Theor Appl Genet (2015) 128:1329–1342
DOI 10.1007/s00122-015-2509-x

ORIGINAL PAPER

Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq

Eudald Illa-Berenguer[1] · Jason Van Houten[1,2] · Zejun Huang[1,3] · Esther van der Knaap[1]

Received: 10 January 2015 / Accepted: 20 March 2015 / Published online: 18 April 2015

Abstract

Key message **Bulk segregant analysis coupled with whole genome sequencing is a powerful approach and cost-effective method to identify loci controlling fruit traits in tomato.**

Abstract Domestication of fruit and vegetable crops was accompanied by selection for weight of the edible parts. Increases in fruit weight are controlled by multiple quantitative trait loci (QTL). To date, only two fruit weight genes have been cloned and a third has been fine-mapped. Genes that control locule number also impact fruit weight and two of them are known. To efficiently identify additional tomato fruit weight (FW) and locule number (LC) loci, six F_2 populations were generated from crosses between closely related tomato accessions for which the alleles of the cloned FW and LC genes were known. We employed the bulk segregant approach coupled to whole genome sequencing (QTL-seq) which led to the identification of three highly significant and newly mapped FW QTL. *fw11.2* was located in the distal part of chromosome 11 above the known loci *fas* and *fw11.3*; *fw1.1* in the pericentromeric region of chromosome 1; and *fw3.3* located ~1.6 Mb below the known fruit weight gene, *SlKLUH/FW3.2*. In addition, we mapped three LC QTL (*lcn2.4*, *lcn5.1*, and *lcn6.1*) although their significance was generally low. To confirm the location of the gene underlying *fw11.2*, we developed additional markers and conducted progeny tests. These results allowed us to narrow down the *fw11.2* QTL to a region of ~750 kb corresponding to 66 candidate genes. Our research approach provided a cost-effective and time-efficient method for the identification of additional genes involved in FW and LC that could be used for both fruit development studies and crop improvement programs.

Communicated by S. Huang.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2509-x) contains supplementary material, which is available to authorized users.

✉ Esther van der Knaap
vanderknaap.1@osu.edu

[1] Department of Horticulture and Crop Science, The Ohio State University/OARDC, Wooster, OH 44691, USA

[2] Present Address: Basic Medical Science Program, Wayne State University School of Medicine, Detroit, MI 48202, USA

[3] Present Address: Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China

Introduction

The process of plant domestication and diversification was accompanied by the selection of features that increase adaptation to cultivation as well as features that are desirable for human consumption and use (Gepts 2004; Purugganan and Fuller 2009). These so-called domestication traits are associated with increase the ease of harvest and include seed retention (non-shattering), increase in fruit and seed size, changes in reproductive strategy and plant architecture, and loss of seed dormancy (Fuller 2007; Meyer and Purugganan 2013). Many studies have demonstrated that domestication traits are genetically controlled (Tanksley 2004; Doebley et al. 2006; Meyer and Purugganan 2013; Olsen and Wendel 2013; Zuo and Li 2014). The underlying loci have been mapped in segregating populations that were derived from crosses between wild and cultivated plants. For example, seed weight QTL were identified in various legume

species including mungbean (Fatokun et al. 1992), cowpea (Fatokun et al. 1992), common bean (Koinange et al. 1996), pea (Timmerman-Vaughan et al. 1996) and soybean (Maughan et al. 1996). Increase in the size of fruit and vegetables appears most relevant in species with edible fruits such as tomato (Grandillo et al. 1999; Tanksley 2004), pepper (Zygier et al. 2005), cucumber (Yuan et al. 2008), melon (Monforte et al. 2014), watermelon (Sandlin et al. 2012) or cherry (Olmstead et al. 2007; Campoy et al. 2014).

The identification of the genes underlying quantitative characters such as fruit weight is an arduous task due to the quantitative nature of the trait. Many loci are thought to control weight, each with a varying effect on the character. Some loci play a major role, explaining up to 30 % of the variance (Frary et al. 2000) whereas others play a smaller role of only 7 % of the total variance (van der Knaap and Tanksley 2003; Huang and van der Knaap 2011). In addition, many loci of even smaller effects are typically not detected as a result of the lack of statistical power and the overshadowing magnitude of the effect exhibited by major loci that are segregating in the population.

The first cloned vegetable fruit weight QTL that resulted from the domestication and selection process is *FW2.2* (Frary et al. 2000). This gene encodes a protein of the cell number regulator family (CNR) (Guo et al. 2010) and is thought to be localized at the plasma membrane (Cong and Tanksley 2006; Libault et al. 2010). We recently cloned the second fruit weight QTL, *FW3.2* (Chakrabarti et al. 2013). The underlying gene encodes a cytochrome P450 and is an ortholog of KLUH that has been shown to regulate organ size in Arabidopsis (Anastasiou et al. 2007). Association mapping of the *fw3.2* locus led to the identification of an SNP, M9, located in the promoter of *SlKLUH* that is highly, significantly correlated with the increase in fruit weight. Further genetic evaluations aimed at testing the functional relevance of the M9 SNP demonstrated that fruit weight was indeed segregating according to the allele of M9 (Chakrabarti et al. 2013). However, in certain populations, the segregation was not as significant or was only observed in backcross populations. These findings suggest that other loci with a larger and/or epistatic effect were segregating, negating the impact of *SlKLUH* on fruit weight. Two other loci also control fruit weight by regulating the number of locules in a fruit. The loci *lc* and *fas* most likely underlie the orthologs of *WUSCHEL* and *YABBY2*, respectively (Lippman and Tanksley 2001; Barrero et al. 2006; Cong et al. 2008; Muños et al. 2011).

With the advent of the reference genome sequence of tomato, we sought to test the utility of a next generation sequencing approach to expedite mapping of quantitative trait loci. In the past, populations derived from distantly related parents were used, featuring many nucleotide polymorphisms that facilitated molecular marker development and fine mapping. Populations derived from wide crosses also yielded many QTL segregating for the traits (Grandillo et al. 1999; van der Knaap and Tanksley 2003). Instead, populations derived from closer related parents should lead to a reduction in the number of segregating loci that are impacting the trait and expedite the fine mapping. Moreover, the number of nucleotide polymorphisms are not limiting when employing a whole genome resequencing approach. When few QTL are predicted to segregate, a bulk segregant analysis (BSA; Michelmore et al. 1991) followed by whole genome sequencing (QTL-seq) should be a good approach to expedite the initial mapping of loci (Takagi et al. 2013).

The main focus of this study was to utilize the QTL-seq approach to map fruit weight and locule number QTL in several segregating tomato populations. In addition, we sought to determine the genetic basis of the minor role of *fw3.2* in certain populations and to evaluate the general feasibility of the QTL-seq method to map loci underlying quantitative traits in tomato and other crops.

