Theor Appl Genet (2015) 128:1747–1763
DOI 10.1007/s00122-015-2544-7

ORIGINAL PAPER

QTL mapping in multiple populations and development stages reveals dynamic quantitative trait loci for fruit size in cucumbers of different market classes

Yiqun Weng[1,2] · Marivi Colle[3] · Yuhui Wang[1] · Luming Yang[1] · Mor Rubinstein[4] · Amir Sherman[4] · Ron Ophir[4] · Rebecca Grumet[3]

Received: 23 March 2015 / Accepted: 18 May 2015 / Published online: 6 June 2015

Abstract

Key message **QTL analysis in multi-development stages with different QTL models identified 12 consensus QTLs underlying fruit elongation and radial growth presenting a dynamic view of genetic control of cucumber fruit development.**

Abstract Fruit size is an important quality trait in cucumber (*Cucumis sativus* L.) of different market classes. However, the genetic and molecular basis of fruit size variations in cucumber is not well understood. In this study, we conducted QTL mapping of fruit size in cucumber using F_2, F_2-derived F_3 families and recombinant inbred lines (RILs) from a cross between two inbred lines Gy14 (North American picking cucumber) and 9930 (North China fresh market cucumber). Phenotypic data of fruit length and diameter were collected at three development stages (anthesis, immature and mature fruits) in six environments over 4 years. QTL analysis was performed with three QTL models including composite interval mapping (CIM), Bayesian interval mapping (BIM), and multiple QTL mapping (MQM). Twenty-nine consistent and distinct QTLs were detected for nine traits from multiple mapping populations and QTL models. Synthesis of information from available fruit size QTLs allowed establishment of 12 consensus QTLs underlying fruit elongation and radial growth, which presented a dynamic view of genetic control of cucumber fruit development. Results from this study highlighted the benefits of QTL analysis with multiple QTL models and different mapping populations in improving the power of QTL detection. Discussion was presented in the context of domestication and diversifying selection of fruit length and diameter, marker-assisted selection of fruit size, as well as identification of candidate genes for fruit size QTLs in cucumber.

Communicated by S. Huang.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2544-7) contains supplementary material, which is available to authorized users.

✉ Yiqun Weng
weng4@wisc.edu; yiqun.weng@ars.usda.gov

✉ Rebecca Grumet
grumet@msu.edu

[1] Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA

[2] USDA-ARS Vegetable Crops Research Unit, Horticulture Department, University of Wisconsin, Madison, WI 53706, USA

[3] Horticulture Department, Michigan State University, East Lansing, MI 48824, USA

[4] Department of Fruit Trees Sciences, Institute of Plant Sciences, Agricultural Research Organization, Volcani Center, Rishon Lezion, Israel

Introduction

Cucumber, *Cucumis sativus* L. ($2n = 2x = 14$), is native to southern Asia and was domesticated from its wild relative *C. sativus* var. *hardwickii* (Candolle 1959; Sebastian et al. 2010; Yang et al. 2012). Cucumber has been cultivated in India for at least 3000 years spreading eastward to China ~2000 years ago and westward to Europe 700–1500 years ago (Whitaker and Davis 1962; Keng 1974; Paris et al. 2012). While the wild cucumber usually bears small, round or spheroid fruits (3–5 cm in diameter), domesticated cucumbers in general have much larger fruits but exhibit significant variations in fruit size, shape and weight. In cultivated cucumber, breeding efforts have resulted in different

market classes adapting to local environments, consuming habits, or processing specifications. Major cucumber market classes with significant commercial production include the European greenhouse (Dutch/English cucumber or European Long) and pickling (European short) cucumbers, the North and South China fresh market cucumbers, North American slicing and pickling (processing) cucumbers, the Beit Alpha (mini) cucumber, and the Japanese Long cucumber. Each market class of cucumbers has unique commercial standards in fruit length and diameter making fruit size an important fruit quality trait for cucumber breeding. For example, ideal US pickling cucumbers have length-by-diameter (*L*/*D*) ratios of approximately 3.0 at harvest with blocky shape, lightly colored skin, warts or tubercles, and an exocarp permeable to brining salt. Small seed cavity size and thick fruit flesh are important in reducing placental hollowness and carpel separation (Kennard and Havey 1995). The premium grade North China fresh market cucumbers should have length of 25–30 cm and diameter of 2.5–3.1 cm (Zhou et al. 2005). Meanwhile, the Canadian Grade #1 parthenocarpic greenhouse cucumber (Dutch type) must have $L \geq 28.0$ cm and $D \geq 4.1$ cm (Canadian Food Inspection Agency Cucumber Standards, http://www.inspection.gc.ca/).

Fruit size measured by fruit length (*L*), diameter (*D*) or the ratio (*L*/*D*), is quantitative in nature. Genetic variances and heritabilities have been analyzed for these traits with *L* in general having a relatively high narrow-sense heritability (e.g., Smith et al. 1978; Strefeler and Wehner 1986; Owens et al. 1985). Kennard and Havey (1995) were the first to conduct QTL mapping to identify QTLs for fruit quality traits (*L*, *D*, seed cavity size, etc.,) in cucumber. With RIL populations developed from a cross between North China type and European greenhouse cucumbers, Yuan et al. (2008) identified 38 QTLs for yield and quality traits including immature fruit length (FL), diameter (FD), flesh thickness (FTH), and seed cavity size (SCD). Miao et al. (2011) identified 10 QTLs for FL, FD, mature fruit length (MFL) and diameter (MFD). More recently, using RIL populations developed from cultivated × wild or semi-wild cucumber inbred lines, 4 QTLs for FL, and 8 QTLs for MFL and MFD were identified by Wang et al. (2014) and Bo et al. (2015), respectively. With an $F_{2:3}$ population derived from CC3 (North China type) × NC76 (US slicing type), Wei et al. (2014) identified 6 QTLs for FL and MFL with one major-effect QTL for both FL and MFL in chromosome 3 explaining ~45 % of phenotypic variations. While these studies have provided some insights into the genetic basis of cucumber fruit development, the relationships of QTLs detected from different studies are not clear. A global picture of cucumber fruit development in different stages is lacking. The resolution of genetic maps in target QTL regions also needs to be improved.

Cucumber fruits are typically harvested while immature, when the fruit is in the middle-to-late phase of rapid fruit growth (~2 weeks after anthesis). Knowledge of genetic and molecular mechanisms of early fruit development is important to understand the formation of fruit yield and quality in cucumber production. A few studies have examined the physiological and molecular biological basis of early fruit development in cucumber. Cucumber fruit is developed from an enlarged inferior ovary. Like many other horticultural crops, early development of cucumber fruit can be divided into three phases: development of the ovary, cell division, and subsequent cell expansion (Gillaspy et al. 1993). The weight of a cucumber fruit can increase by ~200 times in 2 weeks after anthesis (Boonkorkaew et al. 2008; Fu et al. 2010). The transition from rapid cell division to cell expansion occurs ~3–5 days after anthesis (Yang et al. 2013a). In the pickling cucumber cultivar 'Vlaspik', fruit elongation begins almost immediately after pollination, with the most rapid increase occurring approximately 4–12 days post-pollination (dpp); the rapid increase in cell size mirrors the rapid increase in fruit length (Ando and Grumet 2010). The increase in fruit diameter lags somewhat behind the length which occurs primarily between 4 and 16 dpp (Ando and Grumet 2010; Yang et al. 2013a). Cell division and expansion are largely completed by 12–16 dpp with some variation depending on the cultivar and season (Ando et al. 2012; Yang et al. 2013a).

