Theor Appl Genet
DOI 10.1007/s00122-012-1938-z

ORIGINAL PAPER

Comparative mapping in watermelon [*Citrullus lanatus* (Thunb.) Matsum. et Nakai]

Katherine Sandlin · Jason Prothro · Adam Heesacker · Nelly Khalilian · Rebecca Okashah · Wenwen Xiang · Eleni Bachlava · David G. Caldwell · Chris A. Taylor · Danelle K. Seymour · Victoria White · Eva Chan · Greg Tolla · Cathy White · Dolores Safran · Elaine Graham · Steven Knapp · Cecilia McGregor

Received: 2 February 2012 / Accepted: 3 July 2012

Abstract The first single-nucleotide polymorphism (SNP) maps for watermelon [*Citrullus lanatus* (Thunb.) Matsum. et Nakai] were constructed and compared. Three populations were developed from crosses between two elite cultivars, Klondike Black Seeded × New Hampshire Midget (KBS × NHM), an elite cultivar and wild egusi accession, Strain II × PI 560023 (SII × Egusi) and an elite cultivar and a wild citron accession, ZWRM50 × PI 244019 (ZWRM × Citroides). The SII × Egusi and ZWRM × Citroides F_2 populations consisted of 187 and 182 individuals respectively while the KBS × NHM recombinant inbred line (RIL) population consisted of 164 lines. The length of the genetic maps were 1,438, 1,514 and 1,144 cM with average marker distances of 3.8, 4.2, and 3.4 cM for the KBS × NHM, SII × Egusi and ZWRM × Citroides populations, respectively. Shared markers were used to align the three maps so that the linkage groups (LGs) represented the 11 chromosomes of the species. Marker segregation distortion were observed in all three populations, but was highest (12.7 %) in the ZWRM × Citroides population, where Citroides alleles were favored. The three maps were used to construct a consensus map containing 378 SNP markers with an average distance of 5.1 cM between markers. Phenotypic data was collected for fruit weight (FWT), fruit length (FL), fruit width (FWD), fruit shape index (FSI), rind thickness (RTH) and Brix (BRX) and analyzed for quantitative trait loci (QTL) associated with these traits. A total of 40 QTL were identified in the three populations, including major QTL for fruit size and shape that were stable across genetic backgrounds and environments. The present study reports the first SNP maps for *Citrullus* and the first map constructed using two elite parents. We also report the first stable QTL associated with fruit size and shape in *Citrullus lanatus*. These maps, QTL and SNPs should be useful for the watermelon community and represent a significant step towards the potential use of molecular tools in watermelon breeding.

Communicated by M. Havey.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-1938-z) contains supplementary material, which is available to authorized users.

K. Sandlin · J. Prothro · A. Heesacker · N. Khalilian · R. Okashah · C. A. Taylor · S. Knapp
Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, GA 30602, USA

A. Heesacker
Crop and Soil Science Department, Oregon State University, Corvallis, OR 97331, USA

W. Xiang · E. Bachlava · D. K. Seymour · V. White · E. Chan · G. Tolla · C. White · D. Safran · E. Graham · S. Knapp
Monsanto, Woodland, CA 95695, USA

D. G. Caldwell · C. A. Taylor
Monsanto, Chesterfield, MO 63017, USA

D. K. Seymour
Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

C. McGregor (✉)
Department of Horticulture and Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, GA 30602, USA
e-mail: cmcgre1@uga.edu

Introduction

Watermelon [*Citrullus lanatus* (Thunb.) Matsum. et Nakai] was responsible for approximately 9.6 % of the total world vegetable tonnage produced in 2009 (Food and Agriculture

Published online: 09 August 2012

Organization of the United Nations 2011). The overall value of vegetable fresh market production in the United States was estimated to be about $24.5 billion in 2010 with watermelon being responsible for 2 % ($0.49 billion; compare tomatoes value ~ $1.4 billion, or 5.6 %) (United States Department of Agriculture: National Agricultural Statistics Service 2011). In the United States, consumption of watermelon has increased by 37 % from 1980 levels to ~7.2 kg per capita, mainly due to the popularity of seedless cultivars (Wehner 2008).

Watermelon is a member of the Cucurbitaceae family that includes many economically important domesticated species such as cucumber (*Cucumis sativus* L.), melon (*Cucumis melo* L.), squash and pumpkin (including *Cucurbita pepo*, *C. moschata* and *C. maxima*). The genus *Citrullus* (2n = 2x = 22) includes two annual species, *C. lanatus* and *C. rehmii* De Winter, and two perennial species, *C. ecirrhosus* Cogn., *C. colocynthis* (L.) Schrad (Robinson and Decker-Walters 1997). However this classification is not clear-cut and misclassifications occur often (Jarret et al. 1997).

C. lanatus is botanically divided into *C. lanatus* var. *lanatus,* which includes the elite cultivars, and *C. lanatus* var. *citroides (*citron types) (Robinson and Decker-Walters 1997). Egusi type watermelon, cultivated in West Africa for its oil and protein rich seeds were previously classified as *C. lanatus* subsp. *mucosospermus* (Fursa 1972) and sometimes as *C. colocynthis* (Jarret et al. 1997; Robinson and Decker-Walters 1997; Wehner 2008). However, both chloroplast (Dane and Lang 2004; Dane and Liu 2007) and nuclear (Che et al. 2003; Jarret et al. 1997; Nimmakayala et al. 2010) genetic studies clearly identify it as *C. lanatus* var. *lanatus*. Within *C. lanatus*, var. *citroides* showed more genetic diversity than var. *lanatus* (Dane and Lang 2004; Dane and Liu 2007; Levi and Thomas 2005). Egusi types and *citroides* types contain potentiality valuable traits, such as disease resistance, which would be useful for improving elite watermelon varieties (Hashizume et al. 2003; Levi et al. 2001b; McGregor 2011).

The application of marker assisted selection (MAS) in watermelon breeding programs has been limited by a lack of high-throughput DNA markers and genetic mapping information (Levi et al. 2002, 2006). Due to the limited observed marker polymorphisms within *C. lanatus*, and particularly among elite watermelon cultivars (Levi et al. 2001a), mapping studies have focused exclusively on interspecific or inter-subspecific crosses, and no molecular map using two elite parents have been published. The first attempt to construct a genetic map for watermelon was by Navot et al. (1990) using isozyme markers. The map was developed using an interspecific backcross population derived from *C. lanatus* × *C. colocynthis* and contained seven linkage groups. Hashizume et al. (1996) constructed a linkage map (524 cM; 62 markers) spanning 11 linkage groups (LGs) using mainly RAPD markers in a backcross population of a cultivated Japanese C. *lanatus* var. *lanatus* line (H-7) and a non-sweet wild race from South Africa (SA-1, presumably *C. lanatus* var. *citroides*). F_2 and BC_1 populations of the same parents were later used to create two more comprehensive maps using 477 isozymes, RAPD, RFLP, and ISSR markers (Hashizume et al. 2003). The BC_1 population in the latter study used the elite H-7 as the recurrent parent while earlier BC_1 map (Hashizume et al. 1996) was constructed using the wild SA-1 as recurrent parent. The final F_2 map had a length of 2,384 cM, an average interval length of 4.3 cM and 11 linkage groups (LGs), corresponding to the haploid chromosome number in watermelon. However, the distances between markers in some areas were greater than 30 cM. The map produced for the BC_1 population was constructed using markers shown to segregate in the F_2 population and had a length of 1,729 cM with an average marker distance of 7.2 cM. In the original study RAPD markers linked to rind color and flesh color were identified (Hashizume et al. 1996), while the latter BC_1 population was used to identify QTL for four horticulturally important traits, rind hardness, flesh color, rind color and Brix (Hashizume et al. 2003). This was the only mapping study to date that identified QTL linked to traits in watermelon.

Hawkins et al. (2001) used F_2 and F_3 populations, also derived from a cross between a wild *C. lanatus* var. *citroides* (PI 296341) accession and New Hampshire Midget (*C. lanatus* var. *lanatus*) to construct two and five linkage groups consisting of 26 and 13 RAPD markers, respectively. This population was known to be segregating for resistance to fusarium wilt, an economically significant disease in watermelon production (Wehner 2008). However no markers tightly linked to disease resistance were identified. The only two studies thus far to develop recombinant inbred lines (RILs) (Ren et al. 2012; Zhang et al. 2004) similarly used an inter-subspecific cross between *C. lanatus* var. *citroides* (PI 296341) and elite *C. lanatus* var. *lanatus* inbred lines. The Ren et al. (2012) map included 698 markers (SSR, InDel, and structure variation markers) and had a length of ~800 cM and an average distance of 0.8 cM between markers, on 11 linkage groups.