Materials and methods

Plant material

CC4 corresponded to Moneymaker; CC7 corresponded to LA0147; CC9 corresponded to VIR1011; CC37 corresponded to VIR933 (N^{o} 2759 Enano); CC39 corresponded to VIR347 (N^{o} 347 Yablochnyi) and were obtained from Dr. Mathilde Causse (INRA, Avignon, France) (Ranc et al. 2008). LA1589, LA1655, LA2690 and LA2845 were obtained from the TGRC (http://tgrc.ucdavis.edu). LYC0440 and LYC1891 were obtained from IPK Gatersleben, Germany (http://www.ipk-gatersleben.de). Rio Grande, Howard German and Heinz 1439 were obtained from Tomato Growers Supply Company (http://www.tomatogrowers.com). The latter nine accessions were used in a previous study (Rodríguez et al. 2011). Six intraspecific F_2 populations derived from independent crosses were developed and evaluated for fruit weight or locule number (Table 1). One F_2 population (12S143) was a cross between *Solanum lycopersicum* var. *cerasiforme* (SLC) accessions, whereas all other F_2 populations were crosses between *Solanum lycopersicum* var. *lycopersicum* (SLL) and SLC accessions. The three populations that were evaluated for fruit weight only (12S139, 12S141, and 12S143) were segregating for the M9 SNP at the *fw3.2* locus (Chakrabarti et al. 2013) while *fw2.2* was fixed. The three populations that were evaluated primarily for locule number (12S75, 12S76, and 12S97) were segregating for *fas* while *lc* was fixed. To prevent the impact of *fas*, which is known to control locule number and weight, F_2 plants were selected to be heterozygous or homozygous mutant at *fas*. For most of

Table 1 List of accessions used in this study

F_2 population	Parental name	Species category	Fruit weight (g)	Locule number	Fruit weight genes		Fruit shapes genes			
					FW2.2	*FW3.2*	*LC*	*FAS*	*OVATE*	*SUN*
12S139 ($n = 96$)	CC7	*S. lycopersicum* var. *lycopersicum*	87.02 ± 14.20	–	1	1	1	3	3	3
	CC9	*S. lycopersicum* var. *cerasiforme*	15.40 ± 1.12	–	1	3	1	3	3	3
12S141 ($n = 95$)	CC4	*S. lycopersicum* var. *lycopersicum*	63.22 ± 5.53	–	1	1	3	3	3	3
	CC9	*S. lycopersicum* var. *cerasiforme*	15.40 ± 1.12	–	1	3	1	3	3	3
12S143 ($n = 95$)	CC37	*S. lycopersicum* var. *cerasiforme*	22.76 ± 1.84	–	1	1	1	3	3	3
	CC39	*S. lycopersicum* var. *cerasiforme*	27.17 ± 0.78	–	1	3	3	3	3	3
12S75[a] ($n = 94$)	Heinz1439	*S. lycopersicum* var. *lycopersicum*	159.90 ± 20.98	6.08 ± 0.68	1	1	1	3	3	3
	LA1655	*S. lycopersicum* var. *cerasiforme*	48.32 ± 1.61	9.70 ± 0.70	1	3	1	1	3	3
12S76[a] ($n = 78$)	LYC0440	*S. lycopersicum* var. *lycopersicum*	258.57 ± 17.15	14.00 ± 1.13	1	1	1	1	3	3
	LA2845	*S. lycopersicum* var. *cerasiforme*	21.05 ± 3.66	5.23 ± 0.11	3	1	1	3	3	3
12S97[b] ($n = 91$)	LYC1891	*S. lycopersicum* var. *lycopersicum*	77.43 ± 2.69	2.40 ± 0.09	1	1	3	3	3	3
	LA2690	*S. lycopersicum* var. *cerasiforme*	4.05 ± 0.58	2.40 ± 0.26	3	3	3	1	3	3

n equals the size of the population, *1* derived allele, resulting in large fruit or more locules, *3* wild-type allele, resulting in small fruit or fewer locules

[a] Selected to be heterozygous at *fas*

[b] Selected to be homozygous mutant at *fas*

the locule number populations, fruit weight varied extensively in the parents and fruit weight alleles were segregating in the F_2 populations (Table 1). The F_2, F_1 and parents were grown in the experimental fields at the Ohio State University/OARDC Wooster campus in summer 2012. Previously, an F_2 population from a cross between Rio Grande × LA1589 and Howard German × LA1589 let to the identification *fw11.2* (Gonzalo and van der Knaap 2008). To confirm this QTL, the corresponding backcross populations (BC_2 and BC_1F_4, respectively) were generated and evaluated in the present work.

Phenotypic correlations between pairs of traits within an experiment and for trait between experiments were calculated using the Pearson correlation coefficient (rp).

DNA isolation and sequencing of pooled samples

Genomic DNA was extracted from young leaves using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). DNA quantity was measured with the Qubit® 2.0 fluorometer (Invitrogen-Molecular Probes, Eugene, OR). The instrument was calibrated with the Qubit™ dsDNA HS Standards according to the manufacturer's specifications.

QTL-seq approach was applied to three populations; two for fruit weight (12S139 and 12S143) and one for locule number (12S75). For each population, two bulks comprising 10 plants each were generated. The bulks consisted of pooled DNA from plants featuring the most extreme values for the trait. Libraries of an approximate insert size of 300 bp were generated for each bulk and barcoded at the Genome Technology Access Center (GTAC) facility at Washington University (St Louis, MO). Two bulks were pooled per lane on the Illumina HiSeq 2000 and paired-end sequenced at 101 bp. FASTQ files were assessed using the FastQC program (version 0.10.1; Andrews 2010) and filtered for a quality value of at least 28. The summary statistics show genome coverage in each of the bulks (Supplementary Table S1).

Analyses of the short reads

The sequence data were trimmed and filtered prior to analysis. The short reads obtained from the two DNA-bulks were aligned against the tomato genome sequence (The Tomato Genome Consortium 2012) using Bowtie 2 (Langmead and Salzberg 2012). The preset option "--very-sensitive-local" of Bowtie 2 was used. By default, these settings attempt to find either an exact or a 1-mismatch end-to-end alignment for the read before trying the multiseed heuristic. For each read, only the best alignment would be allowed and no multiple alignments would be permitted when mapping to the reference genome. Files were converted to BAM files using SAM tools (Li et al. 2009), sorted and then compared to locate duplicate records using Picard software (http://picard.sourceforge.net). Re-alignment (BAQ) was done to avoid false SNP calls near indels. The resulting files were applied to GATK SNP-calling (McKenna et al. 2010; DePristo et al. 2011). The proportion of alleles corresponding to each of the two parental genomes was calculated by counting the number of reads harboring a SNP compared to the reference genome sequence (AD_ALT) and divided

by the total number of reads (DP) (SNP-index = AD_ALT/DP). This so-called SNP-index was calculated for all positions in the genome. Since the parental genomes were likely to harbor SNP with the reference genome, we filtered out the positions where both bulks presented the same allele leading to a reduction of the number of false positive SNP. To identify regions of the genome associated with fruit weight and locule number, the SNP-indices were subtracted between the bulks (ΔSNP-index = SNP-index_'Largest' − SNP-index_'Smallest') and sliding window analysis was applied averaging the ΔSNP-index within a 1 Mb window size and 10 kb step increment using an in-house developed Python script. The average was plotted for all chromosomes to discern QTL from the regions that did not contribute to the variation for the traits of interest. We expected most of the genomic regions show ΔSNP-index = 0, but only regions with a high absolute ΔSNP-index values would suggest a major contribution to the trait variation.

Marker development and genotyping

PCR-based markers that distinguish wild and cultivated alleles of the known fruit weight and locule number loci, *fw2.2*, *fw3.2*, *fas* and *lc* (Supplementary Table S2), were evaluated in the mapping parents and F_2 populations if the genes were known to be segregating. The new QTL identified by QTL-seq were also mapped in all six populations when we were successful in identifying polymorphic markers in those populations. Indels (≤25 bp) and SNPs that were in the vicinity of the QTL were converted into PCR-based markers (Supplementary Table S2). Kruskal–Wallis test or one-way analysis of variance (ANOVA) was used to test the significance of the alleles with fruit weight or locule number variation in the populations. The degree of dominance of the alleles at a specific locus was calculated as d/a ratio, where $d = \mathrm{Aa} - (\mathrm{AA} + \mathrm{aa})/2$ and $a = (\mathrm{AA} - \mathrm{aa})/2$, where AA is the mean value for the homozygous parent 1, aa is the mean value for the homozygous parent 2 and Aa is the mean value for heterozygous. The percentage of phenotypic variation explained by each QTL (R^2) was estimated by multiple-regression analysis, using as explanatory variables the most significant markers for each QTL.

Data analyses were performed with the R open-source software (version 3.1.1; R Core Team 2014).

