent or in very low frequency in wild genotypes. Comparing cultivated with wild melons, some characters clearly differentiate them: fruits from cultivated melons are larger (ranging from several hundreds of grams to several kilograms), with a thicker edible flesh (in fact, the flesh is the part of the fruit that is consumed nowadays), and have also larger seeds (Pitrat, 2013; Sabato et al. 2015). Among these characteristics, the most important ones are probably the fruit size and the thickness of the edible flesh, as small-seeded cultivars can be found in Asia. It is also remarkable the absence of cucurbitacins in the domesticated fruits, that are responsible for the bitter taste of wild melons. Cultivated melons also show a wide range of fruit shapes (from slightly flat to extremely elongate), whereas wild melons are usually round

to oval, so fruit shape seems to have been also an important trait during the evolution of melon cultivars.

The accession Ames 24294 (TRI) from North Central Regional Plant Introduction Station (Ames, Iowa, USA), previously classified as *C. trigonus*, shows all the typical characteristics of a wild melon (Monforte et al. 2005, Fig. 1), making it an appropriate choice to study the genetics of melon domestication. Thus, we decided to build a mapping population with it and the Spanish cultivar 'Piel de Sapo' (PS), belonging to the *inodorus* market class. PS fruits are large, oval-shaped, with big seeds and a high content of sweet pulp (Monforte et al. 2005, Fig. 1).

The Quantitative Trait Loci (QTL) mapping strategy has revealed itself successful for the genetic dissection of some of these domestication-related traits in other important crops, like rice, maize, sorghum, barley, wheat, soybean, sunflower, and tomato (reviewed by Olsen and Wendel, 2013). In the case of melon, a number of QTL have been identified for most of the traits cited above (reviewed in Diaz et al. 2011 and Monforte et al. 2014). These studies have commonly been carried out in populations coming from the cross between members of the two different subspecies in melon, *melo* and *agrestis*. This is the case too of the mapping (Harel-Beja et al. 2010) and the ultimate identification (Cohen et al. 2014) of the *pH* gene responsible for the fruit acidity that explains the evolutionary history of sweet melons. However, to date, a wild ancestor of the cultivated melon had not been used to generate such populations. So, the aim of this study was to identify QTL responsible for characters with a crucial role in the process of melon domestication using a population derived from a domesticate-by-wild cross (PS×TRI), as well as the validation of their effects in advanced backcross populations.

Materials and methods

F_2 plant material and trait evaluation

An F_2 population composed by 400 plants coming from the intraspecific cross PS (subsp. *melo* var. *inodorus*) × TRI (subsp. *agrestis* var. *agrestis*) was obtained for this study (Fig. 2). PS is an andromonoecious cultivar while TRI is a monoecious melon type. Sex expression shows pleiotropic effects on fruit morphology, as monoecy leads to more elongated fruits (Monforte et al. 2005; Ramamurthy et al. 2015). For that reason, a screening using the CAPS (Cleaved Amplified Polymorphic Sites) marker AluICAPS

(Boualem et al. 2008) was performed according to Diaz et al. (2014) to enrich the population in andromonoecious individuals. In fact, the existence of a significant epistatic interaction between the gene andromonoecious (*a*) and a QTL controlling the FS (Fruit Shape) character has been recently demonstrated (Diaz et al. 2014). This population was maintained *in vitro* at Centre for Research in Agricultural Genomics (CRAG, Barcelona), which allowed us to perform three different experiments in two locations in Spain, Valencia and Zaragoza, in 2011 (E1 and E2, respectively), and in Zaragoza in 2012 (E3). In Zaragoza, 200 F_2 plants together with 5 plants of the F1 and each of the parents (PS and TRI) were randomised in an open field with irrigation at Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA) and allowed to open-pollinate during the summer of both years. Every experimental unit consisted of three plants that were clones of the same genotype, and three fruits were collected from each of them. In Valencia, 113 F_2 plants (all of them andromonoecious) and 5 plants of the F1 and of each parent were grown in drip-irrigated peat pots in a greenhouse at the Polytechnic University of Valencia (UPV). They were self-pollinated by hand and two fruits were collected from each of them.

Twelve traits were evaluated in the F_2 mapping population above described in the three different assays (Table 1). Whole fruits and 100 dried seeds from each fruit were weighed. The remaining characters were measured on digital images of longitudinal sections of the collected fruits using the Tomato Analyzer 3.0 software (Brewer et al. 2006). Principal component analysis (PCA) was used to determine the associations among the morphometric data (30 X and Y coordinates evenly distributed along the perimeter of the fruit section). Two components were extracted (PCA1 and PCA2) as together they accounted for more than 90% of the total variation in the dataset.

Percent mid-parental heterosis (MPH) was calculated using the data derived from a total of 5 replicates of each of the genotypes, F1 and both parents, as:

$$MPH=[(2\overline{F1}-\overline{P1}-\overline{P2})/(\overline{P1}+\overline{P2})]\times 100$$

where $\overline{F1}$, $\overline{P1}$ and $\overline{P2}$ are the mean values of the F1, PS and TRI, respectively, for the trait under study.

Pair-wise Pearson's correlation coefficient was calculated between the 12 traits with the data coming from the F_2 population in each assay independently.

All statistical analyses were performed using JMP v11.1.1 software for Windows (SAS Institute Inc., Cary, NC).

DNA marker analysis in the F_2 mapping population

DNA was isolated from young leaves according to Doyle and Doyle (1990), with slight modifications: 0.2% β-mercaptoethanol was added to the 2% CTAB buffer just before use, the washes were made with a buffer consisting of 76% ethanol and 10mM ammonium acetate, and the pellet was dissolved in 30 µl of 1×TE plus 0.2 µl of 10mg/ml RNAse A (New England Biolabs). The samples were genotyped with a total of 128 SNP-based markers (127 SNPs and 1 CAPS) chosen to cover all the genome among those available in the literature (File S1). Genotyping with SNP markers was carried out using either the Illumina Veracode technology at Center for Genomic Regulation (CGR, Barcelona, Spain), or the Sequenom MassArray iPLEX system, followed by MALDI-TOF mass spectrometry at Servicio de Investigaciones Biomédicas, Unidad Central de Investigación (University of Valencia, Valencia, Spain). Amplifications of the CAPS marker AluICAPS (Boualem et al. 2008) were performed in 25-µl volume solutions containing 20 ng of genomic DNA, 7.5 mM Tris HCl (pH 9.0), 5 mM KCl, 2 mM $(NH_4)_2SO_4$, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 µM of reverse and forward primers and 2.5 U of DNA polymerase (Biotools, Madrid, Spain). PCRs were carried out on a TC-512 thermocycler (Techne, UK) programmed with an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 62ºC for 45 s and 72°C for 1 min, plus a final elongation step at 72°C for 5 min. A 10-µl volume of the AluICAPS PCR products was digested with the restriction enzyme *Alu*I (Fermentas, Fisher Scientific, Madrid, Spain), following the supplier's recommendations, and was separated by agarose gel-electrophoresis.