Marker segregation distortion was common in inter-subspecific *C. lanatus* crosses to date and ranged from 11 to 48 % (Hashizume et al. 1996, 2003; Hawkins et al. 2001; Ren et al. 2012; Zhang et al. 2004). In an attempt to control some of the segregation distortion found in previous studies, Levi et al. (2002) developed a test cross population of (*C. lanatus* var. *citroides* × *C. lanatus* var. *lanatus*) × *C. colocynthis.* This map was later extended from the 205 to 360 markers (Levi et al. 2006) however, segregation distortion remained high, especially for AFLP

and SRAP markers (Levi et al. 2006). The limited amount of genetic diversity in *C. lanatus* var. *lanatus* using older marker technologies has hampered progress. However, the advent of single nucleotide polymorphism (SNP) technology (Henry 2008; Kole and Abbott 2008) now gives us the tools to create maps that were previously not feasible. SNP markers have routinely been used in agricultural breeding programs; in plant and animal variation studies, genome mapping, and association mapping (Kole and Abbott 2008). SNP markers have already proven useful in increasing marker resolution in melon (Deleu et al. 2009) and *Cucurbita pepo* (Esteras et al. 2012; Zraidi et al. 2007). The technology lends itself to automation which makes high-throughput mapping possible and enables comparative mapping across several populations. Typically, linkage maps are limited in their usefulness to the genetic background which they represent, but this can be overcome by comparing maps with varying genetic backgrounds. Consensus maps for Cucurbitaceae have been created in *Cucumis melo* (Perin et al. 2002) and *Cucurbita pepo* (Zraidi et al. 2007). Comparative mapping would be valuable in the *Citrullus lanatus* species, both to overcome the lack of coverage and for the comparison of QTL identified in different genetic backgrounds. The only previous comparative mapping effort in watermelon was between 41 linkage groups of a testcross population [(*C. lanatus var. citroides* × *C. lanatus var. lanatus*) × *C. colocynthis*] and an inter-subspecific BC_2F_2 population with 51 linkage group (Levi et al. 2011). No trait loci were mapped.

Breeding efforts in watermelon have largely concentrated on fruit quality and morphological characteristics. These characteristics include, but are not limited to fruit size and shape, sugar content, flesh color, and rind patterns (Maynard 2001; Wehner 2008). Investigations of the inheritance of fruit morphology and quality traits, date back as far as the 1930 s (Porter 1933, 1937). Since then, many efforts have been made to better understand traits associated with watermelon fruit quality and morphology and many genes have been described that control internal fruit quality and morphology in watermelon (Guner and Wehner 2004). However, no genes have been identified that are responsible for fruit size. This trait has recently become an important consideration for breeders due to increased consumer preference for smaller sized watermelons (Gusmini and Wehner 2007). Watermelon fruit shape and rind thickness also play a role in consumers choice. Consumers usually have a preference for a specific shape, and rind thickness must be maintained as a small percentage of the fruit diameter. Small watermelon fruit must have a very thin rind while larger fruited watermelon has a thicker rind. The thicker rind of the large fruited watermelons help protect the fruit during post harvest handling and shipping (Wehner et al. 2001). An internal fruit characteristic that has received a lot of commercial attention is the Brix value of the fruit flesh. Degrees Brix is a measure of the total soluble solids in watermelon and is highly correlated with the percent sugar (Hashizume et al. 2003; MacGillivray 1947; Maynard 2001). Hashizume et al. (2003) has mapped a single QTL that accounts for 19 % of the variation in Brix in an inter-subspecific BC_1 population. However Brix is a quantitative trait thought to be controlled by several genes and significantly influenced by the environment (Gusmini and Wehner 2005; Hashizume et al. 2003; Porter et al. 1940).

We aim to use SNP markers to create genetic maps for watermelon using elite × elite, elite × egusi and elite × *citroides* populations and to map QTL associated with horticulturally important traits related to fruit size, shape, rind thickness and sugar content in an effort to identify QTL that are stable across populations.

Materials and methods

Plant material

Seed of four elite cultivars (*C. lanatus* var. *lanatus*), Klondike Black Seeded (KBS; PI 635609), New Hampshire Midget (NHM; PI 635617), Strain II (SII; PI 279461) from Japan and ZWRM50 (ZWRM; PI 593359) from China as well as a wild Nigerian egusi type (Egusi; PI 560023) (*C. lanatus* var. *lanatus*) and a wild *C. lanatus* var. *citroides* accession from South Africa (Citroides; PI 244019) were obtained from the United States Department of Agriculture-Agricultural Research Service (USDA-ARS) germplasm collection (Plant Genetic Resources Conservation Unit, Southern Regional Plant Introduction Station, USDA-ARS, Griffin, GA). Three populations were developed; an elite × elite (KBS × NHM) RIL population by single seed descent, an elite × egusi (SII × Egusi) F_2 population and an elite × *citroides* (ZWRM × Citroides) F_2 population. Leaf tissue from the three populations and respective parents were frozen, ground and lyophilized for storage. DNA was extracted using a modified CTAB procedure (Murray and Thompson 1980). The DNA was quantified using the Quant-iT Picogreen DNA reagents kit (Invitrogen, Ltd. Paisley, PA) and normalized to 50 ng/ml for subsequent genotyping.

Single-nucleotide polymorphism (SNP) genotyping

Single-nucleotide polymorphisms assays were identified and subsequently validated at Monsanto (Woodland, CA). Using 454 Sequencing (454 Life Sciences, Roche, Brandford, CT, USA) of elite inbreds that represent the world wide germplasm diversity, SNPs were selected from contigs that were enriched to be single copy regions of the genome. Any contig sequences that were low complexity

or shared 90 % identity to previously characterized repetitive sequences were discarded from further analysis. A single polymorphism was selected per contig and when possible, SNPs segregating as common variation (>5 % allele frequency) were chosen and submitted to Illumina for GoldenGate Assay design (Illumina, San Diego, CA, USA). All SNPs passing minimum design criteria from Illumina were manufactured and validated using the following steps. The 192 samples used for validation consisted of DNA from the SNP discovery as well as independent inbred lines that comprehensively represent world wide germplasm diversity. Any assay that showed complete failure to cluster, excessive compressed clustering, greater than 20 % missing data or greater than 10 % residual heterozygosity in known inbred samples were removed from further analysis. The parents and progeny for the 3 mapping populations were genotyped at Monsanto (St Louis, MO, USA) using an Illumina's GoldenGate® SNP array and BeadStudio® software (Illumina, San Diego, CA, USA). The resultant genotypes for each mapping population were phased for its respective parental genotype and transformed into a locus file.

Map construction, comparative mapping and consensus map

Polymorphic markers were identified for each segregating population among the SNPs genotyped using the GoldenGate array. Markers with segregation distortion ($P < 0.0001$) and excess of missing data (> 50 % datapoints) were removed. Linkage maps were constructed using JoinMap version 4 (Van Ooijen 2006) using regression mapping with default parameters and Kosambi map units and included 388 public SNP markers (EMS 1) as well as 346 propriotary markers, added in order to obtain denser scaffolds. Independence LOD and maximum likelihood were used as grouping method and mapping algorithm, respectively. Default parameters were used with the exception of map building, for which spatial sampling thresholds were changed to 0.1, 0.05, 0.01, 0.005, and 0.001. Linkage groups for each population were drawn using MapChart 2.2 (Voorrips 2002) and LGs in the three populations were visually aligned using shared markers (EMS 2). Chi-square tests for segregation distortion ($P < 0.05$) of mapped markers were calculated during QTL mapping using Windows QTL Cartographer 2.5 (WinQTLCart) (Wang et al. 2011).

A consensus map of public markers for *C. lanatus* was created using BioMERCATOR 2.1 (Arcade et al. 2004) by projecting the SII × Egusi and ZWRM × Citroides onto the KBS × NHM map. In cases where the marker order between maps was inverted, the order of the latter map was used in the consensus map.

Trait evaluation

The KBS × NHM F_7 RIL population (162 lines) and parents were planted in the field in 2010 at the University of Georgia's Plant Science farm in Watkinsville, Georgia (KBS × NHM-GA) and Monsanto's field trial facilities in Woodland, CA (KBS × NHM-CA). At both locations each RIL was represented by a single replicate of eight plants. At the GA location between 1 and 8 fruit were harvested and data collected from two mature fruit (depending on availability) were averaged, while at the CA location, three fruit were harvested and data was collected from a single fruit. The rational for harvesting a number of fruit, but only collecting data from one or two fruit respectively was that harvested fruit were cut open and flesh inspected visually to select fruit at optimum market maturity for phenotyping.

The SII × Egusi F_2 population of 214 plants and parents were planted at the University of Georgia's Plant Science Farm in Watkinsville, GA in the summer of 2007. Two hundred individual ZWRM × Citroides F_2 seeds were planted in the greenhouse at the University of Georgia's campus in Athens, GA during the summer of 2007. One mature fruit was collected and phenotyped for multiple traits in these populations.

Fruit weight (FWT) was measured in kilograms at maturity. Fruit length (FL) was measured in centimeters as the distance between the fruit apex and the point at which the pedicel was attached to the fruit. Fruit width (FWD) was measured in centimeters at the widest part of the fruit as the distance between each edge of the fruit. Rind thickness (RTH) was measured with a digital caliper (Balkamp Manufacturing Corp., Indianapolis, Indiana) in the middle of the fruit, half way between the apex and the pedicel, while degrees Brix (BRX) was measured using a refractometer (Atago Co., Ltd., Tokyo, Japan) from a sample of juice collected from the center of each watermelon. Pearson Correlations and the Shapiro–Wilk test for normality (Shapiro and Wilk 1965) were carried out using JMP 8.0.2 (JMP Version 2009).

Heritability on RIL-mean base was calculated as $h^2 = [\sigma^2_{RIL}/(\sigma^2_{RIL} + \sigma^2_{ENV \times RIL}/r)]$ for the combined environments (Holland et al. 2003; Nyquist and Baker 1991) where σ^2_{RIL} equaled the genetic variance among the genotypes, $\sigma^2_{RIL \times ENV}$ the variance of genotypes by environments interaction, and r the number of environments. Because the data were unbalanced, the value for r was computed as the harmonic mean of the number of environments (Holland et al. 2003).