Heritability estimates

The broad-sense heritability (H^2) of each trait was calculated by variance components method according to Mather and Jinks (1982). The phenotypic variance for the parent lines (V_{P_1} and V_{P_2}) and F_1 progeny (V_{F_1}) is due to environmental factors, whereas the phenotypic variance in the F_2 (V_{F_2}) includes a sum of genetic (V_G) and environmental (V_E) variances. Broad-sense heritability can be estimated by:

$$H^2 = \frac{V_G}{V_{F_2}} = \frac{V_{F_2} - {}^1\!/_4(V_{P_1} + V_{P_2} + 2V_{F_1})}{V_{F_2}}.$$

Epistatic interactions

Epistatic digenic interactions of *fw3.2* with other loci were determined based on the following model of analysis of variance (ANOVA) for each pair of genes that is involved in fruit size variation.

$$Y_{ijl} = \mu + \mathrm{fw3.2}_i + \mathrm{QTL}_j + \mathrm{fw3.2\ QTL}_{ij} + e_{ijl}$$

where μ represents the population mean value; $fw3.2_i$, the effect of the *fw3.2* locus; QTL_j, the effect of the QTL detected here; $fw3.2QTL_{ij}$, the interaction between gene effects; e_{ijl}, are the experimental errors.

Confirming the QTL interval for *fw11.2*

To determine more precisely the interval of *fw11.2* QTL, nine F_2 plants carrying a recombination between markers 13EP232 and 13EP236 were selected from 12S139 population. Seeds from each recombinant F_2 plant were sown and approximately 12 seedlings that were homozygous recombinant and 12 that were homozygous non-recombinant were identified with molecular markers. These plants were transplanted in the field for fruit weight evaluation in summer 2013. Average fruit weights were determined from twenty ripe fruits per plant. Student's *t* test was used to detect significant differences between the homozygous recombinant and non-recombinant genotypes in each F_3 family.

Results

Trait variation in the segregating populations

Three F_2 populations were evaluated for fruit weight (FW) and another three were evaluated for both FW and locule number (LC) (Table 1). The frequency distributions showed continuous variation for both characters (Figs. 1, 2), indicating that FW and LC were quantitatively inherited traits. For FW, the mean values of the F_1 and F_2 plants were skewed towards the small fruited parent. Three of the six families were genotypically selected prior to transplanting in the field to be heterozygous (12S75 and 12S76) or homozygous mutant (12S97) for *fas*. Since *fas* has a large impact on fruit weight, the selection for the mutant allele was expected to result in an average larger fruit in the 12S97 F_2 population. However, the 12S97 family was still

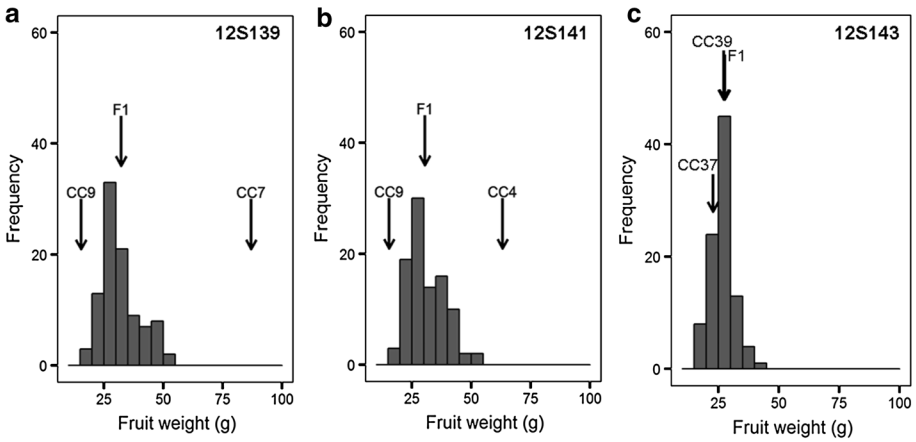

Fig. 1 Frequency distribution of fruit weight in three F_2 tomato populations. **a** 12S139, **b** 12S141, **c** 12S143

skewed towards the small fruited parent (Fig. 2g), suggesting that the small fruit alleles of the segregating loci were dominant over the large fruit alleles. The 12S143 population exhibited transgressive segregation for FW as some plants carried smaller and others larger fruit than either parent (Fig. 1c). The CC37 parent of this population carried slightly smaller fruit compared to the CC39 parent while carrying the mutant alleles of *lc* and *fw3.2*, typically leading to increased FW (Table 1). This suggested that another FW QTL must be segregating in this population and that the effect of increased fruit weight originated with the CC39 parent.

For LC, the segregation in the F_2 populations showed similar trends as for FW. In general, the LC mean values of the F_2 were skewed to the low locule number parent and the F_1 or even smaller (Fig. 2). In contrast, the LC mean values in 12S97 population were higher than the F_1 and both parents. In addition, even though all F_2 plants carried the mutant allele of *fas*, LC varied from 2.25 to 5.40 in this population, suggesting the presence of another locus that enhanced the effect of *fas* on locule number.

Broad-sense heritability was calculated for the six populations (Supplementary Table S3). In the case of FW, heritability estimates showed a wide range (0.12–0.93), with the minimum and maximum H^2 values in the 12S139 and 12S97 populations, respectively. The low H^2 was likely due to high variation in fruit weight for some parents and the F_1 (Supplementary Table S3). The H^2 for LC was higher and ranged between 0.40 and 0.93. Additionally, the correlation between FW and LC was relatively high, suggesting that fruit weight QTL positively increase locule number or vice versa (Fig. 2c, f, i).

Association of the known FW and LC QTL with the traits

The three populations that were primarily studied for variation in fruit weight (12S139, 12S141 and 12S143; Fig. 1) were used in a previous study aimed at validating the segregation of the M9 SNP with the *fw3.2* QTL (Chakrabarti et al. 2013). The alleles of *fw2.2* and *fas* were fixed in all populations whereas *lc* segregated in two of the three populations. To confirm the previous findings using a different population size, F_2 seedlings from the same F_1 were evaluated for fruit weight. As expected, only the 12S141 population showed strong association of the M9 SNP and FW (PVE of 0.35), whereas in the other two families, the segregation was barely or not significant (Table 2). This suggested that other fruit weight QTL were segregating in 12S139 and 12S143. Interestingly, *lc* was associated with fruit weight only in the 12S143 population and not in the 12S141 population. These results combined suggested that genetic background could have a large impact on the FW trait or, in the case of *lc*, that linkage to FW QTL might be the cause of marker-trait association.

In the tomato germplasm, variation in locule number appears to be higher than what can be explained by alleles at the two loci, *lc* and *fas* (van der Knaap et al. 2014). Since LC and FW traits showed significant correlation, we wanted to find out whether the known FW loci *fw2.2* or *fw3.2* segregated for locule number. We genotyped the alleles in the three populations and found that the known FW QTL were not associated with locule number except for *fw2.2* and only in the 12S97 population (Table 2). Taken together, these results suggested the presence of another LC QTL controlling locule number variation in these populations.