To gain insights into the molecular events in early fruit development under rapid fruit growth, Ando et al. (2012) conducted combined morphological analysis with transcriptome profiling of young cucumber fruits at five stages from anthesis through the end of exponential growth. Clustering analysis separated the expression patterns of transcripts by fruit age into three groups corresponding with cell division/pre-exponential growth (0 and 4 dpp), peak exponential expansion (8 dpp), and late/post-exponential expansion stages of growth (12 and 16 dpp); each stage exhibits a unique set of transcripts associated with prevailing biological processes providing us a dynamic view of early fruit development in cucumber. Yang et al. (2013a) characterized the expression of cucumber kinesin genes during early fruit development and revealed the roles of seven kinesin genes (*CsKF1* to *CsKF7)* in exponential cell production and enlargement in cucumber fruit. Jiang et al. (2015) compared the transcriptomes in two near isogenic lines of cucumber with contrasting fruit length and found that microtubule and cell cycle related genes were dramatically activated in the long fruit, and transcription factors were implicated in the fruit length regulation in cucumber. However, the genetic basis of early fruit development in cucumber is largely unknown. We have limited knowledge

on the connections between genetics and physiology or molecular biology of early fruit development in cucumber. From a cucumber breeding perspective, understanding the genetic basis of and identification of molecular markers for fruit size will be important for marker-assisted cucumber breeding. Therefore, the main objective of the present study was to identify QTLs associated with fruit size-related traits in cucumber fruit development.

We previously developed a 783-locus simple sequence repeat (SSR)-based linkage map with an F_2 mapping population derived from two cucumber inbred lines Gy14 and 9930 (Yang et al. 2012, 2013b). Using a recombinant inbred line (RIL) population from the same cross, we also developed a SNP (single nucleotide polymorphism)-based map that contains 11,156 SNP loci (Rubinstein et al. 2015). In the present study, we phenotyped three populations (F_2, F_3 families and RILs) from the Gy14 × 9930 cross in six environments over 4 years for nine fruit size-related traits including the length and diameter of ovary, immature and mature fruits, as well as the ovule number, seed cavity size, and flesh thickness of mature fruits. QTLs underlying these traits were identified with three QTL models. Integration of QTL information from the present and previous studies allowed us to identify 12 consensus QTLs for fruit size in cucumber.

Materials and methods

Plant materials

The two cultivated cucumber inbred lines used to develop segregating populations for this study were Gy14 and 9930. The monoecious 9930 is a typical North China fresh market type cucumber with slim, spiny, dark green immature fruits, and yellow mature fruits; the ratio of length to diameter (*L*/*D*) at the commercial harvest stage is usually >5. The gynoecious Gy14 is a typical North American pickle cucumber that bears blocky and relatively smooth fruits, with a few large spines; it has light green immature and creamy mature fruits with *L*/*D* ratio around 3.0 at the commercial harvest stage (Fig. 1). A single F_1 plant from the cross between Gy14 (female parent carrying *a* allele at each marker locus) and 9930 (male parent carrying *b* allele at each marker locus) was self-pollinated to produce 150 F_2, and F_2-derived F_3 families. Recombinant inbred lines (RILs) were developed through single seed descent. A total of 123 F_6 and 141 F_7 RILs were used during the 2011 and 2012 field trials, respectively, in this study.

Genotyping and linkage map development

Using 92 Gy14 × 9930 F_2 plants, we previously developed a high-density cucumber genetic map with 783 SSR or SNP

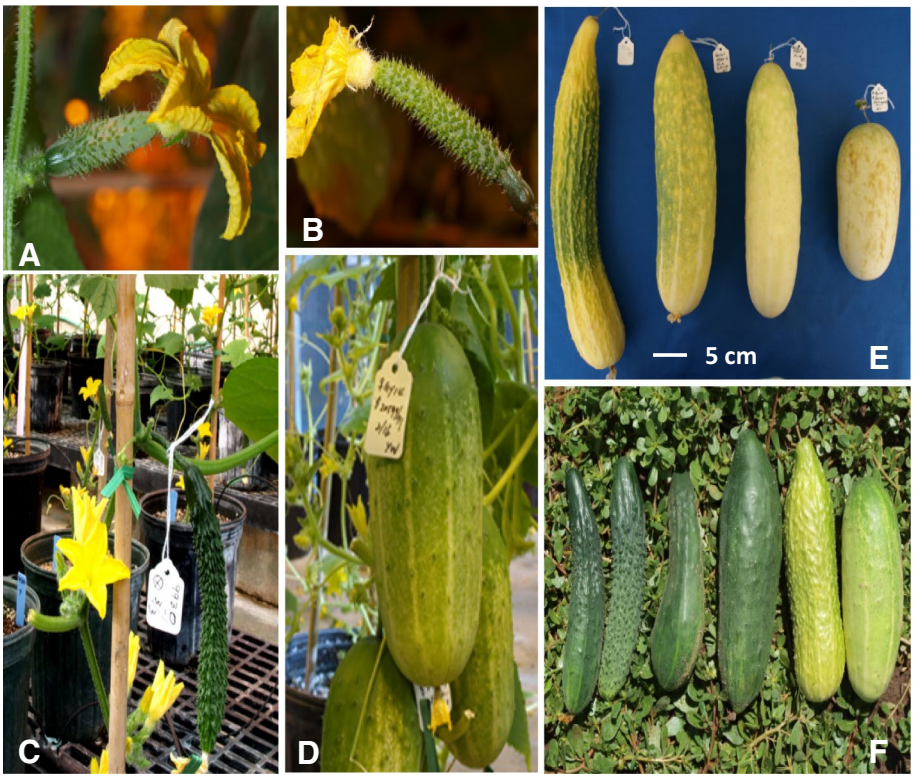

Fig. 1 Fruit morphology of Gy14, 9930 cucumbers and their derivatives. **a** and **b** are ovaries of Gy14 and 9930, respectively. **c** is an immature fruit of 9930. **d** is a mature fruit of Gy14 × 9930 F_1. **e** and **f** are fruits of the F_2 and RIL populations, respectively

markers (Cavagnaro et al. 2010; Yang et al. 2013b). This F_2 was the same population used for fruit size data collection herein. In the present study, this 783-locus map was revised in several aspects: one marker (bin) was kept when there were multiple co-segregating markers and markers with segregating distortion, >5 % missing data, or with dubious double crossovers were removed. Thus, 542 markers were selected for linkage analysis resulting in a new map of 687.8 cM map length with an average of 1.3 cM marker interval. Details of this SSR-based F_2 map are presented in supplemental Table S1. The map was used for QTL analysis with F_2 and F_2-derived F_3 fruit size data.