QTL analysis

QTL detection and analysis was performed using WinQTLCart 2.5 (Wang et al. 2011). The KBS × NHM data

collected at the two locations were maintained as separate sets of data for QTL analysis and all data were analyzed by composite interval mapping (CIM) (Zeng 1993, 1994). The threshold values for all traits were calculated through permutation tests (1,000 permutations, $\alpha = 0.05$) (Churchill and Doerge 1994; Doerge and Churchill 1996). CIM analysis was performed using the standard model (Model 6) with a walk speed of 1 cM and 5 marker cofactors determined by backwards regression. The cofactors within 10 cM on either size to the QTL were excluded from the model. QTL were considered minor, intermediate and major if R^2 was < 10 %, between 10 and 25 % and more than 25 % respectively.

In order to visualize the QTL on the consensus map, the QTL identified in the different populations were projected onto the consensus map through homothetic projection using BioMercator 2.1 (Arcade et al. 2004). Further meta-analysis of QTL was not attempted since the number of independent QTL per trait does not satisfy the 10–40 QTL per trait in < 200 cM criteria (Arcade et al. 2004).

Results

Genetic maps

Three genetic maps were created, an F_6 RIL map for KBS × NHM, and F_2 maps for SII × Egusi and ZWRM × Citroides using 164, 187 and 182 individuals respectively (Table 1). The KBS × NHM RIL map contained 378 SNP markers with a 3.8 cM average distance between markers and a total length of 1,438 cM (Table 1, EMS 2). The SII × Egusi F_2 map had a total length of 1,514 cM with an average distance of 4.2 cM between the 357 markers on the map (Table 1, EMS 2). The shortest total map distance of 1,144 cM was obtained for the ZWRM × Citroides map (Table 1, EMS 2). This map also had the fewest markers, 338 and the smallest average distance between markers (3.4 cM). However, the latter map had the largest gap between two markers (33.04 cM) compared to 22.48 and 27.3 cM for the KBS × NHM and SII × Egusi maps respectively. The number of linkage groups (LGs) for the KBS × NHM map was 13, 14 for the SII × Egusi map, and 16 for the ZWRM × Citroides map. The number of mapped markers that showed segregation distortion varied from 2.8 % for the SII × Egusi map and 12.7 % for the ZWRM × Citroides map. In the ZWRM × Citroides population the distorted markers showed significant over representation of the introgression of Citroides alleles (30/43) ($\chi^2 = 6.7$, $df = 1$, $P < 0.01$). The distorted markers in the KBS × NHM (22/29) ($\chi^2 = 5.8$, $df = 1$, $P < 0.05$) showed preferential introgression of KBS alleles, while the SII × Egusi map showed preferential introgression of Egusi alleles. (9/10) ($\chi^2 = 5.4$, $df = 1$, $P < 0.025$) (1 case of overrepresentation of heterozygotes). In cases where distorted markers were clustered together, alleles from a specific parent were consistently favored at all loci in the cluster (EMS 2).

Comparative mapping and consensus map

Fifty-five of the 734 SNP markers mapped in this study were mapped in all three populations, while 450 were unique to a particular population (ESM 2). Two-hundred-and-twenty-nine markers were shared between two populations. LG 10 was the only LG that did not have at least one marker shared among all three populations. The LGs of the three populations were aligned based on the presence of shared markers (ESM 2). This yielded 11 *Citrullus lanatus* linkage groups, presumably representing the 11 haploid chromosomes of the species. LGs in the three populations were numbered 1–11, with letters ("A" or "B") used to indicate when more than one LG was thought to be from the same chromosome (ESM 2). Two LGs from the ZWRM × Citroides population did not share markers with any other LGs and could therefore not be aligned. These LGs, designated LG 12 and LG 13 contained eight and two

Table 1 Summary of the SNP maps created for the three *C. lanatus* populations

Population	Klondike Black Seeded (KBS) × New Hampshire Midget (NHM)	Strain II × Egusi (PI 560023)	ZWRM 50 × Citroides (PI 244019)
Generation	F_6 RIL	F_2	F_2
Population size	164	187	182
Total Number of Markers	378	357	338
Map Length (cM)	1,438	1,514	1,144
Averaged Distance Between Markers (cM)	3.8	4.2	3.4
Largest Gap (cM)	22.48	27.3	33.04
Number of Linkage Groups	13	14	16
% Segregation Distortion	7.7	2.8	12.7

markers respectively. The order of the shared markers was largely conserved, but rearrangements were observed on LGs 1, 2, 3, 5, 6, and 7. The majority of these rearrangements involved a single pair of adjacent markers that changed positions between the populations (ESM 2).

The projected consensus map of *C. lanatus* had a total length of 1,917.6 cM and included 378 SNP markers (Fig. 1). The average distance between markers was 5.1 cM and the largest gap of 24.4 cM was found on LG 9. ZWRM × Citroides LG 11B could not be projected onto

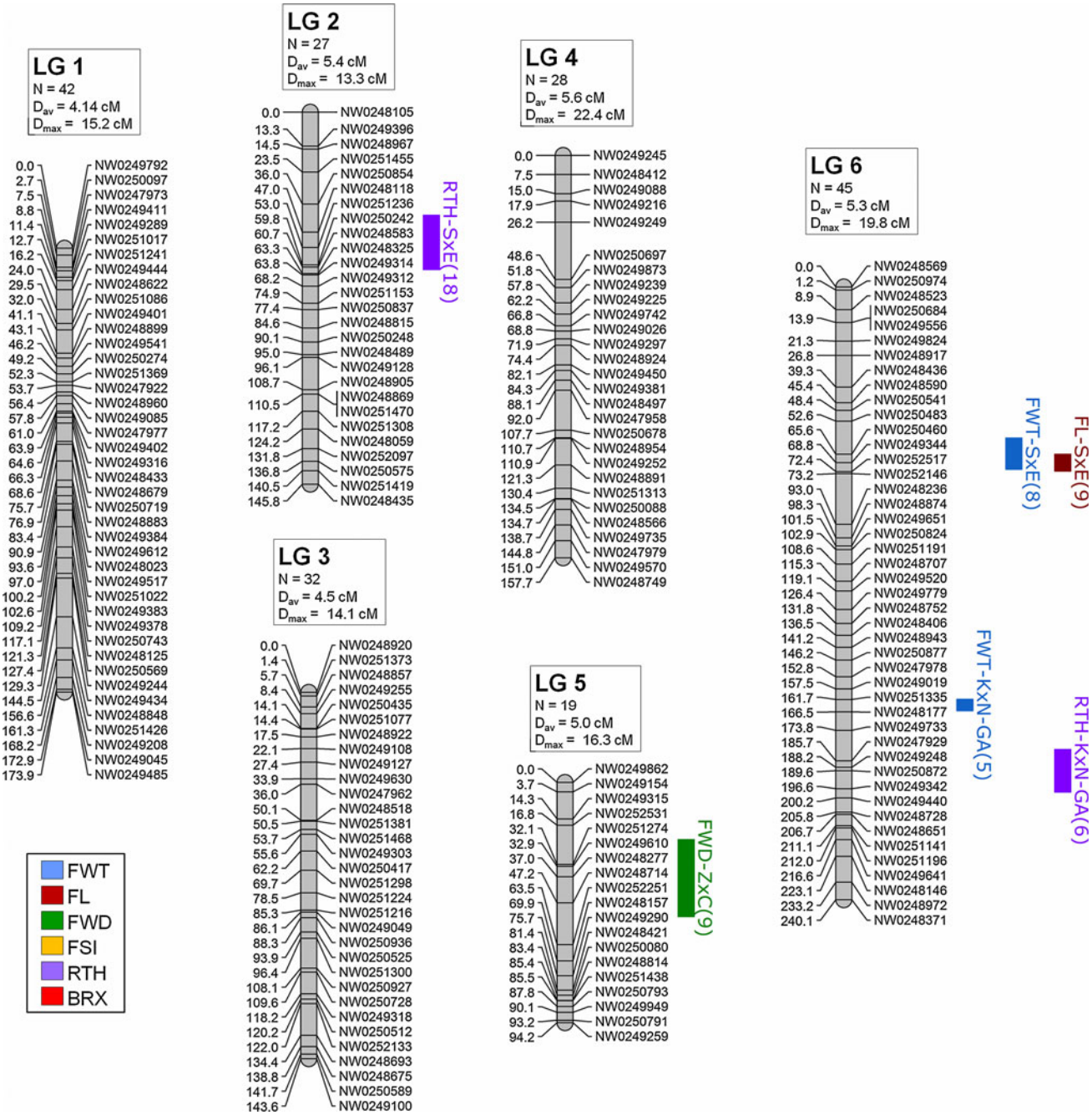

Fig. 1 Consensus SNP map of *Citrullus lanatus*. The average (D_{av}) and maximum distance (D_{max}) between markers for each linkage group are indicated. Projected location of QTL associated with fruit weight (FWT *blue*), fruit length (FL *brown*), fruit width (FWD *green*), fruit shape index (FSI *orange*), rind thickness (RTH *purple*) and Brix (BRX *red*) are indicated as *bars*. The length of the *bars* is equal to the projected 1-LOD support intervals. KxN, SxE and ZxC indicate that the QTL was identified in the Klondike Black Seeded × New Hampshire Midget, the Strain II × Egusi and ZWRM × Citroides populations, respectively. The *number* in *parenthesis* following the QTL name and population represents the percentage of phenotypic variance explained by the QTL (R^2) (color figure online)