Linkage map construction and QTL analysis in the F_2 population

MAPMAKER 3.0 (Lander et al. 1987) was used to construct the linkage map. The distances were calculated with the Kosambi function (Kosambi 1944) and the markers were associated with the group command at minimum LOD 3.0 as described in Diaz et al. (2014). Finally, maps were drawn with MapChart version 2.2 (Voorrips 2002).

QTL analysis was carried out for each experiment (E1, E2 and E3) independently using Windows QTL Cartographer 2.5 (Wang et al. 2007) with the composite interval mapping (CIM, Zeng 1993) procedure. The LOD score threshold for a significant level $p<0.05$ was obtained for each trait by a permutation test with 1000 resamplings.

The QTL terminology system coined by Diaz et al. (2011) was adopted. According to it, the first letters represent the trait abbreviation (Table 1), followed by a "q" that stands for QTL, a letter to name the mapping experiment ("t", in this case), a digit designating the linkage group (LG) to which the QTL maps, a dot and a final number to differentiate several QTL controlling a trait in a certain experiment on the same LG, if that were the case.

QTL-by-environment (QTL×E) interactions were tested by calculating the least square means of each trait in a two-way ANOVA using the genotypic data from the marker at each QTL peak.

Epistatic interactions were assessed with the software QTLNetwork-2.1 (Yang et al. 2008) carrying out 1000 permutations to calculate the appropriate statistic threshold. Significant interactions were double-checked by two-way ANOVA in each environment with the markers closer to the QTL peak as factors.

Advanced backcross (BC) QTL analysis

To verify QTL effects and accelerate the introgression of potentially valuable QTL, the strategy "advanced backcross QTL analysis", conceived by Tanksley and Nelson (1996), which combines the identification of QTL with the ultimate development of a variety, was adopted. The QTL analysis was carried out in several stages of the crossing program: BC1, BC2 and BC3. The populations consisted of 31, 127 and 65 individuals, respectively (File S2), grown in Valencia since 2010 to 2012. In all cases, a single F_2 plant was used to generate each family. Genotyping and map construction was carried out as described above for the F_2 population, employing 93 SNP-based markers (92 SNPs and 1 CAPS) in the case of BC1 and BC2, and 80 SNPs in the BC3 population (File S1).

QTL analysis was performed as reported above with either the composite interval mapping (CIM, Zeng 1993) or the Single-Marker Analysis (Soller and Brody, 1976) procedure.

QTL validation

QTL at LGIV

In 2014, a marker-assisted selection of TRI and PS homozygous genotypes for marker CMPSNP352, located in the QTL region affecting FA, FD, FW, FL and FS on LGIV

(data not shown), was performed among BC3S1 seedlings from the selfing of a single plant heterozygous for the target region belonging to the 12M74 BC3 family (File S2). The experimental design consisted of eight blocks with two plants for each of the two homozygous combinations at the selecting marker (14M10A for TRI and 14M10B for PS alleles), and eight blocks of parental PS too, fully randomized in a field at CITA (Zaragoza).

QTL at LGVI

The effects of the QTL located in LGVI were verified in two different experiments. In 2012, a preliminary assay was performed in a greenhouse at the UPV (Valencia, Spain). Eleven and 15 plants of two BC2S1 lines, 12M124 and 12M125 respectively (File S2), were evaluated for FA, FL, FS, FW, and PA. 12M124 and 12M125 carried the PS and TRI alleles in homozygosis for two markers in the QTL region (FR14P22 and CMPSNP107, both of them tightly linked to the peak marker PS_19-B07, data not shown) respectively.

In 2013, three full-sibling BC2S2 families were tested at CITA in Zaragoza (File S2), two of them homozygous TRI (13M1 and 13M2) and the third one homozygous PS (13M3) for the region containing the QTL in LGVI (markers CMPSNP1021 and A_38-F04 used for the assisted selection, File S1), as well as the parent PS. A randomized block design was adopted, with a total of 15 plants per family, distributed in 5 blocks of 3 plants for sibling family and 10 blocks for the control. The flowers were allowed to open-pollinate and five fruits per plot were collected and phenotypically evaluated for the traits FA, FW, FL, FS and PA.

QTL at LGVIII

In 2012, 50 plants from the BC2S1 family 12M61 (File S2), with the regions harboring some of the putative QTL to validate segregating, together with 17 plants of the PS parent, were grown in drip-irrigated peat pots in a greenhouse in Paiporta (Valencia, Spain). Open-pollination was allowed and two fruits per plant were evaluated for the traits FA, FL, FD, FW, PA and PT. The region represented by this BC2S1 family was that containing the QTL: *faqt8.1*, *flqt8.1*, *fdqt8.1*, *fwqt8.1*, *paqt8.1*, and *ptqt8.1*. The plants were genotyped with the SNP at the LOD peak in the QTL mapping experiment (CI_33-

B09) and only the plants homozygous TRI or heterozygous were selected to be compared to the PS parental.

In 2014, TRI (14M11A) and PS (14M11B) homozygous were selected from a BC3 segregating family (12M76; File S2) using the marker CI_33-B09. The assay was carried out as described above for region at LGIV at CITA (Zaragoza) in 2014 too.

The cosegregation of marker genotypes and the phenotypes was assessed by ANOVA and Tukey mean comparison tests in all the validation experiments. When applicable, the respective mean contrast analyses were performed to determine the effect of the introgression and the background.

Results

Phenotypic variation in domestication-related traits

TRI and PS showed the expected contrasting phenotypes for characters that have played a key role during the domestication process, as FA, FL, FD, FW, PA, PT and SW (Fig. 2, File S3). Even though some variations among trials were observed, the tendency in the means of those traits was consistent. The values of the traits reflecting size, like FA, FL, FD and FW, were always one or two orders of magnitude higher in PS than in TRI, and the PS fruits were more elongated (higher values of FS), with a higher pulp content (higher values of PA and PT) and contain bigger seeds (higher values of SW) than those from TRI.

The values of the F1 showed clearly in the three assays that FA and FW exhibited a consistent recessive mode of inheritance, whereas FS was inherited in dominant (or "overdominant") fashion in this population (Table 2). The remaining traits, FL, FD, PFB, DFB, PA, and PT showed additivity of the genetic effects.

The correlation analysis generated very similar results in the three assays (data not shown), especially between E2 and E3. The strongest correlations were observed between size-related parameters (Table 3), like FW and FA (r=0.99), FL and FA (r=0.98), FW and FL (r=0.95) and FW and FD (r=0.93). In this sense, PCA1 explained mainly size variation as it was clearly and substantially correlated with those same traits, FA (r=0.96), FL and FW (r=0.95), and FD (r=0.93), though it showed a moderate correlation with morphological parameters like FS (r=0.73) and with measurements of the pulp content, like PA (r=0.65). As PCA1 and FA seem to be redundant, from now on, only results on FA will be reported. The pulp-related traits (PA and PT) showed a strong correlation only

between them (r=0.91). In all these cases, the relationships were highly significant ($p \leq 0.001$).