Using 129 Gy14 × 9930 F_7 RILs (the same set as used in 2011 phenotyping), we also developed a ultra-high-density map (Rubinstein et al. 2015) which contains 11,156 SNP loci spanning 600 cM in seven linkage groups. From this map, we selected 458 SNPs that were evenly distributed across seven chromosomes to develop a new map with 605.9 cM map length and 1.3 cM map interval (supplemental Table S2), which was used for QTL analysis of RIL data in this study. Summary statistics of the SSR-based F_2 map and the SNP-based RIL map are presented in supplemental Table S3.

The Gy14 and 9930 cucumber genomes have been sequenced and assembled (Huang et al. 2009; Yang et al. 2012). The physical locations of SSR or SNP markers in the Gy14 and 9930 scaffold and whole genome assemblies were used to verify their genetic map locations. Comparisons of the QTL locations from the F_2, F_3 and RIL data sets were also based on the scaffold or draft genome assembly locations of those associated markers.

Phenotypic data collection

Phenotypic data were collected in six environments over 4 years (2009, 2010, 2011, and 2012) with F_2 individual plants, F_3 families, as well as F_6 or F_7 RILs at three locations. Gy14, 9930 and their F_1 were included in all experiments. Details of the six experiments, WI2009, WI2010, WI2011, WI2012, MI2011 and MI2012, are presented in Table 1. Briefly, the experiment WI2009 was conducted in the Walnut Street Greenhouses of the University of Wisconsin at Madison with 150 Gy14 × 9930 F_2 plants. One self-pollinated fruit was allowed on each plant. All other five experiments were conducted in open fields in two locations: University of Wisconsin Agricultural Research Station at Hancock, WI (F_3 and RILs) over 3 years (2010, 2011 and 2012) and at the Horticulture Teaching and Research Center of Michigan State University at East Lansing, MI (RILs only) in 2011 and 2012. Both locations used the same randomized complete block design (RCBD) consisting of three blocks with five plants per plot. Row spacing was 1.5 m, and plants within a row (plot) were spaced at 60–75 cm. Pollination was facilitated by bees. Local standard commercial production guidelines were followed for insect/weed control and fertilization. In general, 3–10 fruits could be set on each plant under the field conditions.

For the WI2011, MI2011 and MI2012 trials, fruit length (*L*) and diameter (*D*) data were collected at three development stages: ovary length (OvL) and diameter (OvD) at anthesis; immature fruit length (FL) and diameter (FD) 10–12 days post-pollination (dpp), and mature fruit length (MFL) or diameter (MFD) at approximately 35 dpp. For the WI2009, WI2010 and WI2012 experiments, only MFL and MFD data were collected. In addition to fruit length and diameter data at the three stages, ovule number (OvN) per fruit was determined in MI 2011 and 2012; seed cavity size (SCD) and flesh thickness (FTH) were measured in mature fruits in MI2012. SCD was the diameter of endocarp (soft tissue in the center of fruit). For counting of ovule numbers, the fruit was cut longitudinally and number of ovules (seeds) was counted. In F_3 and RIL populations, for each replication, at least five fruits were measured for each trait, and the means per replication were used in statistical analysis.

Statistical analysis of data

Statistical analysis of phenotypic data from all experiments was performed in R (Version 3.1.2, http://www.r-project.org/). For RILs, since multiple-year and multi-location data were collected, variance components were estimated using the restricted maximum likelihood (REML) method (with the *lme4* and *lmerTest* packages in R) (Bates 2010; Kuznetsova et al. 2013). Analysis of variance (ANOVA) was performed to estimate the genetic and environment effects on each trait with different models. For SCD and FTH which only had one-year, one-location data (MI2012), the following model was used: $R_{ir} = \mu + G_i + r_r + \varepsilon_{ir}$. For OvL and OvD, FL and FD with data from three environments (Table 1), a mixed model was used: $R_{ijk} = \mu + G_i + Y_j + r_{jr} + GY_{ij} + \varepsilon_{jir}$. For MFL and MFD, the following mixed model was used: $R_{ijkr} = \mu + G_i + Y_j + L_k + r_r\,(Y_j\,L_k) + GY_{ij} + GL_{ik} + YL_{jk} + GYL_{ijk} + \varepsilon_{ijkr}$, where R is observed value for a given trait, μ grand mean, G genotype, Y year, L location, r block effects, and ε random error. In these models, genotypes were considered as fixed effects, and year, location, and replications as random effects. For OvL, FL, MFL, OvD, FD and MFD, best linear unbiased predictors (BLUPs) were extracted from these models for each genotype and trait and used for the QTL analyses.

Broad sense heritability estimates were calculated from variance components. For a single location, heritability on the plot level was estimated using $h^2 = (\sigma_G^2/\sigma_G^2 + \sigma_{GY}^2 + \sigma_{GL}^2 + \sigma_{GLY}^2 + \sigma_\varepsilon^2)$. The heritability for means across environments on one entry-mean basis

Table 1 Summary of six environments used to assess fruit size data in Gy14 × 9930 F_2, F_3 family and RIL populations over 4 years

Experiments	Location	Season	Experimental design	Traits
WI2009 (F2)	Madison, Wisconsin	2009 fall (August–November), greenhouse trial	150 F_2 plants grown in one greenhouse with one self-pollinated fruit per plant. 92 plants were used for linkage mapping and QTL analysis	MFL, MFD
WI2010 (F3)	Hancock, Wisconsin	2010 summer (June–September), field trial	RCBD: 123 F_3 families (entries), three replications with five plants per rep grown in open field. Data collected from 5–10 fruits per replication per entry	MFL, MFD
WI2011 (F_6 RIL)	Hancock, Wisconsin	2011 summer (June–September), field trial	RCBD: 140 F_6 RILs. three replications with five plants per rep grown in open field. Data collected from 5–10 fruits per plant at three stages	OvL, OvD, FL, FD, MFL, MFD
WI2012 (F_7 RIL)	Hancock, Wisconsin	2012 summer (June–September), field	Same as WI2011 but with 140 F_7 RILs. Only mature fruit data were collected	MFL, MFD
MI2011 (F_6 RIL)	East Lansing, Michigan	2011 summer (June–September), field trial	Same as WI2011	OvL, OvD, FL, FD, MFL, MFD
MI2012 (F_7 RIL)	East Lansing, Michigan	2012 summer (June–September), field trial	Same as WI2012, but data were collected in three stages of fruit development.	OvL, OvD, FL, FD, MFL, MFD, FTH, SCD, OvN

RCBD randomized complete block design. *OvL* ovary length, *OvD* ovary diameter, *FL* fruit length at 12 dpp, *FD* fruit diameter at 12 dpp, *MFL* mature fruit length (35 dpp), *MFD* mature fruit diameter, *FTH* mature fruit flesh thickness and *SCD* seed cavity size

was estimated using the following formula (Marwede et al. 2004): $h^2 = \sigma_G^2/(\sigma_G^2 + \frac{\sigma_{GY}^2}{Ry} + \frac{\sigma_{GL}^2}{Rl} + \frac{\sigma_{GLY}^2}{RyRl} + \frac{\sigma_\varepsilon^2}{RyRlT})$, where σ_G^2 was the genotypic variance, σ_{GY}^2, σ_{GL}^2, and σ_{GLY}^2 were the genotype × environment interaction ($G \times E$) variance, and σ_ε^2 was the residual variance, respectively. Ry, Rl and T were number of years, locations and replicates, respectively. When estimating the heritability of each trait, the $G \times E$ term (σ_{GY}^2, σ_{GL}^2, and σ_{GLY}^2) was kept only when significant interactions were present.