QTL-seq of three populations

We selected three populations for the identification of additional fruit morphology loci using the QTL-seq approach. Two populations were selected to identify FW QTL, 12S139 and 12S143, and one to identify LC QTL, 12S75. For the two fruit weight populations, in both we identified a single QTL, *fw11.2* and *fw1.1*, respectively (Fig. 3). The results for the other chromosomes did not show the presence of any

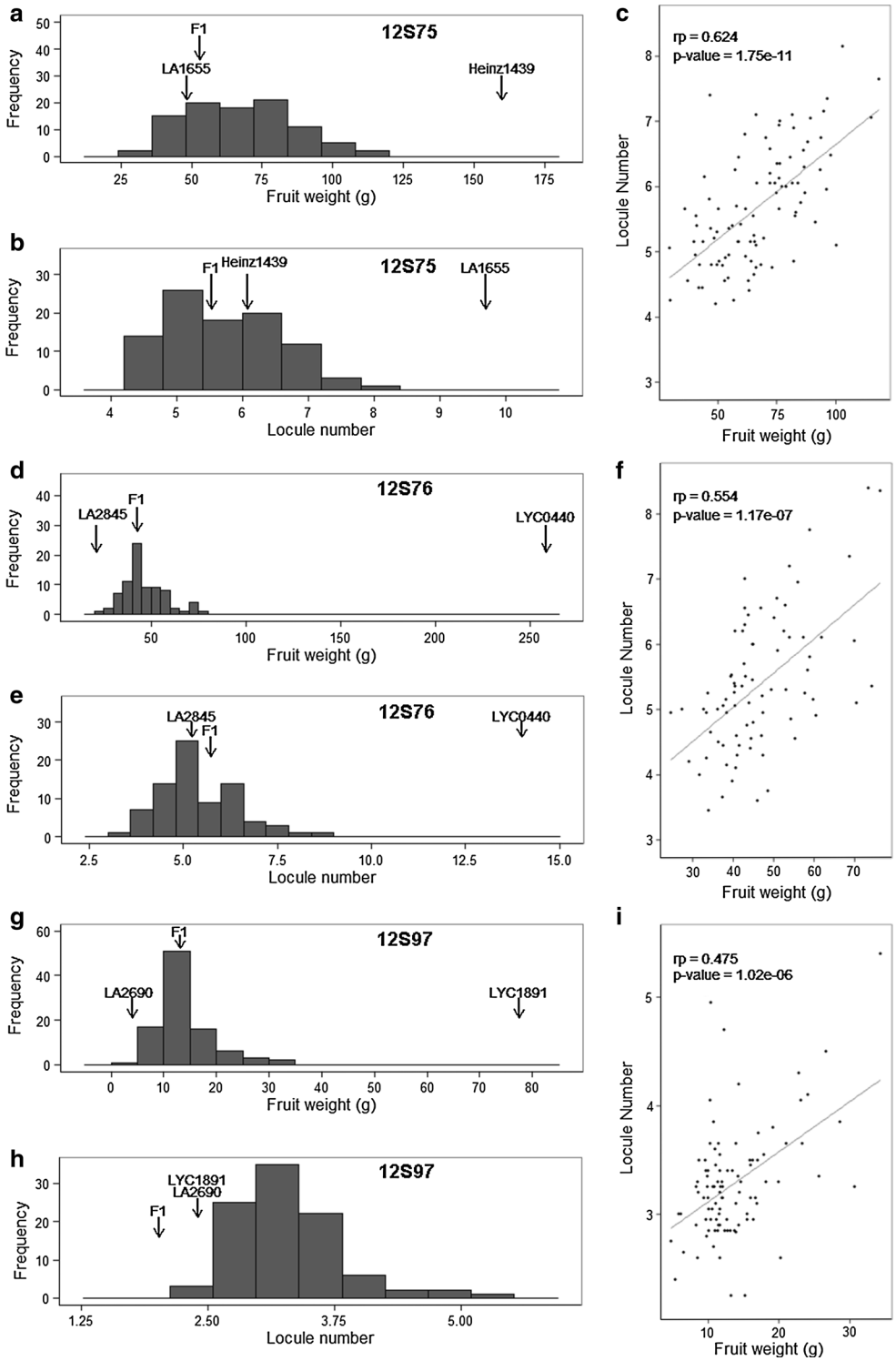

Fig. 2 Fruit weight and locule number distributions in three F_2 populations. FW (**a**) and LC (**b**) distributions and FW to LC correlation (**c**) in 12S75 population. FW (**d**) and LC (**e**) distributions and FW to LC correlation (**f**) in 12S76 population. FW (**g**) and LC (**h**) distributions and FW to LC correlation (**i**) in 12S97 population

additional QTL in these populations (Fig.S1 and S2). Even though *fw11.2* mapped close to *fas* and *fw11.3*, it was distinct from these two known fruit weight loci (Fig. 3a). The *fw1.1* was associated with the pericentromeric region of the chromosome 1 (Fig. 3b). Pericentromeric regions are notoriously large in tomato and occupy 3/4 of the chromosome (Sim et al. 2012). The greatly suppressed recombination frequency of the pericentromeric regions was why *fw1.1* spanned more than two-thirds of the entire chromosome. For LC we identified four QTL on chromosome 2, 3, 5 and 6 (Fig. 4, S3). The *lcn2.4* locus on chromosome 2 was quite distant from *lc* and other previously mapped locule number QTL (Barrero et al. 2006). Therefore, these QTL might represent hitherto unknown loci controlling locule number.

Table 2 Significant markers associated with fruit weight and locule number in each F_2 population

Population	Trait	QTL	Marker	Chr.	CC7[a]	Het.	CC9[a]	*P* value	R^2	*d*/*a*
12S139 CC9 × CC7	Fruit weight	*fw3.2*	FW3.2	3	34.136	31.972	29.150	0.078	0.054	0.132
		fw1.1	13EP405	1	29.766	33.426	30.550	0.240	na	na
		fw3.3	14EP92	3	34.046	32.395	28.298	0.036	0.066	0.426
		lcn6.1	13EP388	6	32.369	31.896	31.510	0.860	na	na
		fw11.2	13EP256	11	36.090	31.888	26.450	2.72E-05	0.177	0.128
	Monomorphic	*fw2.2*	FW2.2							
		lc	LC							
		lcn2.4	13EP372, 13EP374 and 13EP376							
		lcn5.1	13EP348, 13EP354 and 13EP356							
Population	Trait	QTL	Marker	Chr.	CC4[a]	Het.	CC9[a]	*P* value	R^2	*d*/*a*
12S141 CC9 × CC4	Fruit weight	*lc*	LC	2	31.218	31.188	30.730	0.896	na	na
		fw3.2	FW3.2	3	38.700	30.421	26.256	1.55E-07	0.349	−0.331
		fw1.1	13EP405	1	30.109	32.133	30.436	0.416	na	na
		fw3.3	13EP380	3	39.229	29.396	27.439	1.98E-07	0.360	−0.668
		lcn6.1	13EP388	6	31.263	30.454	31.749	0.654	na	na
		fw11.2	13EP254	11	33.308	31.418	29.210	0.132	na	na
	Monomorphic	*fw2.2*	FW2.2							
		lcn2.4	13EP372, 13EP374 and 13EP376							
		lcn5.1	13EP348, 13EP354 and 13EP356							
Population	Trait	QTL	Marker	Chr.	CC37[a]	Het.	CC39[a]	*P* value	R^2	*d*/*a*
12S143 CC39 × CC37	Fruit weight	*lc*	LC	2	29.285	26.412	24.164	3.74E-04	0.160	−0.12
		fw3.2	FW3.2	3	28.102	26.868	24.206	0.016	0.086	0.37
		fw1.1	13EP405	1	23.386	25.907	29.842	1.29E-07	0.292	−0.22
		fw3.3	14EP92	3	28.315	26.790	24.590	0.023	0.079	0.18
		lcn6.1	13EP390	6	27.140	26.395	26.142	0.745	na	na
		fw11.2	13EP581	11	25.317	27.652	26.320	0.100	na	na
	Monomorphic	*fw2.2*	FW2.2							
		lcn2.4	13EP372, 13EP374 and 13EP376							
		lcn5.1	13EP348, 13EP354 and 13EP356							
Population	Trait	QTL	Marker	Chr.	Heinz1439[a]	Het.	LA1655[a]	*P* value	R^2	*d*/*a*
12S75 Heinz1439 × LA1655	Fruit weight	*fw3.2*	FW3.2	3	71.994	69.863	54.406	0.001	0.252	0.76
		fw1.1	13EP401	1	71.323	67.746	62.519	0.248	na	na
		lcn2.4	13EP374	2	60.469	69.170	68.280	0.204	na	na
		fw3.3	13EP378	3	76.983	68.972	49.763	3.10E-06	0.248	0.41
		lcn5.1	13EP348	5	70.575	63.400	70.492	0.192	na	na
		lcn6.1	13EP384	6	70.680	71.014	56.719	0.021	0.071	1.05
	Locule number	*fw3.2*	FW3.2	3	5.662	5.734	5.418	0.321	na	na
		fw1.1	13EP401	1	5.470	5.682	5.857	0.244	na	na
		lcn2.4	13EP374	2	5.214	5.706	6.021	0.003	0.106	−0.22
		fw3.3	13EP378	3	6.089	5.660	5.313	0.023	0.106	−0.11
		lcn5.1	13EP348	5	5.652	5.499	6.026	0.027	0.067	−1.82
		lcn6.1	13EP384	6	5.740	5.822	5.364	0.248	na	na
	Monomorphic	*fw2.2*	FW2.2							
		lc	LC							
		fw11.2	13EP254, 13EP256, 13EP258, and 13EP581							