Linkage map construction

A total of 128 SNP-based markers were incorporated into a genetic map comprised of 12 consensus linkage groups (LGs) that spanned 1,581.6 cM of Kosambi map distance (Fig. 3). The density of the map was of one marker every 12.4 cM with an average length of the interval between any two adjacent markers of 13.6 cM and the 83% of the intervals spanning less than 30 cM.

QTL mapping

A total of 27 QTL with stable effects in at least two F_2 trials were defined (Fig. 3 and Table 4): 10 related to morphological traits (4 for FL, 2 for FD, 3 for FS and 1 for DFB), located in LGII, IV, VI, VIII and XI; 12 to size characters (4 for FA, 4 for PCA1, and 4 for FW), in LGII, IV, VI and VIII; and 5 QTL controlling pulp content (3 for PA and 2 for PT) in LGV, VI and VIII. Only in the case of the QTL *faqt4.1*, a significant G×E interaction ($p<0.05$) was detected, although the QTL was detected in E1 and E2 with changes in the magnitude and dominance effects. LOD score values ranged from 2.86 (*paqt6.1* in E2) to 15.29 (*fsqt2.1* in E3). The tendency in terms of LOD score was maintained in most cases among environments or assays, with E2 being the one showing the highest values most frequently. The percentage of variation explained by the QTL (R^2) detected ranged from 3% (*flqt4.1* in E3) to 46% (*fwqt4.1* in E2). Six out these QTL were already identified in some of the BC populations, where a total of 42 QTL were detected (File S4), 3 in the BC1 (*faqt2.1*, *flqt4*.1, and *fsqt4.1*) and the remaining 3 in the BC3 (*flqt2.1*, *fdqt4.1*, and *fsqt2.1*), all located either in LGII or LGIV. In all cases, the LOD values were higher in the F_2 population than in any of the BC ones for the same QTL, unlike the R^2 values, that resulted higher for most of these common QTL in the BC populations. In some cases, LOD peaks that did not reach the LOD threshold but mapped in the same region that QTL detected in the F_2 population were observed in the BC2 (*paqt5.1*) and BC3 (*faqt6.1*, *faqt8.1*, and *fwqt8.1*) populations (data not shown).

The markers flanking the QTL defined confidence intervals that span from 19.2 (*flqt8.1*, *fdqt8.1* and *pca1qt8.1*) to 56.2 cM (*paqt5.1* and *ptqt5.1*), though 48% of them cover distances shorter than 30 cM (Fig. 3).

The additive effect (a) sign remained invariable across the different environments in all cases (Table 4). For most QTL, the TRI allele decreased the value of the respective character. Interestingly, all the QTL detected in the upper end of LGII (*faqt2.1*, *flqt2.1*, *fsqt2.1*, *fwqt2.1* and *pca1qt2.1*), with the LOD peak on gene *a*, showed a positive value additve effect, what means that the TRI allele increased the value of those traits. For QTL *dfbqt11.1*, the additive effect was negative but since the inverse transformation was carried out with the data to achieve normality, that means that the TRI allele happened to increase the value of DFB. The other transformation carried out does not alter the direction of the additive effect as consisted of a conversion to logarithmic scale of PT data. So, the other case in which a effect had a robust positive sign was in QTL controlling the pulp content in LGV (*paqt5.1* and *ptqt5.1*).

Gene action ratios ranged from -1.49 to 5.59 (Table 4). No inconsistencies were found among the gene action values coming from different assays, being either the same or similar. Approximately 66% of the cases showed an additive gene action. Another 19% of the QTL exhibited either partial or total dominance, and in the remaining 25%, the gene action consisted of overdominance.

Some epistatic interactions between QTL controlling domestication traits and other chromosomic regions have been found. The most robust were those detected between the FW QTL located in LGVI (*fwqt6.1*) and a region upstream in the same chromosome (CMPSNP571), that resulted to be significant in both assays, E2 (p=0.042) and E3 (p=0.018); and the *trans*-interaction between the FW QTL at LGVIII (*fwqt8.1*) and a genomic region located in LGIX (CMPSNP895), that was highly significant in E2 (p=0.007) and very significant in E3 (p=0.02) (File S5). In both cases, the effect of PS alleles increasing FW on the major QTL (*fwqt6.1* and *fwqt8.*1) was stronger when the epistatic QTL was homozygous for PS alleles. No epistatic interactions were detected among the main effect QTL.

QTL validation

Among the detected QTL, we decided to focus on LGs IV, VI and VIII as the effects of those QTLs were quite consistent among trials and the direction of the allelic effects were

according to the expected for a domestication QTL (i. e., the wild allele decrease the fruit morphology trait).

QTL at LGIV

When the TRI and PS alleles in homozygosity for the QTL detected in LGIV were assayed in BC3S1 families (14M10A and B, File S2), significant differences ($p<0.01$) in most characters were observed (Fig. 4), except of FS Furthermore, the plants homozygous for the PS allele (14M10B) did not show any significant difference when compared to the parent PS for any of the traits, indicating that in that advanced backcross family no other QTL involved in fruit morphology was present in the genetic background.

QTL at LGVI

Regarding the QTL found in LGVI, the preparatory assay performed in 2012, allowed us to confirm, even just visually (Fig. 5a), that the BC2S1 family, homozygous TRI for the region (12M125), produced smaller, less elongated (with lower values of the FS index) and with reduced pulp fruits when compared to the family homozygous PS (12M124) for the same QTL (Fig. 5b). The lack of significant differences for FW is probably due to the low number of measures (missing data) as FA average values showed the expected significant differences between the two homozygous families.

In 2013, the QTL were validated in the full-sibling BC2S2 families 13M1-3 (File S2 and Fig. 6a). Three contrasts were performed to address different questions: effect of the introgression in LGVI (13M1 and 13M2 *vs* 13M3), effect of genetic background in lines with the TRI introgression (13M1 *vs* 13M2) and without the TRI introgression (13M3 *vs* PS). With regard to the first question, 13M1 and 13M2 showed highly significant reduction of mean values compared with 13M3 for all the traits (Fig. 6b), verifying the QTL for FA, FL, FS, FW, and PA identified in the F_2 population. The families 13M1 and 13M2 showed differences highly significant for FA and FW ($p<0.001$), very significant for FL ($p<0.01$), and significant for PA ($p<0.05$), indicating the likely presence of other introgressions that may affect those traits, whereas non-significant differences were observed for FS. The effect of the genetic background was also evaluated comparing 13M3 with PS. Highly significant differences were observed for FA, FL and FW ($p<0.001$) but not for FS and PA. Thus, even though some genetic background effects were observed, the current results verify the previously detected QTL on LGVI.

QTL at LGVIII

The differences in all the traits under control of this region (FA, FL, FD, FW, PA and PT) among the plants carrying the TRI allele, either in homozygosis or heterozygosis, and the parent PS, were highly significant and can be observed already in the preliminary assay carried out in 2012 (Fig. 7). In all cases, no differences between the homozygous and heterozygous for the TRI alleles were observed (Fig. 7b), so the presence of only one TRI allele is enough to reduce the average value of the character.