Since there were no significant location and year effects found using ANOVA of RIL data for any traits under investigation, grand means across locations and years for each trait were used to calculate the Spearman's rank correlation coefficient (r_s) among different traits. Calculation of r_s was based on measurement of individual plants for F_2 (WI2009 Experiment), family means for F_3 (WI2010 Experiment), and RIL means for 2011 and 2012 experiments.

QTL analysis

QTL analysis was performed in the R/qtl software package (http://www.rqtl.org/). The individual F_2 plant data, means of each F_3 family, the grand mean or BLUP of each RIL across all environments were used for QTL analysis. QTL detection included preliminary QTL identification using function '*mqmscan*' followed by QTL modeling. The function of '*addqtl*' and '*addint*' were used to scan additional QTLs or QTL pairs. Then function '*refineqtl*' was used to refine QTL locations in the context of multiple QTL model (MQM) (Broman et al. 2003; Arends et al. 2010). The significance of each QTL interval was tested by a likelihood-ratio statistic (LOD). The LOD threshold for declaring significant QTLs for each trait was determined using a permutation test with 1000 repetitions ($P = 0.05$). In MQM analysis, only QTLs detected with a LOD score above the LOD threshold were reported. Information reported by MQM for each QTL included its chromosome location, LOD support value, phenotypic variation (R^2) explained by the QTL, as well as the additive and dominance effects (for F_2 and F_3 data). The support intervals for the map locations of QTL were calculated using a 1.5 LOD drop interval. For detection of epistasis, the genome wide two-locus epistasis interactions were surveyed using the 'interactions' function. The P value cutoff was set as 1×10^{-5} for the total effects and 0.01 for the interaction effects of two examined loci.

To examine the robustness of different models in QTL detection, QTL analysis was also performed with composite interval mapping (CIM) (Zeng et al. 1999) and Bayesian Interval Mapping (BIM) (Yandell et al. 2007) models. The BIM was implemented using the R/qtlbim package (V2.0.7, available at http://www.cran.r-project.org/web/packages/qtlbim/). Bayes factor (BF) profiles were used to

estimate the number of QTLs and the potential interaction among QTLs. Among all possible QTL patterns, the one with the highest posterior mean was chosen as the 'best' pattern for further QTL modeling (Yandell et al. 2007). The BIM model does not provide a LOD threshold and LOD support interval for detected QTLs.

The QTL was named according to its chromosome location and trait name. For example, *qFL1.1* and *qMFL3.1* designated the first QTL for the length of immature and mature fruits in cucumber chromosomes 1 and 3, respectively. A QTL that explained more than 10 % observed phenotypic variations ($R^2 > 10$ %) was considered a major-effect QTL.

A set of consensus QTLs for fruit size (FS) was inferred by integrating the information of QTLs detected with six of the nine traits in the three populations (F_2, F_3 and RIL) via the three QTL models (BIM, CIM and MQM). A consensus QTL was a synopsis of multiple QTLs with the following criteria: (1) were major-effect QTLs ($R^2 > 10$ %) in at least one environment or population; (2) were detected in at least two environments or by at least two QTL models with either grand means or BLUPs that were supported by above-threshold LOD scores, and (3) shared the same or had overlapped 1.5-LOD intervals.

Results

Phenotypic variations of fruit size in Gy14 × 9930 populations

The phenotypic means, standard derivation and range of nine traits from all six experiments are presented in Table 2. A box-plot to depict the scope of genetic variations of fruit length and diameter in F_2, F_3 and RIL populations is shown in Fig. 2. Results from ANOVA and variance component analysis for all traits from the RIL population are presented in supplemental Tables S4 and S5 (online materials), respectively. No significant effects of year, location and genotype × year interactions were found for any trait examined (Table S4) suggesting expression of these traits were highly stable in these environments. Therefore, RIL grand means across the four environments for all traits were used in subsequent data analysis. The frequency distribution of each trait in F_2, F_3 and RIL populations is illustrated in supplemental Fig. S1 (A–K).

The dynamics of fruit length and diameter during fruit development were clearly different. In all environments and at three development stages, the fruit length of Gy14 was consistently shorter than that of 9930 with that of F_1 in the middle (Table 2). The OvL, FL and MFL among F_2, F_3 and RILs all exhibited a normal distribution typical for quantitative traits with the two parents, in general, representing

Table 2 Phenotypic means and range of fruit size-related traits in two parental lines (Gy14 and 9930), their F_1, F_2, F_3, and recombinant inbred lines (RILs) in six environments (WI2009, WI2010, WI2011, WI2012, MI2011, and MI2012)

Traits	GY14	F1	9930	F2 (WI2009, $n = 150$)		F3 (WI2010, $n = 140$)		RIL (WI2011, $n = 123$)		RIL (WI2012, $n = 140$)		RIL (MI2011, $n = 123$)		RIL (MI2012, $n = 140$)	
	Mean*	Mean*	Mean*	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range
OvL	2.1	2.8	3.8	n/a	n/a	n/a	n/a	3.7 ± 0.6	2.5–5.4	n/a	n/a	2.8 ± 0.5	1.9–4.6	2.7 ± 0.4	1.7–4.6
OvD	0.6	0.6	0.7	n/a	n/a	n/a	n/a	0.9 ± 0.1	0.7–1.3	n/a	n/a	0.5 ± 0.1	0.4–0.6	0.5 ± 0.1	0.4–0.9
FL	14.3	20.2	21.8	n/a	n/a	n/a	n/a	17.5 ± 3.3	5.5–24.9	n/a	n/a	16.9 ± 1.8	10.1–21.7	17.5 ± 2.2	11.7–26.5
FD	5.0	5.5	5.6	n/a	n/a	n/a	n/a	4.3 ± 0.6	2.3–6.3	n/a	n/a	4.2 ± 0.5	2.5–6.2	4.3 ± 0.4	3.3–5.1
MFL	18.9	25.5	28.8	23.8 ± 4.0	14.5–46.8	21.5 ± 2.7	15.2–29.9	24.9 ± 3.6	17.6–34.0	25.6 ± 3.8	19.3–38.4	24.3 ± 3.1	18.5–31.1	26.7 ± 3.4	20.1–35.5
MFD	7.9	7.2	5.4	5.7 ± 0.7	4.5–8.0	6.7 ± 0.4	5.5–7.8	7.3 ± 0.6	5.8–8.9	7.4 ± 0.6	6.0–9.7	6.6 ± 0.5	5.4–7.8	7.1 ± 0.6	5.6–8.5
OvN	29.6	46.8	48.6	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	42.2 ± 4.9	29.9–59.7
SCD	4.5	4.1	3.1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	4.0 ± 0.4	2.8–5.2
FTH	1.7	1.7	1.1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	1.5 ± 0.3	0.7–2.3