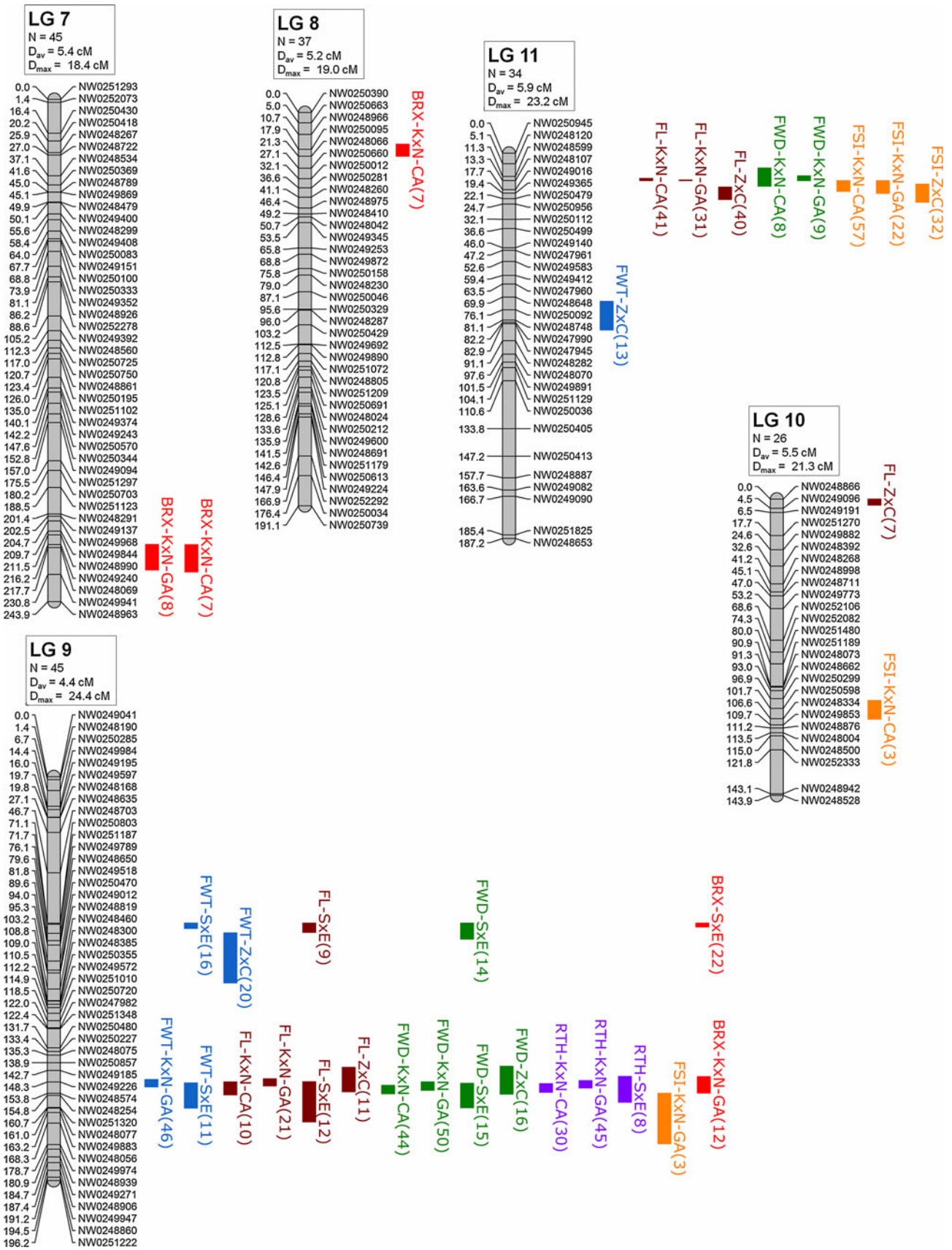

Fig. 1 continued

the consensus map since it had only one common marker with other LGs (EMS 2).

Trait data

Frequency distributions of the fruit weigh (FWT), fruit length (FL), fruit width (FWD), Fruit Shape Index (FSI), rind thickness (RTH) and Brix (BRX) were often found to be non-normal (Fig. 2). QTL analysis was carried out on non-transformed data for all traits, as well as log transformed [$(\log_{10}(x+1)$] for traits that showed non-normal distribution. Transgressive segregation in one or both directions was common (Fig. 2). Significant correlations ($P < 0.05$) were observed between FL, FWD, FSI and RTH in the two locations for the KBS × NHM (Table 2a) population. BRX was not significantly correlated between locations. In all three populations FWT, FL, FWD and RTH were significantly correlated, and as expected FSI is significantly positively correlated with FL and negatively correlated with FWD. However, FSI was not correlated with FWT. BRX at the CA location (KBS × NHM-CA) was not significantly correlated with any other traits at either location, but at the GA location (KBS × NHM-GA) BRX was positively correlated with FWT, FL, FWD and RTH at the same location and to FWD, FSI and RTH in the CA location. In the SII × Egusi populations FSI was not correlated with any other traits and RTH and BRX was also not correlated. FSI and FWT, BRX, and RTH were the only uncorrelated traits in the ZWRM × Citroides population.

Heritabilities (h^2) in the KBS × NHM population was $h^2_{FL} = 0.85$, $h^2_{FWD} = 0.87$, $h^2_{FSI} = 0.88$, $h^2_{RTH} = 0.76$ and $h^2_{BRX} = 0.20$ for the five traits measured at both locations.

QTL mapping

Similar results were obtained for non-transformed and transformed [$(\log_{10}(x+1)$] data and for the sake of brevity only the results from non-transformed data are presented. A total of 40 QTL were identified for the six measured traits in the three populations (Table 3; Fig. 1). Twenty-one were identified in the KBS × NHM population (2 locations), 11 in the SII × Egusi and 8 in the ZWRM × Citroides population. The LOD scores varied from 3.0 to 30.8 (avg. = 9.9) and the R^2 from 2.8 to 56.6 % (avg. = 18.8). The 1-LOD support interval ranges from 0.2 to 24.5 cM with an average of 8.5 cM.

Between two and four QTL were detected in the KBS × NHM population for FWT (R^2: 5.2–45.7 %), FL (R^2: 10.4–40.8 %), FWD (R^2: 8.7–50.0 %), FSI (R^2: 2.8–56.6 %), RTH (R^2: 6.4–45.0 %) and BRX (R^2: 7.0–12.1 %) across the two locations (Table 3; Fig. 1). On LG 9 QTL for FL, FWD and RTH at both locations co-localized with QTL for FWT and BRX at the GA location. QTL associated with FL, FWD and FSI at both locations co-localized in the same genomic region on LG 11. The QTL on LG 7B was the only BRX QTL detected at both the field locations.

In the SII × Egusi population, three QTL were detected for FWT (R^2: 8.2–15.6 %) and FL (R^2: 8.5–12.1 %), two for FWD (R^2: 14.1–14.6 %) and RTH (R^2: 8.4–17.7 %) and one for BRX ($R^2 = 21.6$ %). No QTLs were detected for FSI in this population (Table 3; Fig. 1). QTL for FWT and FL co-localized at three genomic regions, one on LG 6 and two on LG 9B. FWD QTL were also detected in the same two regions of LG 9B (but not LG 6).

Three QTL for FL (R^2: 7.3–39.2) and two for FWT (R^2: 13.2–20.1) and FWD (R^2: 9.2–16.0) and one for FSI (R^2: 31.8) were identified in the ZWRM × Citroides inter-subspecific population. QTL for FL and FWD co-localize on LG 9 but a QTL for FWT on the same LG was located at a different region. On LG 11 QTL for FL and FSI were detected in the same region, while FWT were located in a separate region on the same LG.

Comparative mapping of QTL

Several QTL for traits related to fruit size co-localized within populations as well as across populations (Table 3; Fig. 1). On LG 9, QTL were mapped to two general locations, one in the vicinity of universal marker NW0248650 and the other close to universal marker NW0249185. In the SII × Egusi population FWT, FL, FWD and BRX map to the region close to NW0248650 as well as FWT in the ZWRM × Citroides population. QTL for FL and FWD in all populations mapped to the region close to NW0249185 and in the KBS × NHM and SII × Egusi population FWT and RTH mapped to the same position. In the KBS × NHM-GA population, a BRX QTL was also associated with this region. QTL for FSI and FL were identified at similar positions on LG 11 in the KBS × NHM and ZWRM × Citroides populations.

Discussion

Genetic maps, comparative mapping and consensus map

Owing to the marker density, the broad genetic base represented and public availability of 388 SNP markers (ESM 2), it is suggested that these maps be integrated with the high density SSR Map (Ren et al. 2012) to serve as a reference map for *C. lanatus*. The mapped populations were derived from six different parents, representing elite