Table 2 continued

Population	Trait	QTL	Marker	Chr.	LYC0440[a]	Het.	LA2845[a]	*P* value	R^2	*d/a*
12S76 LA2845 × LYC0440	Fruit weight	*fw2.2*	*FW2.2*	2	53.197	46.814	39.777	0.002	0.153	0.05
		fw1.1	13EP401	1	47.573	47.646	42.384	0.484	na	na
		lcn2.4	13EP372	2	48.867	47.161	43.078	0.140	na	na
		lcn6.1	13EP386	6	47.805	46.505	44.873	0.680	na	na
	Locule number	*fw2.2*	*FW2.2*	2	5.150	5.512	5.250	0.444	na	na
		fw1.1	13EP401	1	4.904	5.428	5.664	0.186	na	na
		lcn2.4	13EP372	2	5.428	5.358	5.401	0.955	na	na
		lcn6.1	13EP386	6	5.439	5.454	5.113	0.734	na	na
	Monomorphic	*lc*	LC							
		fw3.2	FW3.2							
		fw3.3	14EP90, 14EP92, 13EP378, 13EP380, and 13EP382							
		lcn5.1	13EP348, 13EP354 and 13EP356							
		fw11.2	13EP254, 13EP256, 13EP258, and 13EP581							
Population	Trait	QTL	Marker	Chr.	LYC1891[a]	Het.	LA2690[a]	*P* value	R^2	*d/a*
12S97 LA2690 × LYC1891	Fruit weight	*fw2.2*	FW2.2	2	21.613	13.275	10.023	2.76E-09	0.481	−0.44
		fw3.2	FW3.2	3	16.377	14.301	11.360	0.009	0.123	0.17
		fw1.1	13EP401	1	18.294	13.070	11.454	0.002	0.196	−0.53
		lcn2.4	13EP374	2	16.956	13.315	13.595	0.554	na	na
		fw3.3	13EP378	3	15.245	14.439	12.203	0.010	0.051	0.47
		lcn5.1	13EP354	5	14.478	13.894	13.387	0.808	na	na
		lcn6.1	13EP388	6	14.235	13.706	13.424	0.618	na	na
	Locule number	*fw2.2*	FW2.2	2	3.567	3.319	3.054	0.006	0.102	0.03
		fw3.2	FW3.2	3	3.502	3.247	3.201	0.098	0.052	−0.69
		fw1.1	13EP401	1	3.475	3.286	3.183	0.343	na	na
		lcn2.4	13EP374	2	3.456	3.239	3.310	0.828	na	na
		fw3.3	13EP378	3	3.389	3.281	3.253	0.313	na	na
		lcn5.1	13EP354	5	3.353	3.294	3.178	0.632	na	na
		lcn6.1	13EP388	6	2.967	3.312	3.513	0.001	0.141	0.26
	Monomorphic	*lc*	LC							
		fw11.2	13EP254, 13EP256, 13EP258, and 13EP581							

Data were analyzed by single-point one-way ANOVA or Kruskal–Wallis non-parametric ANOVA

Listed below the F_2 population pedigree is the parental cross

Chr. chromosome

Het. mean fruit weight value for the heterozygous, R^2 variance explained by the associated marker, *d/a* degree of dominance, *na* not applicable

[a] Mean fruit weight value for the homozygous parental allele

To confirm that the identified regions were associated with FW or LC, we used the genome sequence data to identify SNPs for molecular marker development. For FW, the two QTL were easily confirmed and showed that each explained 18 and 29 % in phenotypic variance, respectively (Table 2; Supplementary Table S4). For LC, three out of the four QTL were confirmed with markers mapped in the entire population. However, the effect and significance of each of the QTL on the trait was relatively low (Table 2). This result suggested that increased locule number was under the control of numerous loci of relatively small effect in this population. When we associated the LC QTL with FW in the 12S75 population, *lcn3.1* was instead found to have a highly significant effect on FW. Therefore, this region was likely to carry a fruit weight locus that had a pleiotropic effect on locule number. Because of its larger effect on FW, we renamed the locule number QTL to *fw3.3*.

Single marker analysis of the newly identified QTL in all six populations

We wanted to know whether the newly identified QTL, namely *fw1.1*, *fw3.3*, *fw11.2*, *lcn2.4*, *lcn5.1* and *lcn6.1*,

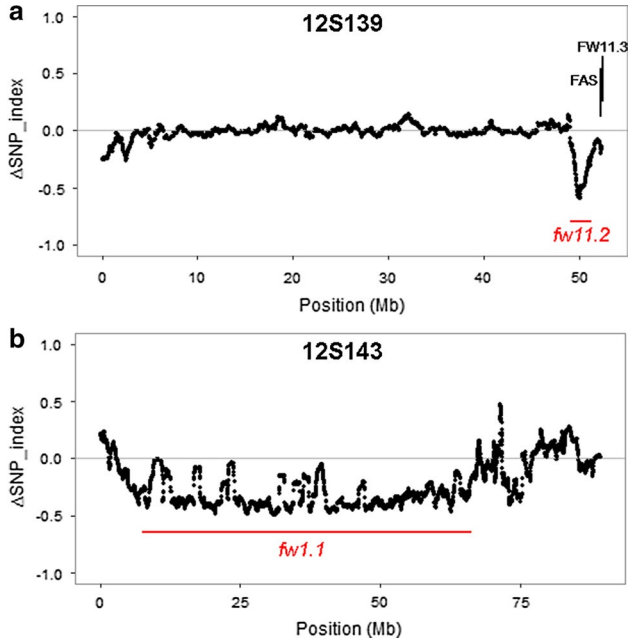

Fig. 3 Identification of fruit weight QTL in 12S139 and 12S143 populations. **a** Average values of ΔSNP-index calculated by sliding window analysis for chromosome 11 in 12S139 population. Known fruit weight and shape genes positions are indicated. **b** Average values of ΔSNP-index calculated by sliding window analysis for chromosome 1 in 12S143 population. *Red lines* indicate the approximate position of QTL (color figure online)