* Based on data across all 4 years in all environments. All numbers are in centimeters (cm). n/a = data not recorded and not available

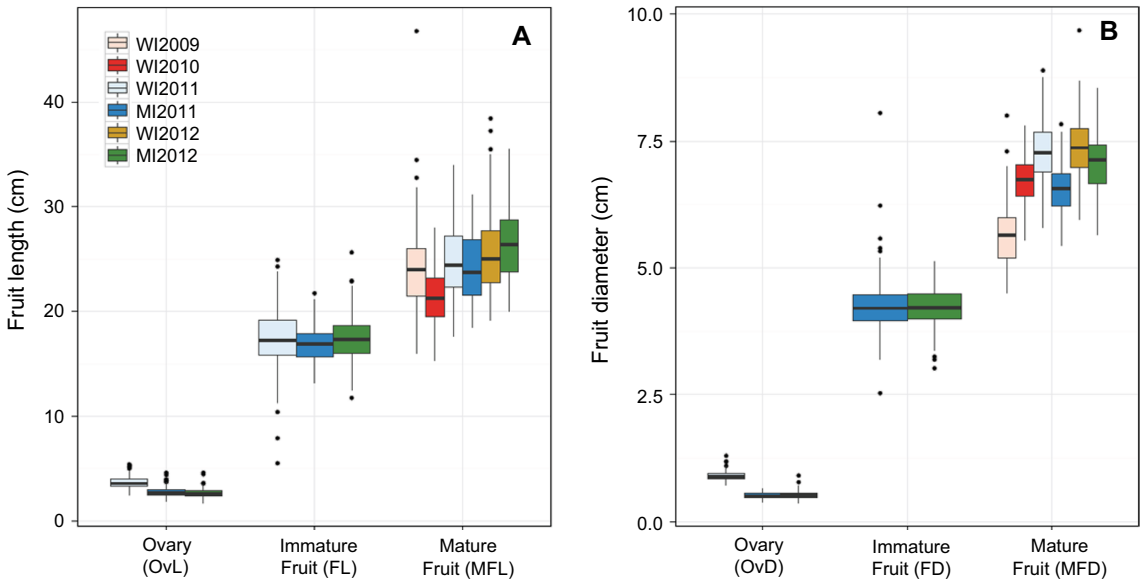

Fig. 2 *Box plots* of fruit length (**a**) and diameter (**b**) based on population means at three development stages (ovary, immature fruit and mature fruit) in six environments (WI2009, WI2010, WI2011, WI2012, MI 2011, MI2012)

the phenotypic extremes and F_1 value as the intermediate (Fig. S1; Table 2). The distribution of ovule number (OvN) was very similar to that of fruit length (Fig. S1K).

There were clear differences in phenotypic variation and distribution of fruit diameter in the segregating populations at three developmental stages. At anthesis, there was practically no difference in OvD among Gy14, 9930 and their F_1 (Table 2), but as fruits developed, Gy14 had increasingly larger fruit diameter than 9930; the F_1 fruit were intermediate. The distribution of MFD in F_3 families was shifted more toward larger values as compared with that of F_2 (Fig. S1B) probably because fruits were usually harvested earlier in the greenhouse experiment (WI2009) than in the field, which may not allow full expression of fruit diameter. Nevertheless, the distribution of OvD, FD and MFD was largely normal in F_2, F_3 and RILs (Fig. S1). This was also true for SCD and FTH that were associated with the fruit diameter (Figs. S1I, J).

Consistent with the no significant genotype × year and genotype × location effects, the grand means of fruit length and diameter at three development stages in two locations and 2 years were largely similar (Table 2). Generally, the broad sense heritability estimates (h^2) for all measured traits except FD were high, ranging from 64 to 92 % (phenotypic mean-based, Table S5). FD had a relatively low broad sense heritability ($h^2 = 0.25$), which was probably due to the fact that fewer replicated trials were conducted and the dates to measure FD data were difficult to control in different environments.

We analyzed the correlations among different traits in different populations and environments using MFL and MFD data. The Spearman's rank correlation coefficients (r_s) of MFL and MFD in F_2, F_3 and RILs are presented in supplemental Table S6. For MFL, there were significant and positive correlations between F_2 and F_3 ($r_s = 0.3949$, $P < 0.001$), F_2 and RIL ($r_s = 0.2669$, $P < 0.01$), as well as between F_3 and RIL ($r_s = 0.5472$, $P < 0.001$) suggesting common genetic mechanisms (and QTLs) playing roles in determining fruit length. For MFD, there was good correlation between F_2 and F_3, as well as between F_3 and RIL; however, no obvious correlation was found between F_2 and RIL (Table S6).

The Spearman's rank correlation coefficients of all traits from the RIL population, as well as MFL and MFD data from F_2 and F_3 are shown in Table 3. Strong positive correlations were observed among OvL, FL, MFL and OvN ($r_s = 0.45$–0.75, $P < 0.001$) suggesting that OvL is a good predictor for FL, MFL and OvN. On the other hand, no correlation was found among OvD, FD, MFD implying that ovary diameter is not a good predictor of fruit diameter at either immature or mature fruit stage. This is consistent with lack of significant differences among genotypes for OvD despite large differences at maturity (Fig. S1; Table 2). These results indicate that factors controlling length are largely determined pre-anthesis, while factors regulating diameter are largely determined post-anthesis. On the other hand, there were strong, positive correlations between OvL and OvD at anthesis ($r_s = 0.4584$, $P < 0.001$). There were also positive correlations between MFL and MFD in F_2 ($r_s = 0.3510$, $P < 0.01$), F_3 ($r_s = 0.3948$, $P < 0.001$), and RIL ($r_s = 0.3510$, $P < 0.001$) populations (Table 3). However, no significant correlation was found between FL and

Table 3 Spearman's rank correlation coefficients among different fruit size-related traits in the Gy14 × 9930 populations

	Population	OvL	FL	MFL	OvD	FD	MFD	SCD	FTH
FL	RIL	0.5471***							
MFL	RIL	0.5936***	0.7542***						
OvD	RIL	0.4584***	ns	ns					
FD	RIL	ns	ns	−0.2893***	ns				
MFD	RIL	ns	ns	0.3510***	ns	ns			
	F_2	n/a	n/a	0.4309***	n/a	n/a			
	F_3	n/a	n/a	0.3948***	n/a	n/a			
SCD	RIL	−0.3301***	−0.4540***	−0.3359***	ns	0.2705**	0.3265***		
FTH	RIL	ns	0.2943***	0.3934***	ns	ns	0.7282***	−0.7495***[a]	
OvN	RIL	0.4708***	0.5330***	0.6509***	ns	−0.2402**	ns	−0.3061***	ns

*** $P < 0.001$; ** $P < 0.01$; *ns* not significant ($n = 124$); *n/a* not applicable (no data)

[a] FTH and SCD data were based on their percentages of MFD (see text)

FD of immature fruits (12 dpp) suggesting that fruit diameter may be differentially regulated at different growth stages.