When the two contrasting homozygous for the region (14M11A and B) were compared, differences were observed for all the traits, except for PA and PT (Fig. 8). These differences were highly significant in the case of FW ($p<0.001$) and very significant for FA, FL and FD ($p<0.01$). When the comparisons were made between the PS homozygous (14M11B) and the parent PS, the effects in FL were highly significant ($p<0.001$), while for FD and PA were very significant ($p<0.01$) and significant ($p<0.05$), respectively. However, for some traits, like FA, FW and PT, those differences were not detected between PS and the homozygous for the PS allele.

Discussion

Throughout the domestication process, a wide range of changes have altered the ancestral melon fruit form leading to the impressive phenotypic diversity of fruits nowadays known as melons. In the current work, we have focused on fruit traits to dissect them genetically and to understand the domestication process in this species. The first striking result is the different behavior of the hybrid F1 depending on the trait. Thus, MPH ranged from highly negative (FA, FW) to highly positive (FS), with low or absent in between (PFB, DFB, PA, PT). The high value of the MPH could be a reflection of the great genetic divergence between the parent genotypes as it has been recently demonstrated at a genome global level with a dendrogram constructed with the data coming from a genotyping array with 768 SNPs (Esteras et al. 2013). However, the fact that the amount of MPH is trait specific is indicating that there may be a biological reason for it. High positive MPH for FS is commonly observed in crosses between distantly related melon genotypes (Monforte et al. 2005), as it is the current case. The most plausible explanation is that alleles inducing elongation are dominant over alleles inducing roundness, i.e., dominance complementation (Fernandez-Silva et al. 2009). Traits showing low MPH or additive mode of action include those related to the pulp content (PA and PT), what may be

443 relevant from a breeding point of view for increasing the edible portion of the fruit, as
444 additive traits are easier to handle in breeding programs. On the other hand, MPH is quite
445 variable among melon hybrids, but in crosses between cultivated and wild genotypes,
446 high negative MPH has also been reported previously (Monforte et al. 2005). This last
447 observation is also common in hybrids between wild and cultivated tomatoes (Monforte
448 et al. 1997; Lippman and Tanksley 2001; Ashrafi et al. 2012; Capel et al. 2015). Fruit
449 development requires a large amount of resources from the plant. Also, small size fruit
450 favors seed dispersion, especially if it is carried out by small animals such as rodents and
451 birds. Therefore, keeping the size of the fruits small should have an evolutionary role in
452 natural populations. The high negative MPH would be one of the ways to restrict the fruit
453 dimensions, especially for outcross species as wild melons are.
454 The strongest correlations have been found among some pair of traits as FL, FW and FA,
455 what could be expected as they are components of the fruit size. Interestingly, the lack of
456 correlation between the pulp-related traits (PA and PT) and the other ones seems to reflect
457 that the fruit size and shape do not influence the relative pulp content. In theory, this
458 would allow to decrease the fruit size without being penalized in terms of pulp relative
459 area. Similarly, SW does not show any correlation with any of the other traits, nor even
460 with those related to the size of the fruit, quite the opposite to what have been suggested
461 by other authors in some melon and pumpkin cultivars (Nerson and Paris, 2000). This is
462 supported by the absence of SW QTL in the same location as fruit-size controlling QTL,
463 what could mean the genetic dissociation of these groups of traits, as it has been suggested
464 in grapevine (Doligez et al. 2013), even if domestication and artificial selection has
465 favored larger and heavier seeds in melon, as in most cultivated plants. The fact that SW
466 QTL have not been detected in the F_2 or BC populations, even when the parents show
467 marked differences in their values for this character, could be due to the existence of a
468 number of QTL with a low to moderate contribution to the phenotypic variance, what
469 would not have been detected as significant.
470 Regarding the genetic map built in this study, no changes in the chromosome assignment
471 have been observed and the relative order of the markers is also maintained in almost all
472 cases when compared to SNP maps recently developed (Diaz et al. 2014; Perpiñá et al.
473 2016) and with the melon consensus map (Diaz et al. 2011; Diaz et al. 2015). The marker
474 density and the gap length are on the order of those maps described in melon with a similar
475 number of markers (Fazza et al. 2013) but, as expected, they are lower in terms of marker
476 density than those reported in highly saturated maps (Deleu et al. 2009; Diaz et al. 2011).

Four clusters of QTLs involved in domestication have been identified, what could be a sign of linkage and/or pleiotropy. The QTL cluster on LGII is most likely due to pleiotropic effects of the gene *a* (Périn et al. 2002; Diaz et al. 2014), responsible for sex determination in melon flowers (Kenigsbuch and Cohen 1990), coding for a member of the 1-aminocyclopropane-1-carboxylic acid synthase gene family (Boualem et al. 2008), and not directly involved in the domestication processes that we are dealing with here. Therefore, only three QTL clusters have been found to be involved in domestication, in general with high R^2, suggesting that the major changes occurred during melon domestication do not have a very complex genetic control, as observed in other crops (Doebley et al. 2006).

Among the domestication traits studied here, the size is one of the most relevant. In this sense, two out of the three FW QTL putatively involved in domestication identified in this study, map in the same chromosomal region as others previously described (Table 5). Thus, *fwqt4.1* maps in the same region harbored by the line SC4-3, that carried introgressions of the exotic PI 161375 accession into PS and showed an important reduction in FW (Eduardo et al. 2007). In the case of the FW QTL found in LGVIII (*fwqt8.1*), the ILs SC8-1 and SC8-2, with introgression in the same chromosome, also reduced FW (Eduardo et al. 2007). Besides, FW effects have been associated to LGVI in the IL SC6-3 (Fernandez-Silva et al. 2010), in the same region where *fwqt6.1* has been mapped here. The strictest definition of a domestication locus implies that the domesticated allele should be fixed, or nearly fixed, in the cultivated germplasm. In melon, a more detailed QTL map and further studies are needed (Monforte et al. 2014) to be able to distinguish between domestication and diversification alleles for FW. Nevertheless, among the FW QTL detected in the current study, *fwqt8.1* has some genetic properties that make it a good candidate to be an actual domestication locus. Its genetic effects are strong (R^2~20% in F_2 populations, and it reduces FW more than 50% when introduced in the PS background, see Fig. 8), the wild allele is dominant over the cultivated allele (d/a ratio is negative in F_2, and in advanced backcross generations, homozygous TRI and heterozygous are indistinguishable regarding the FW, see Fig. 7). The lack of significant differences between the homozygous and heterozygous for the TRI allele at QTL located in LGVIII for any of the traits (Fig. 8b) suggests that the TRI allele is acting as the dominant one, though further experiments would need to be carried out to verify this point.