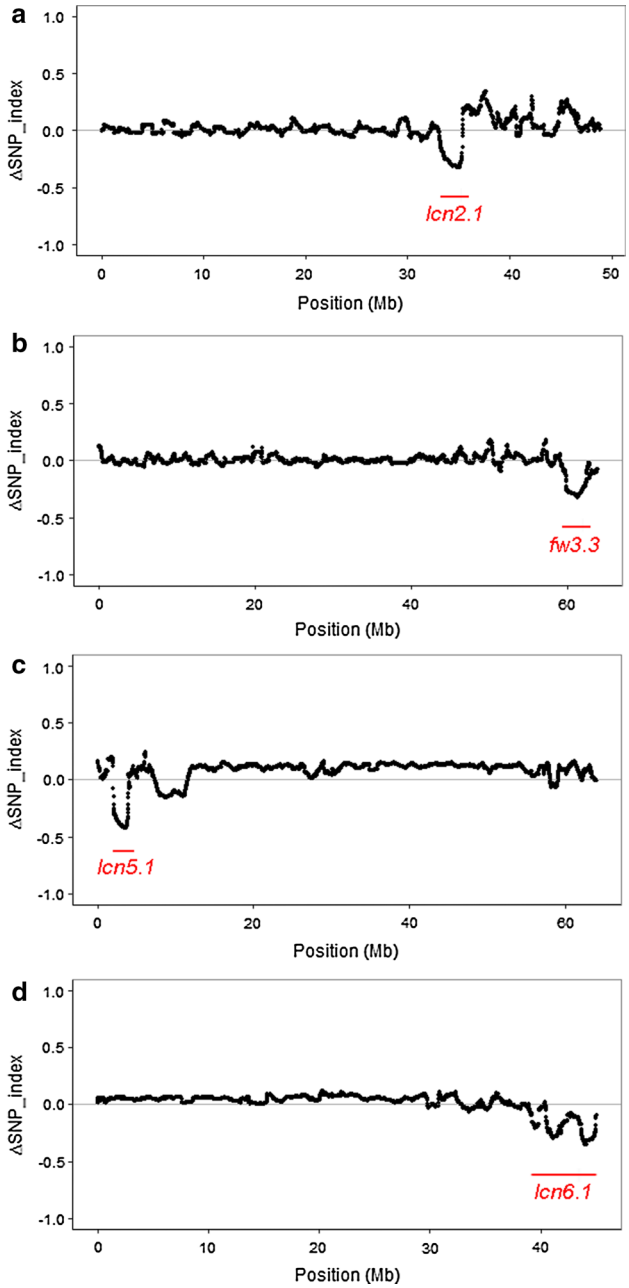

Fig. 4 Identification of locule number and fruit weight QTL in 12S75 population. **a** Average values of ΔSNP-index calculated by sliding window analysis for chromosome 2. **b** Average values of ΔSNP-index calculated by sliding window analysis for chromosome 3. **c** Average values of ΔSNP-index calculated by sliding window analysis for chromosome 5. **d** Average values of ΔSNP-index calculated by sliding window analysis for chromosome 6. *Red lines* indicate the QTL approximate position (color figure online)

explained the variation in FW and/or LC in the other populations that were not used in the QTL-seq experiment. The molecular markers that were developed for the confirmation of the loci in the QTL-seq populations were surveyed in the other populations (Supplementary Table S5). For certain QTL, all markers tested were monomorphic and therefore we could not determine with certainty that these regions were not associated with either FW or LC. For those markers that were segregating, *fw3.3* was associated with FW in the 12S141 in addition to the 12S75 population (Table 2). Even though *fw3.2* and *fw3.3* were both segregating and linked in population 12S141 (15.3 cM and 27 recombinant plants out of 95), the distance between these QTL implied that they represented distinct loci controlling fruit weight. With respect to LC, the most striking QTL was identified in the 12S97 population on chromosome 6. Whereas this QTL was not significantly associated with LC in 12S75, *lcn6.1* explained 14 % of the variance in the 12S97 population. In this population, the *fas* locus was fixed for the mutation and therefore, *lcn6.1* appeared to be a modifier of the mutation in *FAS* by increasing locule number from an average of 3.0–3.5 (Table 2).

Interaction of *fw3.2* with *fw11.2* or *fw1.1*

The effect of *fw3.2* was different in distinct genetic backgrounds. The lack of significant segregation at this locus in

two of three populations was at least in part due to other segregating loci impacting the trait (Fig. 3). To determine whether *fw11.2* and *fw1.1* showed epistatic or additive interaction with *fw3.2*, we plotted the effect of the loci in the different genetic backgrounds (Fig. 5). These data showed that the effect of *fw3.2* appeared negligible when *fw11.2* carried the homozygous cultivated allele, suggesting that these two loci were epistatic to one another. However, two-way ANOVA performed to validate the interactions were not significant for either population (*P* value of 0.2250 and 0.1705, respectively). The lack of significance may be due to the high variation observed in fruit weight in the F_2 population or that the effects of the QTL on the trait were indeed additive.

Fine mapping of *fw11.2*

To further delineate the *fw11.2* QTL, we identified recombinant F_2 plants between the markers 13EP232 and 13EP236, a 1.62-Mb region on chromosome 11. Progeny testing of nine families showed that the QTL was located between 50.6 and 51.3 Mb (Table 3). This region comprised 66 genes encoding a range of functions such as DNA-RNA binding proteins including transcription factors (12 %), other binding proteins (18 %), enzymes (13 %), and many proteins of unknown function (33 %).These fine mapping results were consistent with previous results derived from crosses between cultivated tomato and a wild relative (Supplementary Table S6).

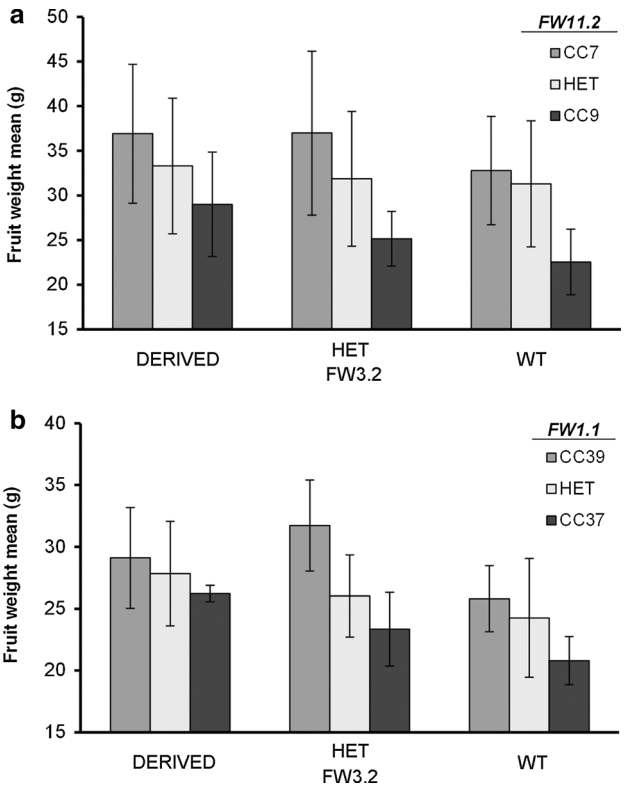

Fig. 5 Digenic interactions between fruit weight QTL using the markers significantly associated to them. **a** *FW3.2* × *FW11.2*, **b** *FW3.2* × *FW1.1*. Alleles at *FW3.2* locus are named WT for the wild-type allele; DERIVED for derived allele and HET for heterozygous. At *FW11.2* locus, *CC7* homozygous for the CC7 allele, *HET* heterozygous, *CC9* homozygous for the CC9 allele. At *FW1.1* locus, *CC37* homozygous for the CC37 allele, *HET* heterozygous, *CC39* homozygous for the CC39 allele

Discussion

Fruit weight and locule number are quantitatively inherited and many loci of both small and large effects control these traits. In the past, distantly related parents were used for the development of F_2 or RIL populations because they offer a high number of segregating loci and nucleotide polymorphisms to associate molecular markers with traits (Grandillo et al. 1999; van der Knaap and Tanksley 2003; Lin et al. 2014). Delineation of each QTL in populations derived from these wide crosses is laborious due to the need to conduct repeated backcrossing to fix background markers. It would typically take more than 5 years after initial QTL identification to delineate a QTL to a defined interval on the chromosome (e.g., van der Knaap and Tanksley 2003; Zhang et al. 2012). With the public release of the tomato reference genome (The Tomato Genome Consortium 2012) and the knowledge of a few genes that control FW and LC (Frary et al. 2000; Cong et al. 2008; Muños et al. 2011; Chakrabarti et al. 2013), genetic studies of quantitatively inherited characters can now be expedited. This is because segregation of the known genes can be avoided or taken into account in the selection of parents for populations. Moreover, whole genome sequencing ensures that the number of nucleotide polymorphisms is not a limiting factor in marker development. Thus, closer related parents can be selected which means fewer SNPs and fewer segregating QTL for the traits. In this study, we used a BSA-whole genome sequencing approach to map loci controlling FW and LC in tomato. We mapped four highly significant QTL: three FW (*fw1.1*, *fw3.3*, *fw11.2*) and one LC (*lcn6.1*) QTL. Within months of the identification of *fw11.2*, we confined its location to a 0.7 Mb region and a reduced number of candidate genes. Pursuing of this strategy should allow us to further reduce this region to few candidate genes within another year. Thus, one of the crucial approaches to the successful implementation of the QTL-seq method in tomato was the selection of parents for this study.