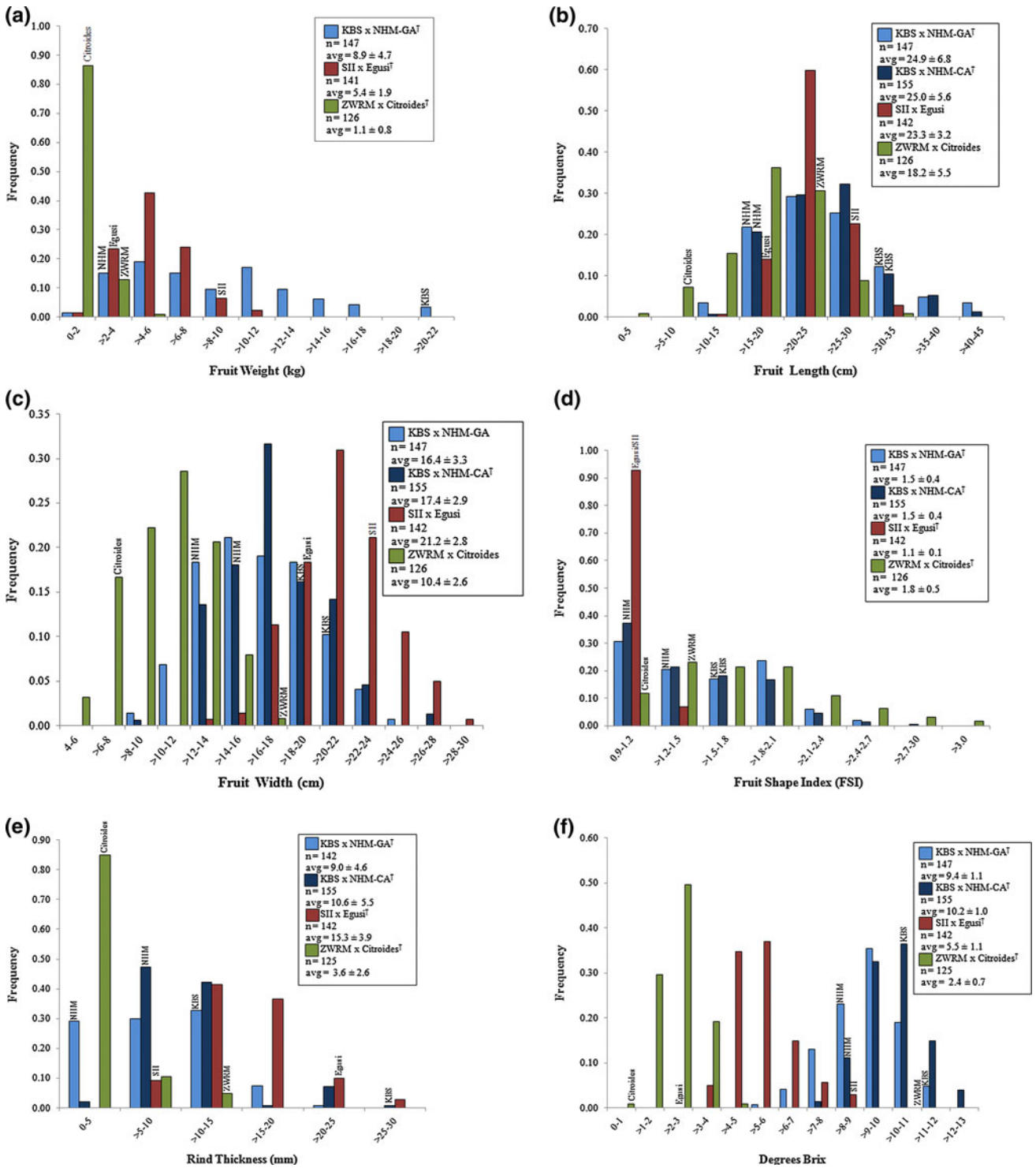

Fig. 2 Frequency distribution for **a** fruit weight (FWT), **b** fruit length (FL), **c** fruit width (FWD), **d** fruit shape index (FSI), **e** rind thickness (RTH), **f** Brix (BRX) in the Klondike Black Seeded × New Hampshire Midget RIL populations grown in Georgia (KBS × NHM-GA) population and California (KBS × NH–CA), and the Strain II × PI 560023 (SII × Egusi) and ZWRM50 × PI 244019 (ZWRM × Citroides) F_2 populations. Normality was tested using the Shapiro–Wilk test (Shapiro and Wilk 1965) and non-normal distributions ($P < 0.05$) are indicated (†) (color figure online)

cultivars (KBS × NHM), an egusi type (SII × Egusi) and an inter-subspecific cross between *C. lanatus* var. *lanatus* and *C. lanatus* var. *citroides* (ZWRM × Citroides). To our knowledge, these are the first published maps for *C. lanatus* var. *lanatus* × *C. lanatus* var. *lanatus* crosses and the first publically available SNP maps for the genus.

Table 2 Pearson correlations for fruit weight (FWT), fruit length (FL), fruit width (FWD), fruit shape index (FSI), rind thickness (RTH) and Brix (BRX) in the (a) Klondike Black Seeded × New Hampshire Midget (KBS × NHM), (b) Strain II × PI 560023 (SII × Egusi) and (c) ZWRM50 × PI 244019 (ZWRM × Citroides) populations; starred numbers (*) indicate significant ($P < 0.05$) correlations

(a)

KBS x NHM	FWT (GA)	FL (GA)	FWD (GA)	FSI (GA)	RTH (GA)	BRX (GA)	FL (CA)	FWD (CA)	FSI (CA)	RTH (CA)
FL (GA)	0.79*									
FWD (GA)	0.84*	0.43*								
FSI (GA)	0.14	0.69*	0.34*							
RTH (GA)	0.66*	0.47*	0.71*	–0.08						
BRX (GA)	0.54*	0.37*	0.53*	–0.03	0.39*					
FL (CA)	0.50*	0.75*	0.19*	0.64*	0.26*	0.15				
FWD (CA)	0.59*	0.23*	0.77*	–0.35*	0.65*	0.45*	0.18*			
FSI (CA)	0.04	0.48*	–0.34*	0.79*	–0.19*	–0.18*	0.75*	–0.48*		
RTH (CA)	0.56*	0.37*	0.59*	–0.06	0.61*	0.29*	0.38*	0.57*	–0.03	
BRX (CA)	0.11	0.07	0.06	0.02	–0.07	0.11	0.02	–0.05	0.03	–0.04

(b)

SII x Egusi	FWT	FL	FWD	FSI	RTH
FL	0.90*				
FWD	0.93*	0.91*			
FSI	–0.08	–0.05	0.00		
RTH	0.46*	0.52*	0.48*	0.00	
BRX	0.38*	0.39*	0.38*	0.00	0.15

(c)

ZWRM x Citroides	FWT	FL	FWD	FSI	RTH
FL	0.66*				
FWD	0.86*	0.60*			
FSI	–0.03	0.62*	–0.22*		
RTH	0.35*	0.33*	0.34*	–0.05	
BRX	0.30*	0.24*	0.34*	–0.05	0.28*

The number of markers present in more than one map is high (~39 %), including 7 % universal markers found in all three maps. This allowed for the alignment of the three maps to create 11 LGs that we regard as representing the 11 haploid chromosomes of the species. A relatively small number of re-arrangements were observed when comparing marker order among the maps. These re-arrangements often represent a discrepancy in the order of a single pair of markers and it cannot be ruled out that these are due to genotyping errors, missing values, or the inclusion of distorted markers (Hackett and Broadfoot 2003; Zraidi et al. 2007). Re-arrangements were also detected in the only previous comparative mapping effort involving *C. lanatus* (Levi et al. 2011), but since no markers are shared between the two studies it is not possible to compare the location of these re-arrangements.

Two LGs, representing 10 SNP markers in the ZWRM × Citroides population did not align with any LGs in other populations. An investigation of the SNP genotypes in the KBS × NHM and SII × Egusi populations showed that 8 out of the 10 and 7 out of 10 markers were not polymorphic in the two populations, respectively. Null alleles were not observed for any of these markers. Since all the markers in LG 12 are fixed in the KBS × NHM population, this might represent a genomic region involved in domestication and/or the location of an undesirable trait.

The genome coverage of the ZWRM × Citroides map (1,144 cM) is less than the previous inter-subspecific F_2 map (2,385 cM) and BC_1 map (1,729 cM) produced by Hashizume et al. (2003) and the BC_2F_2 map (2,162 cM) produced by Levi et al. (2011), but more than the ~800 cM reported by Ren et al. (2012). The average

Table 3 Genomic regions associated with QTL for fruit weight (FWT), fruit length (FL), fruit width (FWD), rind thickness (RTH) and Brix (BRX) in the Klondike Black Seeded × New Hampshire Midget (KBS × NHM), Strain II × PI 560023 (SII × Egusi) and ZWRM50 × PI 244019 (ZWRM × Citroides) populations