Our data show that the *fwqt4.1* and *fwqt8.1* are major QTL with a recessive mode of action (Table 4). Furthermore, the positions of all FW QTL found in this work agree with those reported by Monforte et al. (2014) in the meta-analysis carried out anchoring most of the size and shape QTL previously published in melon on the genome sequence. That study allows us to verify that FW QTL described here colocalize also with homeologous sequences of tomato genes controlling the shape and/or size, like SUN (in the case of *fwqt2.1* and *fwqt4.1*), OFP, CNR, and WOX (in the case of *fwqt8.1*). The *OFP* (Ovate Family Proteins) genes regulate the tomato elongation (Liu et al. 2002), whereas the *CNR* (Cell Number Regulator) family of genes controls the fruit size in tomato (Frary et al. 2000), among other species. The WOX (WUSCHEL-related homeobox) family of genes has been described to control the tomato locule number (Muños et al. 2011). In the regions containing size QTL at LGIV (*pca1qt4.1* and *fwqt4.1*), a tomato homeologous gene that regulates the fruit size has been found in the melon sequence (*CmCNR2*). Similarly, other members of the CNR family (*CmCNR4*, *6*, *7*) are located in the region at LGVI that harbors size-related QTL (*faqt6.1*, *fwqt6.1* and *pca1qt6.1*). In the case of *fwqt8.1*, homeologous genes to some that control the size in tomato have been discovered, like *Wuschel* (*CmWuschel3*) and *KLU*/*FW3.2* (*Cmfw3.2-5*), the latter regulating the cell division in tomato fruits.

In the case of FS, most of the QTL discovered in this study have been previously observed in the same or adjacent regions using different mapping populations (Table 5): *fsqt2.1* (Périn et al. 2002; Paris et al. 2008; Diaz et al. 2014), and *fsqt6.1* (Diaz et al. 2011). Other QTL for FL (*flqt4.1*) found here match those previously reported in the same location (Eduardo et al. 2007; Diaz et al. 2011). Interestingly, a few genes that have been reported to play a role in the tomato fruit shape have been identified in this same (*CmSUN10*) and neighboring region (*CmOFP4*, *5*) in the melon genome (Table 6) (Garcia-Mas et al. 2012). This is partly a reflection of those traits being the most frequently included in QTL identification studies.

Genes controlling the pulp content could have played an important role in the melon domestication history as well. QTL for traits related to pulp content, such as PT, have been found in a limited number of studies (Obando et al. 2008; Paris et al. 2008), and nonetheless in the same positions as those reported here in LGV and VIII. The QTL for PT located in LGV and LGVIII coincide with those reported by Obando et al. (2008) and Paris et al. (2008), respectively (Table 5). A QTL for relative pulp thickness has also been reported in the LGVI region (Obando et al. 2008) in which we have found a PA QTL (but

not for PT). The lack of more matches in the case of QTL for pulp content-related traits is probably due to the fact that this character has not been frequently measured in the published melon QTL studies.

For most traits, the TRI alleles decrease the value of such characters. Therefore, these QTL are good candidates for searching gene regions involved in the processes of melon domestication and evolution. Among the QTL in which the TRI alleles act increasing the value of the character are those controlling the pulp content at LGV (*paqt5.1* and *ptqt5.1*). This robust positive a effect could render direct benefits from a breeding perspective, as an increase in the edible pulp content is always a desirable feature, regardless the size of the fruit.

In many cases, genotypes at a marker locus accounted for a considerable degree of the total phenotypic variation, reaching up to 46% in the case of the one linked to *fwqt4.1*. In fact, there were only two minor QTL (R^2<10%) when all the environments are considered, *ptqt8.1* and *pca1qt4.1*. In all cases, except for *flqt4.1* and only in E3, the R^2 values were greater than 5%. This, together with the fact that the QTL×E is not significant in almost all cases, makes them good specific targets for breeding programs. In fact, the idea of using wild accessions in melon breeding is becoming more attractive since Qi et al. (2013) found that the genome-wide reduction in genetic variability caused by domestication and bottleneck is larger in some members of the *Cucurbitacea* family (i.e. watermelon and cucumber) than in some grain crops (rice, maize and soybean).

The additive mode of gene action is the most frequently observed in this study, followed by dominance (partial or complete). However, real overdominace is a very rare event. In this sense, the apparently overdominant *loci* could be the result of the linkage of several partially dominant QTL that affect the same trait.

The most robust QTL detected in all three trials were those located in the regions at LGIV, VI and VIII. Most of them were validated in advanced backcross families (BC2S1, BC2S2 and BC3). In those cases, the phenotypic differences observed in plants carrying the exotic allele were always statistically significant when compared to those that carried the PS allele or to the parental PS itself. At LGIV, only the QTL for FS (*fsqt4.1*) did not show any significant difference when the two homozygous lines (14M10, Fig. 5) were compared, what could be due to the fact that the introgression in the plant used to originate them did not harbor the whole genomic region containing such QTL.

The effects of all QTL at LGVI were verified in the validation assays, even when the effects of introgressions other than that of the QTL at LGVI were revealed by comparing

578 13M1 and 13M2. It is still valid to conclude that the differences found between 13M1-2
579 and 13M3 are mainly due to the target introgression of QTL at LGVI, as they are between
580 three and six orders of magnitude higher than the ones between 13M1 and 13M2 (data
581 not shown). This is especially true in the case of FS and PA, in which the effects of those
582 undesired introgressions are non-significant and just significant, respectively. This could
583 be indicative of the selection of the PS allele (or the alleles of other cultivars, in general)
584 in modern western melon cultivars during domestication.
585 A possible explanation for the lack of verification of the QTL *paqt8.1* and *ptqt8.1* is that
586 the genetic intervals of these two QTL are not entirely included in the families used for
587 the validation (14M11, Fig. 3). On the other hand, the homozygous TRI and the parent
588 PS showing similar mean values for some traits is due to the special environmental
589 conditions of the summer of 2014 in Zaragoza, with no particularly high temperatures,
590 what affected PS fruit set and growing, as this cultivar is quite sensitive to relatively low
591 temperatures. In any case, that did not invalidate the comparison between both contrasting
592 homozygous.
593 Currently, introgression lines of these QTL are being generated, what will predictably
594 contribute to their fine mapping and future molecular cloning.
595
596 **Conclusions**
597
598 The QTL mapped here provide insights into the genetic basis of traits under domestication
599 during melon crop evolution. The comparison of genome-wide variations of cultivated
600 and wild melon accessions as those used in this work, could shed some light on this by
601 identifying domestication sweeps hopefully in these same areas and, more than likely,
602 also in others controlling domestication-related traits not addressed here. A deeper
603 knowledge of the domestication process gains importance when it comes to choose the
604 germplasm to study. In this sense, the comparison of Asian and African wild accessions,
605 and the cultivars and landraces derived from them, would clarify some unanswered
606 questions, such as, whether several independent domestication events have occurred, and,
607 in that case, whether the same genomic regions have been the target in both domestication
608 histories.
609 Domestication and selection have reshaped and hugely altered the size of melon fruits.
610 So, the identification of the QTL responsible for it, will allow breeders to perform this