A critical consideration for the successful implementation of QTL-seq and gene identification is also the size of the population. Whereas nearly 100 F_2 individuals were sufficient to delineate a single QTL (12S139 and 12S143), this

Table 3 Progeny test of *fw11.2* recombinant plants

Parental pedigree	Progeny	Genotype of parent												
		SL2.40chr3	SL2.40chr11											
		58,852,276	49,734,559	49,813,364	50,274,001	50,304,751	50,437,500	50,581,429	50,695,179	50,922,056	51,071,400	51,073,674	51,336,278	51,353,247
		12EP239	13EP232	13EP244	13EP234	13EP250	13EP117	13EP254	13EP256	13EP258	13EP108	12EP900	13EP240	13EP236
12S139-1	13S90	3	1	1	1	1	1	2	–	2	2	2	3	3
12S139-7	13S91	3	1	1	1	1	2	2	2	2	2	2	2	2
12S139-60	13S78	3	1	1	1	1	1	1	2	2	2	2	2	2
12S139-64	13S92[c]	3	1	1	2	2	2	2	2	2	2	2	2	2
12S139-66	13S79	3	2	2	2	2	2	1	1	1	1	1	1	1
12S139-78	13S80	3	1	1	1	1	1	1	2	2	2	2	2	2
12S139-80	13S81	3	2	2	2	2	2	1	1	1	1	1	1	1
12S139-90	13S82	3	2	2	1	1	1	1	1	1	1	1	1	1
12S139-94	13S83	3	2	2	2	2	2	2	2	2	2	2	1	1

Parental pedigree	Progeny	Fruit weight progeny test								*P* value
		CC7[a]				CC9[b]				
		n	Mean	Variance	s.d	*n*	Mean	Variance	s. d.	
			μ	σ^2	σ		μ	σ^2	σ	two-tails
12S139-1	13S90	11	53.60	79.95	8.94	11	47.59	31.74	5.63	0.074
12S139-7	13S91	7	34.84	4.18	2.04	10	30.89	8.88	2.98	0.009
12S139-60	13S78	11	60.18	250.31	15.82	11	50.03	75.61	8.70	0.081
12S139-64	13S92[c]	8	36.03	21.88	4.68	8	33.41	36.12	6.01	0.348
12S139-66	13S79	12	37.97	50.64	7.12	12	34.48	31.42	5.61	0.196
12S139-78	13S80	12	47.55	50.31	7.09	12	39.31	9.34	3.06	0.002
12S139-80	13S81	11	43.43	92.29	9.61	11	39.83	102.90	10.14	0.402
12S139-90	13S82	11	58.32	123.35	11.11	10	57.55	239.36	15.47	0.895
12S139-94	13S83	12	33.97	22.93	4.79	10	30.04	10.43	3.23	0.039

Chromosome position of markers based on tomato genome assembly SL2.40 from SGN (www.solgenomics.net). Units bp. Marker score, 1 homozygous for CC7; 2 heterozygous; 3 homozygous for CC9

Chr. Chromosome, *n* number of tomato plants, *Mean* average fruit weight value, *sd* standard deviation

[a] Average fruit weight for parental genotype 1 (CC7)

[b] Average fruit weight for parental genotype 3 (CC9)

[c] Unexpected segregation in field test. This family needs re-progeny testing

size was too small for the reliable identification of three LC QTL segregating in one population (12S75). Moreover, better delineation of the QTL using more recombinant F_2 plants for progeny testing is desirable when the goal is to clone the underlying gene. Thus, even for the populations that segregate for a single QTL, a larger population would have resulted in a shorter interval of the location of the gene in the first generation of progeny testing (the F_3), which would speed up the eventual cloning of the gene. In contrast, when the goal is to identify closely linked markers without further gene identification, a population size of 400 is desirable. In all, a researcher needs to consider available growing space, labor costs, the ease of trait evaluation and the number of individual populations to decide on the optimal F_2 population size and reach the intended goal. For tomato FW and LC traits, a population size of approximately 200 individuals is preferred when undertaking a QTL-seq approach for the identification of the underlying genes.

It is broadly accepted that the power and precision of genetic mapping is significantly affected by the magnitude of trait heritability (Flint et al. 2005). QTL detection is easier for traits where the number of contributing loci is low and heritability is high (Moreau et al. 1998). Broad-sense heritability (H^2) has been used as an index of reliability of phenotypic selection for genetic characteristics (Holland et al. 2003). However, our data suggest that H^2 may not be a reliable parameter in predicting putative QTL. For instance, in 12S139 and 12S143 populations, broad-sense heritability for fruit weight was 0.11 and 0.83, respectively. Yet, *fw11.2* QTL was detected on 12S139 population despite of its low H^2 value. Our relatively low heritability values were likely due to the high variation in the trait values for the parents. More replication of parents and F_1 could have improved the accuracy of the heritability values in this study.

The feasibility to detect QTL by QTL-seq can also be affected by the abundance of sequencing reads at a specific locus. The minimum read depth must assure enough coverage across the genome, as this is required to detect variants, but also to adequately assess the contribution of each parent to the bulked DNAs. In tomato, the relatively small size of its genome (950 Mb; The Tomato Genome Consortium 2012) has allowed us to re-sequence DNA-bulks at 20-fold coverage, which was sufficient to detect variants (Dohm et al. 2008). Those results show that QTL-seq is a cost-efficient approach for de novo QTL identification in tomato. This conclusion is in line with conclusions from the first publication describing a QTL-seq approach in rice (Takagi et al. 2013), another crop species with a compact genome. It remains to be determined whether QTL-seq can be successfully applied in other crop species with larger genomes. But for genomes the size of tomato or smaller, the method is highly effective to identify markers linked to quantitative traits of agronomic interest.

Author contribution statement EIB conducted the analyses, research and constructed the figures and tables; JVH extracted DNA from the six populations and developed the scripts for the analyses of the QTL-Seq data; ZH developed the locule number populations and identified the *fw11.2* locus in one of the populations; EvdK initiated the project, supervised the research and analyses, and wrote the paper with EIB.

Acknowledgments This work was supported by the Agriculture and Food Research Initiative competitive grant no. 2013-67013-21229 of the USDA National Institute of Food and Agriculture and by a National Science Foundation grant IOS 0922661 to EvdK.

Conflict of interest The authors have declared that there is no conflict of interest.