Trait	LG	Population	Suggested QTL name	Greenhouse/ field	Location	Year	Left marker	Position (cM)[a]	Maximum LOD	R^2 (%)	Additive effect[b]	Dominance effect	1-LOD support interval (cM)	1-LOD support interval (cM)
FWT	6	KBS × NHM	*fwt6.2*	Field	Watkinsville, GA	2010	NW0251335	162.2 (162.2)	3.6	5.2	−1.085		161.6	166.1
FWT	9	KBS × NHM	*fwt9.2*	Field	Watkinsville, GA	2010	NW0249226	75.7 (148.3)	22.3	45.7	3.252		74.5	78.0
FWT	6	SII × Egusi	*fwt6.1*	Field	Watkinsville, GA	2007	NW0250460	63.5 (66.3)	4.9	8.2	0.669	−0.442	55.0	67.5
FWT	9B	SII × Egusi	*fwt9.1*	Field	Watkinsville, GA	2007	NW0250803	0.0 (71.1)	8.5	15.6	0.982	0.145	0.0	3.0
FWT	9B	SII × Egusi	*fwt9.2*	Field	Watkinsville, GA	2007	NW0248254	98.1 (155.7)	6.3	11.4	−1.042	−0.291	90.2	104.2
FWT	9	ZWRM × Citroides	*fwt9.1*	Greenhouse	Athens, GA	2007	NW0248650	60.5 (85.7)	6.4	20.1	0.540	−0.247	53.7	69.2
FWT	11A	ZWRM × Citroides	*fwt11.1*	Greenhouse	Athens, GA	2007	NW0248648	50.9 (81.1)	5.1	13.2	0.307	0.494	47.0	54.6
FL	9	KBS × NHM	*fl9.2*	Field	Woodland, CA	2010	NW0249226	78.7 (151.3)	7.6	10.4	1.856		75.7	81.9
FL	11	KBS × NHM	*fl11.1*	Field	Woodland, CA	2010	NW0248599	13.3 (13.3)	25.1	40.8	3.651		12.3	13.3
FL	9	KBS × NHM	*fl9.2*	Field	Watkinsville, GA	2010	NW0249226	75.7 (148.3)	13.1	21.1	3.183		74.2	77.6
FL	11	KBS × NHM	*fl11.1*	Field	Watkinsville, GA	2010	NW0248599	13.3 (13.3)	19.4	31.0	3.806		13.1	13.3
FL	6	SII × Egusi	*fl6.1*	Field	Watkinsville, GA	2007	NW0250460	63.5 (66.3)	3.8	8.5	1.313	−0.207	62.9	68.0
FL	9B	SII × Egusi	*fl9.1*	Field	Watkinsville, GA	2007	NW0250803	0.0 (71.1)	3.9	8.7	1.173	0.464	0.0	5.3
FL	9B	SII × Egusi	*fl9.2*	Field	Watkinsville, GA	2007	NW0248254	98.1 (155.7)	5.1	12.1	−1.727	−0.750	89.5	112.7
FL	9	ZWRM × Citroides	*fl9.2*	Greenhouse	Athens, GA	2007	NW0249185	102.6 (146.5)	5.5	11.1	1.114	2.972	98.5	107.8
FL	10	ZWRM × Citroides	*fl10.1*	Greenhouse	Athens, GA	2007	NW0248866	0.0 (0.0)	4.0	7.3	1.284	2.161	0.0	2.8
FL	11A	ZWRM × Citroides	*fl11.1*	Greenhouse	Athens, GA	2007	NW0249365	20.4 (19.4)	18.1	39.2	4.396	3.231	17.0	24.0
FWD	9	KBS × NHM	*fwd9.2*	Field	Woodland, CA	2010	NW0249226	79.7 (152.3)	24.8	43.7	1.920		77.3	81.4
FWD	11	KBS × NHM	*fwd11.1*	Field	Woodland, CA	2010	NW0250945	11.0 (11.0)	7.7	8.7	−0.871		7.2	16.1
FWD	9	KBS × NHM	*fwd9.2*	Field	Watkinsville, GA	2010	NW0249226	76.7 (149.3)	25.5	50.0	2.410		75.7	79.7
FWD	11	KBS × NHM	*fwd11.1*	Field	Watkinsville, GA	2010	NW0248599	12.3 (12.3)	6.9	9.0	−1.000		11.0	13.3
FWD	9B	SII × Egusi	*fwd9.1*	Field	Watkinsville, GA	2007	NW0250803	0.0 (71.1)	7.4	14.1	1.343	0.496	0.0	9.2
FWD	9B	SII × Egusi	*fwd9.2*	Field	Watkinsville, GA	2007	NW0248254	99.1 (156.6)	7.3	14.6	−1.710	−0.574	90.4	103.9
FWD	5	ZWRM × Citroides	*fwd5.1*	Greenhouse	Athens, GA	2007	NW0248277	18.9 (37.0)	4.0	9.2	−1.208	0.372	14.2	32.2
FWD	9	ZWRM × Citroides	*fwd9.2*	Greenhouse	Athens, GA	2007	NW0249185	101.6 (145.2)	6.6	16.0	1.202	0.876	98.1	108.8
FSI	10	KBS × NHM	*fsi10.1*	Field	Woodland, CA	2010	NW0250598	33.1 (33.1)	3.2	3.0	−0.068		29.3	38.3
FSI	11	KBS × NHM	*fsi11.1*	Field	Woodland, CA	2010	NW0248107	15.3 (15.3)	30.8	56.6	0.299		13.4	18.8
FSI	9	KBS × NHM	*fsi9.1*	Field	Watkinsville, GA	2010	NW0248056	99.7 (172.3)	3.3	2.8	−0.068		81.2	105.7
FSI	11	KBS × NHM	*fsi11.1*	Field	Watkinsville, GA	2010	NW0248599	15.3 (15.3)	16.9	21.5	0.305		13.4	19.5
FSI	11A	ZWRM × Citroides	*fsi11.1*	Greenhouse	Athens, GA	2007	NW0249365	20.4 (19.4)	13.7	31.8	0.345	0.312	15.4	25.5
RTH	9	KBS × NHM	*rth9.1*	Field	Woodland, CA	2010	NW0249226	78.7 (151.3)	13.2	29.3	3.110		76.5	80.7
RTH	6	KBS × NHM	*rth6.1*	Field	Watkinsville, GA	2010	NW0247929	188.7 (188.7)	5.1	6.4	−1.168		181.3	197.9

Table 3 continued

Trait	LG	Population	Suggested QTL name	Greenhouse/ field	Location	Year	Left marker	Position (cM)[a]	Maximum LOD	R^2 (%)	Additive effect[b]	Dominance effect	1-LOD support interval (cM)	1-LOD support interval (cM)
RTH	9	KBS × NHM	*rth9.1*	Field	Watkinsville, GA	2010	NW0249226	75.7 (148.3)	25.0	45.0	3.260		75.0	78.6
RTH	2	SII × Egusi	*rth2.1*	Field	Watkinsville, GA	2007	NW0250854	42.6 (51.8)	6.1	17.7	2.345	−0.432	33.2	51.1
RTH	9B	SII × Egusi	*rth9.1*	Field	Watkinsville, GA	2007	NW0248254	97.1 (154.8)	3.6	8.4	−1.775	−0.811	86.5	100.8
BRX	7B	KBS × NHM	*brx7.1*	Field	Woodland, CA	2010	NW0248069	34.5 (227.2)	3.0	7.3	0.266		26.5	36.2
BRX	8	KBS × NHM	*brx8.1*	Field	Woodland, CA	2010	NW0250095	17.9 (17.9)	3.2	7.0	−0.262		15.4	21.3
BRX	7B	KBS × NHM	*brx7.1*	Field	Watkinsville, GA	2010	NW0248069	32.5 (224.5)	3.4	8.4	0.333		26.4	35.5
BRX	9	KBS × NHM	*brx9.2*	Field	Watkinsville, GA	2010	NW0249226	76.7 (149.3)	5.0	12.1	0.404		73.1	81.0
BRX	9B	SII × Egusi	*brx9.1*	Field	Watkinsville, GA	2007	NW0250803	0.0 (71.1)	8.5	21.6	0.653	0.025	0.0	2.2

[a] Number in parenthesis refer to position on consensus map (Fig. 1)

[b] Negative values indicate that the effect is contributed by the allele from male parent

distance between markers was smaller (3.4 cM) than the 4.3 cM (F_2), 7.2 cM (BC_1) and 7.9 cM in the Hashizume et al. (2003) and Levi et al. (2011) maps respectively, but larger than the 0.8 cM in the RIL population (Ren et al. 2012). Both maps derived from the intra-subspecific crosses (KBS × NHM and SII × Egusi) covered a larger genomic region than the inter-subspecific map indicating suppressed recombination in the latter population. The intra-subspecific maps had a larger average gap distance but smaller maximum gap distance and lower segregation distortion (<8 %) than the ZWRM × Citroides map. The marker distortion observed in the latter population (12.4 %) was similar to what was observed by Hashizume et al. (2003) (11 %) and much lower than the 24.7 % found by Levi et al. (2011). However, in the present study, severely distorted markers ($P < 0.0001$) were removed before attempting to construct the maps and were therefore not included in the calculations of percentage segregation distortion. The <8 % segregation distortion observed in the intra-subspecific populations was similar to what has been found in *C. pepo* (~5–10 %) (Esteras et al. 2012; Zraidi et al. 2007). Segregation distortion is common in mapping studies and has been attributed to various causes, including aneuploidy, chromosomal re-arrangements, lethal/deleterious alleles or pollen competition (Tang et al. 2010; Zraidi et al. 2007). The lower segregation distortion in the intra-subspecific crosses is expected because this phenomenon is more common in wide (inter-specific) crosses due to differences in chromosome structure that leads to preferential segregation (Buckler et al. 1999). The 12.7 % segregation distortion in the inter-subspecific ZWRM × Citroides population is less than the ~30 % often observed in inter-specific crosses (Tang et al. 2010), but similar to the 10.2 % observed in the inter-subspecific *Cucurbita pepo* subsp. *pepo* (Zucchini) × *Cucurbita pepo* subsp. *ovifera* (Scallop) population (Esteras et al. 2012), supporting the sometimes controversial infraspecific classification of *C. lanatus*. The direction of segregation distortion favored KBS and wild alleles (Egusi and Citroides), but not exclusively. As would be expected, the distorted markers were often clustered together and markers in a specific cluster favored alleles from the same parent. Ren et al. (2012) reported that in the inter-subspecific RIL population, all the markers in segregation distortion regions (SDR) originated from the elite parent (*C. lanatus* var. *lanatus*). This is in contrast to the results in the present study. However, the selection pressure for elite phenotypes (fruit and seed production) during RIL production might have contributed to the results (Ren et al. 2012). The directionality of distortion is important in plant breeding since it will affect allele introgression of loci in the distorted region. In a region surrounding NW0250460 on LG 6 (EMS 2), elite alleles are favored in both the SII × Egusi

and the ZWRM × Citroides populations. It is therefore anticipated that introgressing desirable wild traits located in this region into elite material would require larger population sizes than predicted by expected Mendelian ratios.

A consensus SNP map was created for *C. lanatus* with BioMercator by using the KBS × NHM map as reference for projection (Fig. 1). The KBS × NHM population was chosen as the reference based on the fact that the vast majority of watermelon breeding involves crosses between elite parents, although crosses with egusi and *citroides* are used, especially for disease resistance introgression. The consensus map can be used as a reference map for use of these SNPs in other populations. However it should be noted that the marker order and distances between markers were calculated based on common markers between maps, rather than recombination frequencies, and should therefore be further validated.