611 selection in an even more directed way. A step further consists of detecting the linkage
612 drag associated with the introgression from wild accessions in order to minimize it.
613
614 **Acknowledgments**
615
616 We thank to S Casal, A Mercader and M Amit for technical support and D. L. Goodchild
617 for reviewing the English language. This work was supported by the Spanish Ministry of
618 Economy and Competitiveness/FEDER grants AGL2012-40130-C02-02, AGL2015-
619 64625-C2-2-R to AJM, AGL2014-53398-C2-2-R to BP, AGL2015-64625-C2-1-R,
620 Centro de Excelencia Severo Ochoa 2016-2020, and the CERCA Programme/Generalitat
621 de Catalunya to JGM and MM. AD was supported by a JAE-Doc contract from CSIC.
622
623
624 **References**
625
626 Ashrafi H, Kinkade MP, Merk HL, Foolad MR (2012) Identification of novel quantitative
627 trait loci for increased lycopene content and other fruit quality traits in a tomato
628 recombinant inbred line population. Mol breeding 30:549-567
629 Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin, H., Sari MA, Collin
630 F, Flowers JM, Pitrat M, Purugganan, M. D, Dogimont C, Bendahmane A (2008) A
631 conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in
632 melons. Science 321:836-838
633 Brewer MT, Lang L, Fujimura K, Dujmovic N, Gray S, van der Knaap E (2006)
634 Development of a controlled vocabulary and software application to analyze fruit
635 shape variation in and other plant species. Plant Phys 141:15-25

636 Capel C, Fernández del Carmen A, Alba JM, Lima-Silva V, Hernández-Gras F, Salinas
637 M, Boronat A, Angosto T, Botella MA, Fernández R, Granell, A, Capel J, Lozano R
638 (2015) Wide-genome QTL mapping of fruit quality traits in a tomato RIL population
639 derived from the wild-relative species *Solanum pimpinellifolium* L. Theor Appl
640 Genet 128:2019-203
641 Cohen S, Itkin M, Yeselson Y, Tzuri G, Portnoy V, Harel-Baja R, Lev S, Sa'ar U,
642 Davidovitz-Rikanati R, Baranes N, Bar E, Wolf D, Petreikov M, Shen S, Ben-Dor S,
643 Rogachev I, Aharoni A, Ast T, Schuldiner M, Belausov E, Eshed R, Ophir R,
644 Sherman A, Frei B, Neuhaus, HE, Xu Y, Fei Z, Giovannoni J, Lewinsohn E, Tadmor
645 Y, Paris HS, Katzir N, Burger Y, Schaffer AA (2014) The *PH* gene determines fruit
646 acidity and contributed to the evolution of sweet melons. Nat Commun 5:4026
647 Deleu W, Esteras C, Roig C, Gonzalez-To, M, Fernandez-Silva, I, Gonzalez-Ibeas D,
648 Blanca J, Aranda MA, Arus P, Nuez F, Monforte, AJ, Pico MB, Garcia-Mas J (2009)
649 A set of EST-SNPs for map saturation and cultivar identification in melon. BMC
650 Plant Biol 9:90
651 Dhillon NPS, Ranjana R, Singh K, Eduardo I, Monforte AJ, Pitrat M, Dhillon NK,
652 Singh PP (2007) Diversity among landraces of Indian snapmelon (*Cucumis melo*
653 var. *momordica*). Genet Res Crop Evol 54:1267-1283
654 Diaz A, Fergany M, Formisano G, Ziarsolo P, Blanca J, Fei Z, Staub JE, Zalapa JE,
655 Cuevas HE, Dace G, Oliver M, Boissot N, Dogimont C, Pitrat M, Hofstede R, van
656 Koert, Harel-Beja R, Tzuri G, Portnoy V, Cohen S, Schaffer A, Katzir N, Xu Y,
657 Zhang H, Fukino N, Matsumoto S, Garcia-Mas J, Monforte AJ (2011) A consensus
658 linkage map for molecular markers and Quantitative Trait Loci associated with
659 economically important traits in melon (*Cucumis melo* L.). BMC Plant Biol 11:111

660 Diaz A, Forment J, Argyris JM, Fukino N, Tzuri G, Harel-Beja R, Katzir N, Garcia-Mas
661 J, Monforte AJ (2015). Anchoring the consensus ICuGI genetic map to the melon
662 (*Cucumis melo* L.) genome. Mol Breeding 35:188
663 Diaz A, Zarouri B, Fergany M, Eduardo I, Álvarez JM, Picó B, Monforte AJ (2014)
664 Mapping and introgression of QTL involved in fruit shape transgressive segregation
665 into 'Piel de Sapo' melon (*Cucumis melo* L.). Plos One 9:e104188
666 Doebley JF, Gaut BS, Smith BD (2006). The molecular genetics of crop domestication.
667 Cell 127:1309-1321
668 Doligez A, Bertrand Y, Farnos M, Grolier M, Romieu C, Esnault F, Dias S, Berger G,
669 Francois P, Pons T, Ortigosa P, Roux C, Houel C, Laucou V, Bacilieri R, Peros JP
670 (2013) New stable QTLs for berry weight do not colocalize with QTLs for seed traits
671 in cultivated grapevine (*Vitis vinifera* L.). BMC Plant Biol 13:217
672 Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13-15
673 Eduardo I, Arús P, Monforte AJ, Obando J, Fernández-Trujillo JP, Martínez JA, Alarcón
674 AL, Álvarez JM, van der Knaap E. 2007. Estimating the genetic architecture of fruit
675 quality traits in melon using a genomic library of near isogenic lines. J Am Soc Hortic
676 Sci 132:80-89
677 Esteras C, Formisano G, Roig C, Diaz A, Blanca J, García-Mas, J, Gómez-Gillamón ML,
678 López-Sesse A, Lázaro A, Monforte A, Pico B (2013) SNP genotyping in melons:
679 Genetic variation, population structure and linkage disequilibrium. Theor Appl Genet
680 126:1285-1303
681 FAO (2016) Statistics Division of Food and Agriculture Organization of the United
682 Nations (FAOSTAT). http://faostat.fao.org/