References

Anastasiou E, Kenz S, Gerstung M, MacLean D, Timmer J, Fleck C, Lenhard M (2007) Control of plant organ size by *KLUH/CYP78A5*-dependent intercellular signaling. Dev Cell 13:843–856. doi:10.1016/j.devcel.2007.10.001

Andrews S (2010) FastQC a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Barrero LS, Cong B, Wu F, Tanksley SD (2006) Developmental characterization of the fasciated locus and mapping of *Arabidopsis* candidate genes involved in the control of floral meristem size and carpel number in tomato. Genome 49:991–1006. doi:10.1139/g06-059

Campoy J, Le Dantec L, Barreneche T, Dirlewanger E, Quero-García J (2014) New insights into fruit firmness and weight control in sweet cherry. Plant Mol Biol Rep. doi:10.1007/s11105-014-0773-6

Chakrabarti M, Zhang N, Sauvage C, Muños S, Blanca J, Canizares J et al (2013) A cytochrome P450 regulates a domestication trait in cultivated tomato. Proc Natl Acad Sci USA. doi:10.1073/pnas.1307313110

Cong B, Tanksley S (2006) FW2.2 and cell cycle control in developing tomato fruit: a possible example of gene co-option in the evolution of a novel organ. Plant Mol Biol 62:867–880. doi:10.1007/s11103-006-9062-6

Cong B, Barrero LS, Tanksley SD (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet 40:800–804. doi:10.1038/ng.144

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498. doi:10.1038/ng.806

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309–1321. doi:10.1016/j.cell.2006.12.006

Dohm JC, Lottaz C, Borodina T, Himmelbauer H (2008) Substantial biases in ultra-short read data sets from high-through put DNA sequencing. Nucl Acids Res 36:e105. doi:10.1093/nar/gkn425

Fatokun CA, Menancio-Hautea DI, Danesh D, Young ND (1992) Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. Genetics 132:841–846

Flint J, Valdar W, Shifman S, Mott R (2005) Strategies for mapping and cloning quantitative trait genes in rodents. Nat Rev Genet 6:271–286. doi:10.1038/nrg1576

Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88. doi:10.1126/science.289.5476.85

Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates: recent archaeo botanical insights from the old World. Ann Bot 100:903–924. doi:10.1093/aob/mcm048

Gepts P (2004) Crop domestication as a long-term selection experiment. In: Janick J (ed) Plant breeding reviews, vol 24. John Wiley & Sons, Inc, New York, pp 1–44. doi:10.1002/9780470650288.ch1

Gonzalo MJ, van der Knaap E (2008) A comparative analysis into the genetic bases of morphology in tomato varieties exhibiting elongated fruit shape. Theor Appl Genet 116:647–656. doi:10.1007/s00122-007-0698-7

Grandillo S, Ku HM, Tanksley SD (1999) Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor Appl Genet 99:978–987. doi:10.1007/s001220051405

Guo M, Rupe MA, Dieter JA, Zou J, Spielbauer D, Duncan KE, Howard RJ, Hou Z, Simmons CR (2010) Cell number regulator 1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. Plant Cell 22:1057–1073. doi:10.1105/tpc.109.073676

Holland JB, Nyquist WE, Cervantes-Martínez CT (2003) Estimating and interpreting heritability for plant breeding: an update. In: Janick J (ed) Plant breeding reviews, vol 22. John Wiley & Sons, Inc., New York, pp 9–112. doi:10.1002/9780470650202.ch2

Huang Z, van der Knaap E (2011) Tomato *fruit weight 11.3* maps close to fasciated on the bottom of chromosome 11. Theor Appl Genet 123:465–474. doi:10.1007/s00122-011-1599-3

Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common bean. Crop Sci 36:1037–1045. doi:10.2135/cropsci1996.0011183X003600040037x

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Meth 9:357–359. doi:10.1038/nmeth.1923

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP (2009) The sequence alignment/map format and SAM tools. Bioinformatics 25:2078–2079. doi:10.1093/bioinformatics/btp352

Libault M, Zhang X-C, Govindarajulu M, Qiu J, Ong YT, Brechenmacher L, Berg RH, Hurley-Sommer A, Taylor CG, Stacey G (2010) A member of the highly conserved FWL (tomato *FW2.2*-like) gene family is essential for soybean nodule organogenesis. Plant J 62:852–864. doi:10.1111/j.1365-313X.2010.04201.x

Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z et al (2014) Genomic analyses provide insights into the history of tomato breeding. Nat Genet 46:1220–1226. doi:10.1038/ng.3117

Lippman Z, Tanksley SD (2001) Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom Genet 158:413–422

Mather K, Jinks JL (1982) Biometrical genetics, the study of continuous variation, 3rd edn. Chapman and Hall, London

Maughan PJ, Maroof MAS, Buss GR (1996) Molecular-marker analysis of seed-weight: genomic locations, gene action, and evidence for orthologous evolution among three legume species. Theor Appl Genet 93:574–579. doi:10.1007/BF00417950

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The genome analysis toolkit: a map reduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. doi:10.1101/gr.107524.110

Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet 14:840–852. doi:10.1038/nrg3605

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832. doi:10.1073/pnas.88.21.9828

Monforte AJ, Diaz AI, Caño-Delgado A, van der Knaap E (2014) The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon. J Exp Bot 65:4625–4637. doi:10.1093/jxb/eru017

Moreau L, Charcosset A, Hospital F, Gallais A (1998) Marker-assisted selection efficiency in populations of finite size. Genetics 148:1353–1365

Muños S, Ranc N, Botton E, Bérard A, Rolland S, Duffé P et al (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near *WUSCHEL*. Plant Physiol 156:2244–2254. doi:10.1104/pp.111.173997

Olmstead JW, Iezzoni AF, Whiting MD (2007) Genotypic differences in sweet cherry fruit size are primarily a function of cell number. J Amer Soc Hort Sci 132:697–703

Olsen KM, Wendel JF (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. Annu Rev Plant Biol 64:47–70. doi:10.1146/annurev-arplant-050312-120048

Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. Nature 457:843–848. doi:10.1038/nature07895

R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

Ranc N, Muños S, Santoni S, Causse M (2008) A clarified position for *Solanum lycopersicum* var. cerasiforme in the evolutionary history of tomatoes (solanaceae). BMC Plant Biol 8:130. doi:10.1186/1471-2229-8-130

Rodríguez GR, Muños S, Anderson C, Sim S-C, Michel A, Causse M, Gardener BBM, Francis D, van der Knaap E (2011) Distribution of SUN OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol 156:275–285. doi:10.1104/pp.110.167577

Sandlin K, Prothro J, Heesacker A, Khalilian N, Okashah R, Xiang W et al (2012) Comparative mapping in watermelon [*Citrullus lanatus* (Thunb.) Matsum. et Nakai]. Theor Appl Genet 125:1603–1618. doi:10.1007/s00122-012-1938-z

Sim S-C, Durstewitz G, Plieske J, Wieseke R, Ganal MW, Van Deynze A, Hamilton JP, Buell CR, Causse M, Wijeratne S, Francis DM (2012) Development of a large snp genotyping array and generation of high-density genetic maps in tomato. PLoS One 7:e40563. doi:10.1371/journal.pone.0040563

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C et al (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J 74:174–183. doi:10.1111/tpj.12105

Tanksley SD (2004) The genetic developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell 16:S181–S189. doi:10.1105/tpc.018119

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641. doi:10.1038/nature11119

Timmerman-Vaughan GM, McCallum JA, Frew TJ, Weeden NF, Russell AC (1996) Linkage mapping of quantitative trait loci controlling seed weight in pea (*Pisum sativum* L.). Theor Appl Genet 93:431–439. doi:10.1007/BF00223187

van der Knaap E, Tanksley SD (2003) The making of a bell pepper-shaped tomato fruit: identification of loci controlling

fruit morphology in yellow stuffer tomato. Theor Appl Genet 107:139–147. doi:10.1007/s00122-003-1224-1

van der Knaap E, Chakrabarti M, Chu Y, Clevenger JP, Illa-Berenguer E, Huang Z, Keyhaninejad N, Mu Q, Sun L, Wang Y, Wu S (2014) What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. Front Plant Sci 5:227. doi:10.3389/fpls.2014.00227

Yuan XJ, Li XZ, Pan JS, Wang G, Jiang S, Li XH et al (2008) Genetic linkage map construction and location of QTLs for fruit-related traits in cucumber. Plant Breed 127:180–188. doi:10.1111/j.1439-0523.2007.01426.x

Zhang N, Brewer M, van der Knaap E (2012) Fine mapping of *fw3.2* controlling fruit weight in tomato. Theor Appl Genet 125:273–284. doi:10.1007/s00122-012-1832-8

Zuo J, Li J (2014) Molecular genetic dissection of quantitative trait loci regulating rice grain size. Annu Rev Genet 48:99–118. doi:10.1146/annurev-genet-120213-092138

Zygier S, Chaim AB, Efrati A, Kaluzky G, Borovsky Y, Paran I (2005) QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theor Appl Genet 111:437–445. doi:10.1007/s00122-005-2015-7