Traits and QTLs

Pairwise correlations between phenotypic traits associated with fruit size (FWT, FL, FWD) were significant in all three populations as well as between the two experimental locations for the KBS × NHM RIL population (Table 2). As can be expected from such correlations, the QTL associated with the three traits were often located in similar genomic regions within populations (Table 3; Fig. 1). Fruit trait QTL co-localization has been reported for several vegetables (deVicente and Tanksley 1993; Tanksley 2004), including other members of *Cucurbitaceae* (Esteras et al. 2012; Monforte et al. 2004; Yuan et al. 2008a, b; Zalapa et al. 2007) and can be due to pleiotropic effects (single gene affecting multiple traits) or tightly linked genes. In melon, fruit shape is associated with two major pleiotropic genes, one affecting sex expression and one controlling carpel number (Fernandez-Silva et al. 2010; Monforte et al. 2004; Perin et al. 2002), but a number of other genes also play a smaller role. Co-localization of fruit traits is also common in cucumber (Yuan et al. 2008a, b) and *C. pepo* (Esteras et al. 2012) and showed some stability across populations with different genetic backgrounds (Yuan et al. 2008a). Fine mapping of the regions associated with these QTL will be needed to determine whether pleiotropic effects are also involved in fruit size and shape traits in watermelon.

In this study QTL regions were identified that were stable across populations, as well as some that were unique to a specific population. Inconsistent QTL detection across genetic backgrounds and environments are common and in a specific population only a subset of QTL will usually be detected (Blanc et al. 2006). However, the different populations in the current study were grown in different environments (greenhouse or field and different years) and had different genetic backgrounds, confounding the genetic and environmental factors contributing to phenotypes. It is therefore not possible to determine which of these two factors are responsible for the inconsistencies. We will therefore concentrate this discussion on those QTL that were stable across different populations or environments since stable QTL are highly desirable for marker assisted selection in plant breeding.

We identified a single genomic region on LG 9 (consensus map: 140.7–167.7 cM) where QTL associated with FL and FWD co-localize in all three populations, as well as FWT and RTH in two of the populations (Table 3; Fig. 1). The percent of phenotypic variation (R^2) explained by QTL for the three traits ranged from 11.4–45.7 % (*fwt9.2*), 10.4–21.1 % (*fl9.2*), 14.6–50.0 % (*fwd9.2*) and 8.4–34 % (*rth9.1*), while the highest R^2 was consistently found in the KBS × NHM population. Additional QTL for FWT (consensus map LG9: 71.1 cM–100.2 cM) and FL (consensus map LG11: 12.3 cM–22.5 cM) were associated with two other regions in two of the three populations (*fwt9.1* and *fl11.1*). Fruit weight has recently become an important consideration for watermelon breeders because consumer preference is shifting away from the traditionally large fruited watermelons to smaller sized fruit. Watermelon varieties that produce fruit that fit into the weight category most preferred by consumers must be a priority for breeders (Gusmini and Wehner 2007).

Genes associated with watermelon fruit size have not been previously described, but it has been suggested that approximately 25 genes might be involved (Wehner et al. 2001). Watermelon fruit shape is usually classified as round, oval, blocky, or elongate (Wehner et al. 2001). The incompletely dominant elongate fruit gene (*O*) is thought to be responsible for elongate (*OO*), blocky (*Oo*) and round (*oo*) fruit (Poole and Grimball 1945; Weetman 1937; Wehner et al. 2001). However, in a recent study of a cross between Mountain Hoosier (round) and Calsweet (elongate) the progeny phenotypes did not fit the expected segregation ratio for a single incompletely dominant gene (Kumar 2009). We identified QTL associated with FSI on three different LGs (Fig. 1), confirming that fruit shape is controlled by more than one gene. However, we did detect a major QTL for FSI (*fsi11.1*) in the KBS × NHM and the ZXWM × Citroides populations (consensus map LG 11: 13.4–23.8 cM) (Table 3; Fig. 1). No QTL for FSI was detected in this region (or any other) in the SII × Egusi population, but that is not surprising since both parents have round fruit shapes and very little segregation was observed in the progeny (Fig. 2d). We postulate that *fsi11.1* is the location of the *O* gene, but that other genes also play a (lesser) role in determining fruit shape in watermelon.

Transgressive segregation in at least one direction was observed for all six traits (Fig. 2). This phenomenon, where

the progeny have trait values outside the range of the parents is generally associated with antagonistic additive effects (deVicente and Tanksley 1993; Rieseberg et al. 1999) although other mechanisms have been suggested (Rieseberg et al. 1999). Antagonistic additive effects were observed for 78 % (14/18) of the trait–population–location combinations where more than one QTL was detected. It is specifically apparent for the FSI trait where the round NHM and Citroides contribute QTL alleles increasing the FSI (*fsi9.1* and *fsi10.1*). The progeny of both of these populations exhibited phenotypes that were much more elongated than either of the parents. These results reiterate the hidden potential in apparently inferior or unadapted phenotypes in plant breeding programs.

In the KBS × NHM RIL population, one minor QTL for BRX was stable across locations (*brx7.1*; consensus map LG7: 216.2–229.5 cM), despite the fact that no significant correlation was detected between BRX at the two locations (Table 2; Fig. 1). However, two additional QTL were not stable across locations (*brx8.1* and *brix9.2*) and none of the BRX QTL was stable across populations. The Brix QTL described by Hashizume et al. (2003) was in *C. lanatus* var. *lanatus* × *C. lanatus* var. *citroides* background but since there are no shared markers between the maps it is not possible to determine whether any of the BRX QTL in the current study are at the same location.

BRX is thought to be polygenic and is known to be not only influenced by genetic background and environment (Gusmini and Wehner 2005; Hashizume et al. 2003; Porter et al. 1940), but also by fruit maturity. In the SII × Egusi, the ZWRM × Citroides and KBS × NHM-GA populations there was a significant correlation between BRX and fruit size traits. A positive correlation between BRX and fruit weight was reported by Showalter (1961) but it was thought to be related to fruit maturity. The parents of all three crosses varied greatly in time to maturity, making it difficult to harvest all fruit at optimum maturity, especially for crosses involving wild accessions (SII × Egusi, the ZWRM × Citroides). The mean BRX of the ZWRM × Citroides F_2 population (2.4 ± 0.7) was shifted more towards the Citroides parent (Brix: 1.4) than the ZWRM parent (Brix: 10.3) (Fig. 2). Hashizume et al. (2003) found the same phenomenon in F_1 progeny of their *C. lanatus* var. *lanatus* × *C. lanatus* var. *citroides* population and attributed it to partial dominance of the *citroides* allele. Data on the F_2 population was not available, but mean Brix of the BC_1 population was close to the midparent value (Hashizume et al. 2003). The ZWRM × Citroides population is currently being advanced to develop a RIL population. During this process seed germination percentage was recorded for seed harvested from the F_2 fruit. Seed germination percentage is related to fruit maturity, at least in elite watermelon cultivars (Demir and Mavi 2004; Nerson 2002). We did not find a significant correlation between seed germination percentage and BRX in the ZWRM × Citroides population (data not shown), suggesting that the shift in BRX phenotypic distribution is not due to immature fruit. However, it cannot be ruled out that some of the BRX QTL identified in this study are influenced by fruit maturity, especially those that are co-localized with fruit size QTL (*brx9.1* and *brix9.2*). The current effort to develop a ZWRM × Citroides RIL population should shed more light on this issue.

We have created the first SNP maps for *Citrullus lanatus*, including a map of elite cultivars. The consensus SNP map will be a valuable resource for mapping in the species and should increase the ability to compare results among future studies. We have identified major QTL for fruit size and shape that are stable across genetic backgrounds and environments. These QTL should be prime candidates for marker assisted selection in watermelon.

Acknowledgments This research was funded by Monsanto. The authors thank Hussein Abdel-Haleem for assistance with heritability calculations.

References

Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J (2004) BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. Bioinformatics 20:2324–2326

Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor Appl Genet 113:206–224

Buckler ES, Phelps-Durr TL, Buckler CSK, Dawe RK, Doebley JF, Holtsford TP (1999) Meiotic drive of chromosomal knobs reshaped the maize genome. Genetics 153:415–426

Che K-P, Liang C-Y, Wang Y-G, Jin D-M, Wang B, Xu Y, Kang G-B, Zhang H-Y (2003) Genetic assessment of watermelon germplasm using the AFLP technique. HortScience 38:81–84

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Dane F, Lang P (2004) Sequence variation at cpDNA regions of watermelons and related wild species: implications for evolution of *Citrullus* haplotypes. Am J Bot 91:1922–1929

Dane F, Liu J (2007) Diversity and origin of cultivated and citron type watermelon (*Citrullus lanatus*). Genet Resour Crop Evol 54:1255–1265

Deleu W, Esteras C, Roig C, Gonzalez-To M, Fernandez-Silva I, Gonzalez-Ibeas D, Blanca J, Aranda M, Arus P, Nuez F, Monforte A, Pico M, Garcia-Mas J (2009) A set of EST-SNPs for map saturation and cultivar identification in melon. BMC Plant Biol 9:90

Demir I, Mavi K (2004) The effect of priming on seedling emergence of differentially matured watermelon (*Citrullus lanatus* (Thunb.) Matsum and Nakai) seeds. Sci Horticult 102:467–473

deVicente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genet Mol Biol 142:285–294

Esteras C, Gomez P, Monforte AJ, Blanca J, Vicente-Dolera N, Roig C, Nuez F, Pico B (2012) High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping. BMC Genomics 13:80