683 Fazza AC, Dallagnol LJ, Fazza AC, Monteiro CC, Lima BM, Wassano DT, Camargo
684 LEA (2013) Mapping of resistance genes to races 1, 3 and 5 of *Podosphaera xanthii*
685 in melon PI 414723. Crop Breed Appl Biot 13:349-355
686 Fernandez-Silva I, Moreno E, Eduardo I, Arús P, Álvarez JM, Monforte AJ (2009) On
687 the genetic control of heterosis for fruit shape in melon (*Cucumis melo* L.). J Hered
688 100:229-235
689 Fernandez-Silva I, Moreno E, Essafi A, Fergany M, Garcia-Mas J, Martín Hernández
690 AM, Álvarez JM, Monforte AJ (2010) Shaping melons: agronomic and genetic
691 characterization of QTLs that modify melon fruit morphology. Theor Appl Genet
692 121:931-940
693 Frary A, Nesbitt TC, Grandillo S, van der Knaap E, Cong B, Liu JP, Meller J, Elber R,
694 Alpert KB, Tanksley SD. 2000. *fw2.2*: a quantitative trait locus key to the evolution
695 of tomato fruit size. Science 289:85-88
696 Garcia-Mas J, Monforte AJ, Arús P (2004) Phylogenetic relationships among *Cucumis*
697 species based on the ribosomal internal transcribed spacer sequence and
698 microsatellite. Plant Syst Evol. 248: 191-203
699 markers
700 Harel-Beja R, Tzuri G, Portnoy V, Lotan-Pompan M, Lev S, Cohen S, Dai N, Yeselson
701 L, Meir A, Libhaber SE, Avisar E, Melame T, van Koert P, Verbakel H, Hofstede R,
702 Volpin H, Oliver M, Fougedoire A, Stalh C, Fauve J, Copes B, Fei Z, Giovannoni J,
703 Ori N, Lewinsohn E, Sherman A, Burger J, Tadmor Y, Schaffer AA, Katzir N (2010)
704 A genetic map of melon highly enriched with fruit quality QTL and EST markers,
705 including sugar and carotenoid metabolism genes. Theor Appl Genet 121:511-533
706 Kenigsbuch D, Cohen Y (1990) The inheritance of gynoecy in muskmelon. Genome
707 33:317-320

708 Kirkbride JHJr (1993) Biosystematic monograph of the genus *Cucumis* (*Cucurbitaceae*).
709 Parkway Publishers, Boone, North Carolina
710 Kosambi DD (1944) The estimation of map distances from recombination values. Ann
711 Eugen 12:172-175
712 Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits
713 using RFLP linkage maps. Genetics 121:185-199
714 Leida C, Moser C, Esteras C, Sulpice R, Lunn JE, de Langen F, Monforte AJ, Pico B
715 (2015) Variability of candidate genes, genetic structure and association with sugar
716 accumulation and climacteric behavior in a broad germplasm collection of melon
717 (*Cucumis melo* L.). BMC Genet. 16:28
718 Lippman Z, Tanksley SD (2001) Dissecting the genetic pathway to extreme fruit size in
719 tomato using a cross between the small-fruited wild species Lycopersicon
720 pimpinellifolium and L. esculentum var. Giant Heirloom. Genetics 158:413-422
721 Liu JP, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes
722 underlying the cause of pear-shaped tomato fruit. Proc Natl Acad. Sci. USA
723 99:13302-13306
724 Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication
725 and diversification. Nat Rev Genet 14:840-852
726 Monforte AJ, Asins MJ, Carbonell EA (1997) Salt tolerance in *Lycopersicon* species VI.
727 Genotype-by-salinity interaction in quantitative trait loci detection: constitutive and
728 response QTLs. Theor Appl Genet 95:706-713
729 Monforte AJ, Diaz A, Caño Delgado A, van der Knaap E (2014) The genetic basis of fruit
730 morphology in horticultural crops: lessons from tomato and melon. J Exp Bot
731 65:4625-4637

732 Monforte AJ, Eduardo I, Abad S, Arús P (2005) Inheritance mode of fruit traits in melon:
733 Heterosis for fruit shape and its correlation with genetic distance. Euphytica 144:31-
734 38
735 Muños S, Ranc N, Botton E, Bérard, A., Rolland, S., Duffé, P, Carretero Y, Le Paslier
736 MC, Delalande C, Bouzayen, Brunel, D, Causse M (2011) Increase in tomato locule
737 number is controlled by two SNPs located near WUSCHEL. Plant Physiol 156:2244-
738 2254
739 Nerson H, Paris HS (2000) Relationship between fruit size and seed size in cucurbits.
740 Cucurbit Genetics Coop Rept 23:64-67
741 Obando J, Fernández-Trujillo JP, Martínez JA, Alarcón AL, Eduardo I, Arús P, Monforte
742 AJ (2008) Identification of melon fruit quality quantitative trait loci using near-
743 isogenic lines. J Am Soc Hortic Sci 133:139-151
744 Olsen KM, Wendel JF (2013) A bountiful harvest: genomic insights into crop
745 domestication phenotypes. Annu Rev Plant Biol 64:47-70
746 Paris MK, Zalapa JE, McCreight JD, Staub JE (2008) Genetic dissection of fruit quality
747 components in melon (*Cucumis melo* L.) using a RIL population derived from exotic
748 × elite US Western Shipping germplasm. Mol Breeding 22:405-419
749 Périn C, Hagen LS, Giovinazzo N, Besombes D, Dogimont C, Pitrat M (2002) Genetic
750 control of fruit shape acts prior to anthesis in melon (*Cucumis melo* L.). Mol Genet
751 Genomics 266:933-941
752 Perpiñá G, Esteras C, Gibon Y, Monforte AJ, Picó B (2016) A new genomic library of
753 melon introgression lines in a cantaloupe genetic background for dissecting desirable
754 agronomical traits. BMC Plant Biol 16:154

Pitrat M (2012) Domestication and diversification of melon. In: Sari N, Solmaz I, Aras V (eds) Cucurbitaceae 2012, Proceedings of Xth Eucarpia meeting, Antaley (Turkey), pp 31-39

Pitrat M (2013) Phenotypic diversity in wild and cultivated melons (*Cucumis melo*). Plant Biotechnol 30:273-278

Qi J, Liu X, Shen D, Miao H, Xie B, Li X, Zeng P, Wang S, Shang Y, Gu X, Du Y, Li Y, Lin T, Yuan J, Yang X, Chen J, Chen H, Xiong X, Huang K, Fei Z, Mao L, Tian L, Städler T, Renner SS, Kamoun S, Lucas WJ, Zhang, Z, Huang S (2013) A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. Nat Genet 45:1510-1515

Ramamurthy RK, Waters BM (2015) Identification of fruit quality and morphology QTLs in melon (*Cucumis melo*) using a population derived from *flexuosus* and *cantalupensis* botanical groups. Euphytica 204:163-177

Roy A, Bal SS, Fergany M, Kaur S, Singh H, Malik AA, Singh J, Monforte AJ, Dhillon NPS (2012) Wild melon diversity in India (Punjab State). Genet Resour Crop Evol 59:755-767

Sabato D, Esteras C, Grillo O, Pico B, Bacchetta G (2015) Seeds morpho-colourimetric analysis as complementary method to molecular characterization of melon diversity. Sci. Hortic-Amsterdam 192:441-452

Sebastian P, Schaefer H, Telford IRH, Renner SS (2010) Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. Proc Natl Acad Sci USA 107:14269-14273

Soller M, Brody T (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor Appl Genet 47:35-39