The SCD was positively correlated with both FD and MFD, but negatively correlated with OvL and FL. Meanwhile FTH was positively correlated with FL, MFL, and MFD ($r_s = 0.2943$–0.7282, $P < 0.001$), but there was no significant correlation between FTH and SCD, which was counterintuitive. The reason for this was that absolute FTH and SCD values were used in calculation of r_s and variations of MFD in the RIL population were not considered. If the SCD and FTH data were converted into percentages of MFD for each RIL for calculation of the correlation, the r_s between SCD and FTH would be −0.7495*** ($P < 0.001$), which was expected.

QTL analysis

For QTL analysis, MFL and MFD data from a single fruit of each F_2 plants (WI2009) and F_3 family means (WI2010), as well as grand means of nine traits for RILs across four environments were used. The BLUPs for six traits from the RIL population (OvL, FL, MFL, OvD, FD, and MFD) were also used in QTL analysis. Genotypic data for 82 F_2 and 92 F_3 families were based on 542 SSRs (Table S1); genotypic data for 140 RILs were based on 458 SNPs (Table S2).

A whole genome scan for QTLs was first conducted with single QTL model (SCANONE in R/qtl). Then three QTL models (MQM, CIM, and BIM) were used to refine QTL number and location. The chromosomal locations for QTLs detected in chromosomes 1, 4 and 6 were further refined by the addition of more markers in target regions for QTL analysis. Information for QTLs detected with all three QTL models and with both grand means and BLUPs are provided in supplemental Table S7. For comparison purpose, the QTLs detected with MQM for six traits in the RIL population based on the data from individual (2011 and 2012) year and location (WI and MI) were also provided in Table S7. No unique QTLs were detected with individual data sets; in fact, more QTLs were detected when grand means were used. Therefore, subsequent data analysis and discussions were based on results from grand means or BLUPs.

For each QTL, the peak location, support LOD score, R^2 value, additive and dominance (F_2 and F_3 only) effects were estimated. The 1.5-LOD support interval for each QTL was also provided, which is highlighted on the F_2-based SSR (Table S1) and RIL-based SNP (Table S2) genetic maps (defined by flanking markers of the interval). The summary of distinct QTLs for all traits is presented in Table 4. Each QTL was assigned a name; if multiple QTLs for the same trait detected by different QTL models or populations were located at the same or nearby locations, the same name was assigned.

QTLs detected with different mapping populations, QTL models and summary statistics

Since MFD and MFL data were collected in F_2, F_3 and RIL populations, we compared the power of QTL detection for the two traits in three populations (Table S7). With MQM, for MFL, 2 and 4 QTLs were detected in F_3 and RILs, respectively (none in F_2). All 4 MFL QTLs in the RIL population were consistently identified with three QTL models (MQM, BIM and CIM), of which, two were unique, and two (*qMFL1.1* and *qMFL6.1*) were shared with F_3 (Table S7). Interestingly, in the F_3 population, the two shared QTLs were major-effect, interacting QTLs ($R^2 = 35.2$ % for *qMFL1.1*; $R^2 = 19.0$ % for *qMFL6.1*; $R^2 = 6.1$ % for interaction), whereas in the RIL population, both were minor-effect QTLs ($R^2 < 10$ %) with no significant interactions (Table 4). On the other hand, the two major-effect,

interacting QTLs, *qMFL3.1* and *qMFL4.1* ($R^2 > 20$ % for each; $R^2 = 7.7$ % for interaction) detected in the RIL population were not detected in F_3 by MQM; but both loci and their interaction were picked by BIM or CIM (Table 4; Table S7) suggesting the different power in QTL detection by different QTL models.

Four distinct MFD QTLs were detected in the F_2 (*qMFD2.1*) and F_3 (*qMFD1.1*, *qMFD2.1*, *qMFD6.1*, and *qMFD7.1*) populations. Three QTLs (*qMFD2.1*, *qMFD4.1*, and *qMFD6.1*) were detected in the RIL population with all three models (Table 4). In the RIL population, the MQM failed to detect *qMFD1.1* and *qMFD7.1*, which were interacting loci in the F_3 (R^2 for interaction was 9.3 %). However, *qMFD7.1* was picked by both CIM (LOD = 2.4) and BIM in the RIL population; BIM also detected interaction between *qMFD6.1*, and *qMFD7.1* (Table S7) indicating BIM may have better power in detecting QTL interactions. Only one MFD QTL was detected in F_2, this was probably due to the small population size (n = 82) used for QTL analysis and limited number of fruits (only one) for data collection from each plant.

Using data from the RIL population, we were able to compare the power of QTL detection among three QTL models. Overall, for grand means, the MQM, CIM and BIM models detected 17 QTLs +1 interaction (for 7 traits, LOD >2.8), 21 QTLs (for 9 traits, LOD >3.0), and 26 QTLs + 4 interactions (for 9 traits, no LOD threshold), respectively (Table S7). All 17 QTLs detected with MQM were also detected by BIM and CIM, and in most cases, the magnitude of effects, peak locations, and phenotypic variations explained were very similar (Table S7). The interaction for MFL between Chr3 and Chr4 detected by MQM was also detected by BIM. Notably, seven QTLs in RILs and F_3 families were not detected by MQM, but were picked by both BIM and CIM suggesting these QTLs were possibly true. As such, these QTLs were included in Table 4. Interestingly, when the high-density map was used for QTL analysis, the two QTLs for OvN were detected again with MQM (see below). For detection of QTL interactions, the BIM model seems to be more powerful.

When BLUPs were used in QTL analysis in the RIL population for OvL, FL, MFL, OvD, FD, and MFD, the MQM, CIM and BIM models detected 14 QTLs +1 interaction, 11 QTLs (LOD > 3.0) + no interaction, and 20 QTLs + 5 interactions, respectively (Table S7). All 14 QTLs identified by MQM with BLUPs were also detected with grand means (Table 4). Four QTLs uniquely detected by BLUPs but not by grand means were *qOvD7.1*, *qFD2.1*, *qFD5.1*, and *qFD6.1*. However, the three FD QTLs could only be detected with either BIM or CIM.

To summarize, the RIL population detected more consistent QTLs. All three models seem to be able to detect major QTLs with largely consistent results in number and location of QTLs. MQM appeared more robust and provided more information, whereas BIM exhibited better power in detecting interacting QTLs.

Fruit size QTLs

As shown in Table 4, 29 distinct QTLs were detected for nine traits by three QTL models in three populations with grand means or BLUPs with one QTL for FTH, 2 for OvN, 3 each for OvL, FL and FD, 4 each for MFL, OvD, and SCD, and 5 for MFD. Of them, *qMFD1.1* was only detected in F_3; all other MFD or MFL QTLs detected in F_3 were shared with RILs.