Fernandez-Silva I, Moreno E, Essafi A, Fergany M, Garcia-Mas J, Martín-Hernandez A, Álvarez J, Monforte A (2010) Shaping melons: agronomic and genetic characterization of QTLs that modify melon fruit morphology. Theor Appl Genet 121:931–940

Food and Agriculture Organization of the United Nations (2011) Crop production. http://www.faostat.fao.org/site/567/default.aspx#ancor. Accessed 18 July 2011

Fursa TB (1972) K sistematic roda *Citrullus* Schrad. [On the taxonomy of genus *Citrullus* Schrad.]. Botanicheski Zhurnal 57:31–41

Guner N, Wehner TC (2004) The genes of watermelon. HortScience 39:1175–1182

Gusmini G, Wehner TC (2005) Foundations of yield improvement in watermelon. Crop Sci 45:141–146

Gusmini G, Wehner TC (2007) Heritability and genetic variance estimates for fruit weight in watermelon. HortScience 42:1332–1336

Hackett CA, Broadfoot LB (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity 90:33–38

Hashizume T, Shimamoto I, Harushima Y, Yui M, Sato T, Imai T, Hirai M (1996) Construction of a linkage map for watermelon (*Citrullus lanatus* (Thunb.) Matsum & Nakai) using random amplified polymorphic DNA (RAPD). Euphytica 90:265–273

Hashizume T, Shimamoto I, Hirai M (2003) Construction of a linkage map and QTL analysis of horticultural traits for watermelon [*Citrullus lanatus* (Thumb.) Matsum & Nakai] using RAPD, RFLP and ISSR markers. Theor Appl Genet 106:779–785

Hawkins LK, Dane F, Kubisiak TL, Rhodes BB, Jarret RL (2001) Linkage mapping in a watermelon population segregating for Fusarium wilt resistance. J Am Soc Hort Sci 126:344–350

Henry RJ (2008) Plant genotyping II: SNP technology. CABI, Wallingford

Holland JB, Nyquist WE, Cervantes-Martinez CT (2003) Estimating and interpreting heritability for plant breeding: an update. Plant Breed Rev 22:9–111

Jarret RL, Merrick LC, Holms T, Evans J, Aradhya MK (1997) Simple sequence repeats in watermelon (*Citrullus lanatus* (Thunb.) Matsum. & Nakai). Genome 40:433–441

JMP Version 8.0.2 (2009) SAS Institute Inc., Cary, NC, 1989–2009

Kole C, Abbott AG (2008) Principles and practices of plant genomics, vol 1. Genome mapping Science Publishers, Enfield

Kumar R (2009) Inheritance of fruit yield and other horticulturally important traits in watermelon [*Citrullus lanatus* (Thunb.) Matsum. & Nakai]. North Carolina State University, Raleigh

Levi A, Thomas CE (2005) Polymorphisms among chloroplast and mitochondrial geneomes of *Citrullus* species and subspecies. Genet Resour Crop Evol 52:609–617

Levi A, Thomas CE, Keinath AP, Wehner TC (2001a) Genetic diversity among watermelon (*Citrullus lanatus* and *Citrullus colocynthis*) accessions. Genet Resour Crop Evol 48:559–566

Levi A, Thomas CE, XP XPZ, Joobeur T, Dean RE, Wehner TC, Carle BR (2001b) A genetic linkage map for watermelon based on randomly ampliWed polymorphic DNA markers. J Hort Sci 126:730–737

Levi A, Thomas C, Joobeur T, Zhang X, Davis A (2002) A genetic linkage map for watermelon derived from a testcross population: (*Citrullus lanatus* var. *citroides* × *C. lanatus* var. *lanatus*) × *Citrullus colocynthis*. Theor Appl Genet 105:555–563

Levi A, Thomas CE, Trebitsh T, Salman A, King J, Karalius J, Newman M, Reddy OUK, Xu Y, Zhang X (2006) An extended linkage map for watermelon based on SRAP, AFLP, SSR, ISSR, and RAPD markers. J Am Soc Hort Sci 131:393–402

Levi A, Wechter P, Massey L, Carter L, Hopkins D (2011) An extended genetic linkage map for watermelon based on a testcross and a BC_2F_2 population. Am J Plant Sci 2:93–110

MacGillivray JH (1947) Soluble solids content of different regions of watermelon. Plant Physiol 22:637–640

Maynard DN (2001) An introduction to the watermelon. In: Maynard DN (ed) Watermelon characteristics, production and marketing. ASHS Press, Alexandria, pp 9–20

McGregor CE (2011) *Citrullus lanatus* germplasm of Southern Africa. Isr J Plant Sci (in press)

Monforte AJ, Oliver M, Gonzalo MJ, Alvarez JM, Dolcet-Sanjuan R, Arús P (2004) Identification of quantitative trait loci involved in fruit quality traits in melon (*Cucumis melo* L.). Theor Appl Genet 108:750–758

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Resour 8:4321–4325

Navot N, Sarfatti M, Zamir D (1990) Linkage relationships of genes affecting bitterness and flesh color in watermelon. J Hered 81:162–165

Nerson H (2002) Effects of seed maturity, extraction practices and storage duration on germinability in watermelon. Sci Horticult 93:245–256

Nimmakayala P, Tomason YR, Jeong J, Ponniah SK, Karunathilake A, Levi A, Perumal R, Reddy UK (2010) Genetic reticulation and interrelationships among *Citrullus* species as revealed by joint analysis of shared AFLPs and species-specific SSR alleles. Plant Genetic Resour 8:16–25

Nyquist WE, Baker R (1991) Estimation of heritability and prediction of selection response in plant populations. Crit Rev Plant Sci 10:235–322

Perin C, Hagen S, De Conto V, Katzir N, Danin-Poleg Y, Portnoy V, Baudracco-Arnas S, Chadoeuf J, Dogimont C, Pitrat M (2002) A reference map of *Cucumis melo* based on two recombinant inbred line populations. Theor Appl Genet 104:1017–1034

Périn C, Hagen L, Giovinazzo N, Besombes D, Dogimont C, Pitrat M (2002) Genetic control of fruit shape acts prior to anthesis in melon (*Cucumis melo* L.). Mol Genet Genomics 266:933–941

Poole CF, Grimball PC (1945) Interaction of sex, shape, and weight genes in watermelon. J Agric Res 71:533–552

Porter DR (1933) Watermelon breeding. Hilgardia 7:533–552

Porter DR (1937) Inheritance of certain fruit and seed characters in watermelons. Hilgardia 10:489–509

Porter DR, Bission CS, Allinger HW (1940) Factors affecting the total soluble solids, reducing sugars, and sucrose in watermelons. Hilgardia 13:31–66

Ren Y, Zhao H, Kou Q, Jiang J, Guo S, Zhang H, Hou W, Zou X, Sun H, Gong G, Levi A, Xu Y (2012) A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. PLoS One 7:e29453

Rieseberg LH, Archer MA, Wayne RK (1999) Transgressive segregation, adaptation and speciation. Heredity 83:363–372

Robinson RW, Decker-Walters DS (1997) Cucurbits. CAB International Publishing, Oxon

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52:591–611

Showalter RK (1961) Specific gravity, weight, and solids relationships in watermelons. Florida St Hort Soc 74:268–271

Tang S, Okashah R, Knapp S, Arnold M, Martin N (2010) Transmission ratio distortion results in asymmetric introgression in Louisiana Iris. BMC Plant Biol 10:48

Tanksley SD (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell 16:S181–S189

United States Department of Agriculture—National Agricultural Statistics Service (2011) 2010 Agricultural Statistics. http://www.nass.usda.gov/Publications/Ag_Statistics/2010/index.asp. Accessed 09 July 2011

Van Ooijen JW (2006) JoinMap®4 Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V, Wageningen

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Wang S, Basten CJ, Zeng ZB (2011) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh

Weetman LM (1937) Inheritance and correlation of shape, size and color in the watermelon *Citrullus vulgaris* Schrad. Iowa Agric Exp Station Ann Bull 228:224–256

Wehner T (2008) Watermelon. In: Prohens J, Nuez F (eds) Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae, and Cucurbitaceae. Springer, New York, pp 381–418

Wehner TC, Shetty NV, Elmstron GW (2001) Breeding and seed production. In: Maynard DN (ed) Watermelons: characteristics, production, and marketing. ASHS Press, Alexandria, pp 27–73

Yuan X, Pan J, Cai R, Guan Y, Liu L, Zhang W, Li Z, He H, Zhang C, Si L, Zhu L (2008a) Genetic mapping and QTL analysis of fruit and flower related traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Euphytica 164:473–491

Yuan XJ, Li XZ, Pan JS, Wang G, Jiang S, Li XH, Deng SL, He HL, Si MX, Lai L, Wu AZ, Zhu LH, Cai R (2008b) Genetic linkage map construction and location of QTLs for fruit-related traits in cucumber. Plant Breed 127:180–188

Zalapa J, Staub J, McCreight J, Chung S, Cuevas H (2007) Detection of QTL for yield-related traits using recombinant inbred lines derived from exotic and elite US Western Shipping melon germplasm. Theor Appl Genet 114:1185–1201

Zeng ZB (1993) Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. PNAS 90:10972–10976

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468

Zhang R, Xu Y, Yi K, Zhang H, Liu L, Gong G, Levi A (2004) A genetic linkage map for watermelon derived from recombinant inbred lines. J Am Soc Hort Sci 129:237–243

Zraidi A, Stift G, Pachner M, Shojaeiyan A, Gong L, Lelley T (2007) A consensus map for *Cucurbita pepo*. Mol Breed 20:375–388