780 Tanaka K, Akashi Y, Fukunaga K, Yamamoto T, Aierken Y, Nishida H, Long CL,
781 Yoshino H, Sato YI, Kato K (2013) Diversification and genetic differentiation of
782 cultivated melon inferred from sequence polymorphism in the chloroplast genome.
783 Breeding Sci 63:183-196
784 Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the
785 simultaneous discovery and transfer of valuable QTLs from unadapted germplasm
786 into elite breeding lines. Theor Appl Genet 92:191-203
787 Voorrips RE (2002) MapChart software for the graphical presentation of linkage maps
788 and QTLs. J Hered 93:77-78
789 Wang S, Basten CJ, Zeng ZB (2007) Windows QTL cartographer 2.5. Department of
790 Statistics, North Carolina State University, Raleigh, NC
791 Yang J, Hu CC, Hu H, Yu RD, Xia Z, Ye XZ and Zhu J (2008) QTLNetwork: mapping
792 and visualizing genetic architecture of complex traits in experimental populations.
793 Bioinformatics 24:721-723
794 Zalapa JE, Staub JE, McCreight JD, Chung SM, Cuevas H (2007) Detection of QTL for
795 yield-related traits using recombinant inbred lines derived from exotic and elite US
796 Western Shipping melon germplasm. Theor Appl Genet 114:1185-1201
797 Zeng ZB (1993) Theoretical basis of separation of multiple linked gene effects on
798 mapping quantitative trait loci. Proc Natl Acad Sci USA 90:10972-10976
799

800

801 **Figure legends**

802

803 **Figure 1.** Longitudinal sections of typical fruits from "Piel de Sapo" (PS) and Trigonus
804 (TRI) parent genotypes. The rule scale on the left is in cm.
805 **Figure 2.** Variability in fruit morphology domestication-related traits among parents
806 "Piel de Sapo" (a) and "Trigonus" (b); F_1 hybrid (c); and several F_2 plants (d).
807 **Figure 3.** Linkage map obtained in the F_2 population based on SNP markers (on the right).
808 Distances in cM are indicated on the left. QTL controlling domestication-related traits
809 QTL responsible for shape (open bars); size (solid bars) and pulp content (hatched bars)
810 traits are on the right of the linkage groups. Length of QTL bars represents the two-LOD
811 support intervals (from peak marker) calculated by CIM using Windows QTL
812 Cartographer 2.5. QTL nomenclature is according Diaz et al. (2011).
813

814 **Figure 4.** Validation of the effects of QTL detected in LGIV**.** a) Longitudinal sections of
815 fruits coming from homozygous TRI (14M10A) and homozygous PS (14M10B) plants
816 selected from a BC3S1 family, and the parent PS. b) Average values for the traits under
817 the control of QTL in that region (FA, FL, FD, FS, and FW) for PS and both homozygous
818 of the BC3S1 family 14M10. Mean comparison have been done between homozygous
819 (left side of each graph) and between the PS homozygous and the parent PS (right side of
820 each graph). Means within columns followed by different letters are significantly
821 different in pairwise comparisons ($p<0.05$).
822 **Figure 5.** Preliminary validation of homozygous for the region containing QTL at LGVI.
823 a) Longitudinal sections of the fruits coming from BC2S1 plants homozygous for PS
824 (12M124) and TRI (12M125). b) Average values for the traits under the control of QTL
825 in that region (FA, FL, FS, FW, and PA) in12M124 and 12M125. Means within columns
826 followed by different letters are significantly different in pairwise comparisons ($p<0.05$).
827 **Figure 6.** Phenotypic validation of QTL located at LGVI**.** a) Longitudinal sections of
828 fruits coming from the parent PS and three BC2S2 families: two homozygous TRI (13M1
829 and 13M2) and one homozygous PS (13M3). b) Average values for the traits under the
830 control of QTL in that region (FA, FL, FS, FW, and PA) in PS and 13M1, 2 and 3. Mean
831 comparisons have been done between the two TRI homozygous and the PS homozygous
832 (left side of each graph), between the two TRI homozygous (middle part of each graph),
833 and between the PS homozygous and the parent PS (right side of each graph). Means

within columns followed by different letters are significantly different in pairwise comparisons ($p<0.05$).

Figure 7. Preliminary validation of QTL in LGVIII**.** a) Longitudinal sections of representative fruits coming from the parent PS and the segregating BC2S1 family 12M61, homozygous for the TRI allele (right) and heterozygous (left). b) Average values for traits FA, FL, FD, FW, PA, and PT in PS and the homozygous TRI and heterozygous of the BC2S1 family 12M61 plants. Means within columns followed by different letters are significantly different in pairwise comparisons ($p<0.05$).

Figure 8. Validation of the effects of the QTL detected in LGVIII. a) Longitudinal sections of representative fruits from the parent PS and the two BC3 families: the homozygous for the TRI allele (14M11A), and the homozygous for the PS allele (14M11B). b) Average values for the traits under the control of QTL in that region (FA, FL, FD, FW, PA, and PT) in PS, 14M11A and 14M11B. Mean comparison have been done between both homozygous (left side of each graph) and between the PS homozygous and the parent PS (right side of each graph). Means within columns followed by different letters are significantly different in pairwise comparisons ($p<0.05$).

Table captions

Table 1. Description and abbreviation of the traits under study in the F_2 population derived from the cross PS×TRI.

Table 2. Percent mid-parental heterosis (%MPH) for the traits measured in the parents (PS and TRI) and the F_1 coming from the cross PS×TRI in three different field assays (E1: Valencia 2011; E2: Zaragoza 2011; E3: Zaragoza 2012).

Table 3. Linear correlation (Pearson) between the 12 traits evaluated in an F_2 population derived from the cross PS×TRI in each assay independently (E1: Valencia, 2011; E2: Zaragoza, 2011; E3: Zaragoza, 2012), though for clarity purposes, only data from E3 are shown.

Table 4. QTL for domestication-related traits identified in an F_2 population coming from the cross PS×TRI in three different experiments: E1 (Valencia, 2011); E2 (Zaragoza, 2011) and E3 (Zaragoza, 2012).

Table 5. Putative homologous QTL to those found in this study.

Table 6. Genes found in the regions containing the QTL identified in the PS×TRI F_2 population.

868 **Supplemental Files**

869 **File S1**. Markers used in the mapping of the F_2 population and the QTL verification
870 experiments.

871

872 **File S2.** Genealogical diagram with all the families used in the mapping, QTL
873 identification and validation experiments.

874 **File S3.** Mean values for the traits measured in the parents (PS and TRI), the F1 and the
875 F_2 population coming from the cross PS×TRI in three different field assays (E1: Valencia
876 2011; E2: Zaragoza 2011; E3: Zaragoza 2012).

877 **File S4.** QTL identified in the backcross populations (BC1, BC2, and BC3).

878 **File S5.** Verification of significant epistatic interactions between major FW QTL and
879 other genomic regions by two-way ANOVA. a) *fwqt6.1* (AI_19-F11)×CMPSNP571 and
880 vice versa, both markers located in LGVI, for two assays, E1: Valencia 2011 (a.1) and
881 E3: Zaragoza 2012 (a.2); b) *fwqt8.1* (CI_33-B09)×CMPSNP895 and *viceversa*, CI_33-
882 B09 in LGVIII and CMPSNP895 in LGIX, for two assays, E1: Valencia 2011 (b.1) and
883 E3: Zaragoza 2012 (b.2). All the epistatic interactions were significant ($p<0.05$), except
884 for the one represented in b.1, which was very significant ($p<0.01$).

885