QTLs for ovary length (OvL) and diameter (OvD) Three QTLs, *qOvL1.1* ($R^2 = 17.4$ %), *qOvL3.1* ($R^2 = 9.2$ %) and *qOvL4.1* ($R^2 = 18.7$ %) were detected for OvL (Table 4). The ovary length of Gy14 was shorter than that of 9930 (Table 2), but the additive effects of the Gy14 alleles at all three loci were positive (*a* allele contributing to increase of ovary length). Since the three QTLs together could only explain 28.8–38.0 % phenotypic variations, additional QTLs or QTL interactions not detected in the RIL population may exist.

The four QTLs, *qOvD2.1*, *qOvD5.1*, *qOvD6.1* and *qOvD7.1* together explained up to 35.7 % of phenotypic variations of OvD in the RIL population. Interestingly, no significant differences of ovary diameter were observed among Gy14, 9930 and their F_1, but transgressive segregation of OvD was obvious in the RIL population (Table 2, Fig. S1F). This could be explained by the opposite additive effects among the four QTLs; but additional QTLs not detected in this study for OvD are possible.

QTLs for immature fruit length (FL) and diameter (FD) No QTL for FD was identified in any populations using grand mean; 3 FD QTLs were detected with BLUPs by either BIM or CIM, but only 18.2 % total phenotypic variations could be explained. The failure to detected major-effect QTLs for FD may be due to fewer environments for data collection or the possible inconsistency in use of criteria (dates after anthesis) to measure FD across years and locations. It is also possible that there were no significant genetic variations of FD in the RIL population at the time of data collection (10-12 dpp) because the increase of fruit diameter in Gy14 occurred mainly in later development stage (compare OvD, FD and MFD between Gy14 and 9930 in Table 2).

For FL, three QTLs, *qFL3.1*, *qFL4.1* and *qFL6.1* were able to explain a maximum 36.8 % of phenotypic

Table 4 Summary of fruit size QTLs detected with BIM (B), CIM (C) and MQM (M) in F_2, $F_{2:3}$ and RIL populations across six environments in four years with grand mean (GM) and BLUPs. Values were from MQM analysis or CIM in cases MQM data were not available

Target traits[a]	Mapping pop	QTL Models	Data used	Chr (LG)	QTL Loci	Peak Pos (cM)	Nearest marker	LOD score[b]	1.5 LOD interval				R^2 (%)[c]	Additive effect (*a*)	Dominance effect (*d*)	QTL interactions	Total var explained (%)
									Left locus	Left cM	Right locus	Right cM					
OvL	RIL	B, C, M	GM, BLUP	1	*qOvL1.1*	33.4	SNP.7985	6.2	SNP.7425	31.3	SNP.8593	35.2	17.4	0.10		RIL, *qOvl1.1* with *qOvl4.1* ($R^2 = 7.0$ %)	28.8–38.0
	RIL	B, C, M	GM, BLUP	3	*qOvL3.1*	46.8	SNP.142445	3.0	SNP.37745	44.3	SNP.117741	54.3	9.2	0.13			
	RIL	B, C, M	GM, BLUP	4	*qOvL4.1*	66.8	SNP.59961	5.8	SNP.58453	58.3	SNP.63725	71.6	18.7	0.19			
FL	RIL	B, C	GM, BLUP	3	*qFL3.2*	67.2	SNP.44901	3.4	SNP.43341	63.0	SNP.45957	70.4	8.6	0.39		None	32.0–36.8
	RIL	B, C, M	GM, BLUP	4	*qFL4.1*	68.4	SNP.59961	6.0	SNP.58453	58.3	SNP.150557	75.6	17.0	0.92			
	RIL	B, C, M	GM, BLUP	6	*qFL6.1*	69.0	SNP.93865	3.4	SNP.144625	43.5	SNP.92777	77.3	9.2	−0.63			
MFL	F_3	B, C, M	F3 mean	1	*qMFL1.1*	22.7	SSR21336	10.0	UW019729	22.0	SSR16472	27.8	35.2	2.13	−1.2	F_3, *qMFL1.1* with *qMFL6.1* ($R^2 = 6.1$ %)	49.2
	F_3	B	F3 mean	3	*qMFL3.1*	45.7		3.2					7.8	1.11	−0.4		
	F_3	B, C	F3 mean	4	*qMFL4.1*	62.5	UW084421	2.4	SSR18719	0.00	UW042029	108.40	7.1	0.85	0.7		
	F_3	B, C, M	F3 mean	6	*qMFL6.1*	37.6	SSR03932	6.0	SSR17604	26.5	UW085138	53.1	19.0	−1.33	−0.1		
	RIL	B, C, M	GM, BLUP	1	*qMFL1.1*	33.4	SNP.7985	4.9	SNP.112117	15.4	SNP.122737	43.9	8.2	1.08		RIL, *qMFL3.1* with *qMFL4.1* ($R^2 = 7.7$ %)	48.2–58.3
	RIL	B, C, M	GM, BLUP	3	*qMFL3.1*	10.8	SNP.31413	12.0	SNP.125329	8.6	SNP.162797	13.8	22.9	1.48			
	RIL	B, C, M	GM, BLUP	4	*qMFL4.1*	69.9	SNP.63725	11.4	SNP.136581	64.2	SNP.127269	78.4	21.6	1.29			
	RIL	B, C, M	GM, BLUP	6	*qMFL6.1*	50.4	SNP.90045	8.5	SNP.144625	43.5	SNP.91177	54.3	8.5	−1.39			
OvD	RIL	B, C, M	GM, BLUP	2	*qOvD2.1*	13.6	SNP.159433	4.5	SNP.19601	11.4	SNP.113369	17.6	12.8	−0.01		RIL, *qOvD2.1* with *qOvD6.1* ($R^2 = 12.2$ %)	23.5–35.7
	RIL	B, C	GM	5	*qOvD5.1*	79.5	SNP.77021	2.4	SNP.72793	68.3	SNP.77021	80.9	6.7	0.02			
	RIL	B, C, M	GM, BLUP	6	*qOvD6.1*	52.0	SNP.141049	4.6	SNP.88809	45.1	SNP.93613	73.5	12.9	0.01			
	RIL	B, C, M	BLUP	7	*qOvD7.1*	23.0	SNP.163837	3.7	SNP.128737	16.2	SNP.105149	68.3	10.3	−0.01			
FD	RIL	C	BLUP	2	*qFD2.1*	32.5	SNP.126917	2.9	SNP.107169	7.0	SNP.23701	45.5	4.4	0.02		RIL, *qFD5.1* with *qF6.1* ($R^2 = 14.2$ %)	0–18.2
	RIL	B	BLUP	5	*qFD5.1*	46.4	SNP.70369	4.8					17.5	−0.03			
	RIL	B	BLUP	6	*qFD6.1*	48.0	SNP.90457	4.1					14.6	0.01			
MFD	F_2	B, C, M	F2 plant	2	*qMFD2.1*	74.0	UW085339	3.8	SSR00507	59.1	SSR16462	88.1	16.5	−0.25	−0.4	None	16.5
	F_3	B, C, M	F3 mean	1	*qMFD1.1*	69.3	UW008891	7.4	UW375302	64.0	SSR14445	77.0	25.7	−0.23	0.0	F_3, *qMFD1.1* with *qMFD7.1* ($R^2 = 9.3$ %)	46.5
	F_3	B, C	F3 mean	2	*qMFD2.1*	68.3	UW083965	4.3	SSR00507	58.70	UW083904	78.38	10.3	−0.16	−0.2		
	F_3	B, C, M	F3 mean	6	*qMFD6.1*	48.9	UW083827	3.3	SSR03932	37.6	UW085171	63.0	10.2	−0.19	0.2		
	F_3	B, C, M	F3 mean	7	*qMFD7.1*	37.0	UW085407	6.2	MU6481-1	20.1	SSR14797	51.4	20.9	−0.15	0.3		
	RIL	B, C, M	GM	2	*qMFD2.1*	74.0	SNP.28169	3.4	SNP.142889	69.1	SNP.27309	79.9	9.8	−0.16		RIL, *qMFD6.1* with *qMFD7.1* ($R^2 = 7.5$ %)	24.0–25.6
	RIL	B, C, M	GM, BLUP	4	*qMFD4.1*	64.2	SNP.136581	4.9	SNP.58453	58.3	SNP.63725	71.5	9.3	0.12			
	RIL	B, C, M	GM, BLUP	6	*qMFD6.1*	46.6	SNP.150873	5.4	SNP.88809	45.1	SNP.91177	54.3	16.5	−0.21			
	RIL	B, C	GM, BLUP	7	*qMFD7.1*	46.4	SNP.103657	2.4	SNP.100505	20.8	SNP.106789	52.7	6.3	−0.13			
OvN	RIL	B, C	GM	1	*qOvN1.1*	43.6	SNP.165749	3.9	SNP.8593	35.2	SNP.138053	46.7	12.3	2.04		None	29.4
	RIL	B, C	GM	6	*qOvN6.1*	60.3	SNP.91625	2.9	SNP.141049	52.9	SNP.137389	68.1	12.4	−2.05			

Table 4 continued

Target traits[a]	Mapping pop	QTL Models	Data used	Chr (LG)	QTL Loci	Peak Pos (cM)	Nearest marker	LOD score[b]	1.5 LOD interval				R^2 (%)[c]	Additive effect (*a*)	Dominance effect (*d*)	QTL interactions	Total var explained (%)
									Left locus	Left cM	Right locus	Right cM					
SCD	RIL	B, C, M	GM	2	*qSCD2.1*	73.0	SNP.28301	4.4	SNP.28601	70.3	SNP27309	79.9	9.9	−0.11		None	49.4
	RIL	B, C, M	GM	3	*qSCD3.1*	90.0	SNP.49125	7.3	SNP.22341	88.0	SNP.51973	94.2	17.2	−0.16			
	RIL	B, C, M	GM	6	*qSCD6.1*	63.8	SNP.135301	5.5	SNP.92461	54.9	SNP.95017	65.5	12.6	0.13			
	RIL	B, C, M	GM	7	*qSCD7.1*	46.4	SNP.103657	5.0	SNP.102221	39.4	SNP.147781	50.5	11.3	−0.13			
FTH	RIL	B, C, M	GM	6	*qFTH6.1*	58.2	SNP.168061	5.1	SNP.89429	45.3	SNP.94117	70.3	18.7	−0.12		None	18.7

[a] *OvL* ovary length, *OvD* ovary diameter, *FL* immature fruit length, *FD* immature fruit diameter, *MFL* mature fruit length, *MFD* mature fruit diameter, *SCD* seed cavity size and *FTH* flesh thickness mature fruit

[b] For RILs, LOD threshold at $P = 0.05$ with MQM for OvL, OvD, FL, FD, MFL, MFD, OvN, SCF and FTH was 2.7, 2.8, 2.7, 2.6, 2.7, 2.9, 2.8, 2.7 and 2.8, respectively, based on 1,000 permutations. LOD thresholds for MFL (F_2), MFD (F_2), MFL (F_3) and MFD (F_3) were 3.1, 3.5, 3.6, and 3.6, respectively. QTLs with lower thresholds were included when detected with more than one method or population

[c] Percentages of phenotypic variations explained by the detected QTLs

[d] Phenotypic variations explained by all QTLs in that population

variations. The QTL *qFL6.1* contributed negatively to fruit length, whereas other two showed positive additive effects.

QTLs for length (MFL), diameter (MFD) and associated traits (OvN, SCD and FTH) Four QTLs in four chromosomes were detected for MFL in both F_3 and RIL populations (Table 4) accounting for 49.2 % (F_3) and 58.3 % (RIL) phenotypic variations observed. The two QTLs detected in the F_3 population, *qMFL1.1* ($R^2 = 27.8$ %) and *qMFL6.1* ($R^2 = 19.0$ %) were major-effect, and interacting QTLs that accounted for ~6.1 % phenotypic variations, whereas other two, *qMFL3.1* and *qMFL4.1* were minor-effect QTLs. The interaction plot at the two loci is illustrated in Fig. 3a. In F_3, all but *qMFL6.1* showed positive additive effects on mature fruit length. Interestingly, in the RIL, the same set of four QTLs showed the opposite in the magnitudes of effects: *qMFL1.1* and *qMFL6.1* were major-effect QTLs ($R^2 > 20$ % for each); *qMFL3.1* and *qMFL4.1* were minor and interacting QTLs ($R^2 = 7.7$ % for interactions, Table 4).

There were five QTLs for mature fruit diameter (MFD) with two major-effect, interacting QTLs, *qMFD1.1* ($R^2 = 25.7$ %) and *qMFD7.1* ($R^2 = 20.9$ %) that was only detected with MQM in the F_3 population ($R^2 = 9.3$ % from the interaction that is illustrated in Fig. 3b). The remaining three QTLs, *qMFD2.1 (*detected in F_2 and RIL), *qMFD4.1* (detected only in RIL) and *qMFD6.1* (detected in F_3 and RIL) each could explain ~10 % phenotypic variations. The four QTLs detected in F_3 all showed negative additive effects whereas in the RIL *qMFD4.1* had large positive additive effect on MFD (Table 4). This can probably explain why the four QTLs in RIL could only explain ~25 % phenotypic variations versus the 46.5 % in the F_3.

Four QTLs for SCD in four chromosomes (total $R^2 = 49.4$ %) were consistently identified with three QTL models. The Gy14 alleles at three loci (*qSCD2.1*, *qSCD3.1* and *qSCD7.1*) contributed to smaller seed cavity sizes (negative additive effects). The MQM model failed to detect any significant QTL for OvN in the RIL, but two QTLs, *qOvN1.1* and *qOvN6.1*, were consistently detected with BIM and CIM models (Table S7) contributing to 29.4 % phenotypic variations. Finally, only one QTL, *qFTH6.1* was detected for flesh thickness ($R^2 = 18.7$ %).

Refinement of QTL locations in chromosomes 1, 4 and 6

From Table 4, chromosomes 1, 4 and 6 seemed to harbor most major-effect and consistent QTLs for fruit size-related traits. To refine the map locations of QTLs, additional markers from the ultra-high-density SNP-based RIL map (Rubinstein et al. 2015) were selected to increase marker density on the framework RIL genetic map in target regions

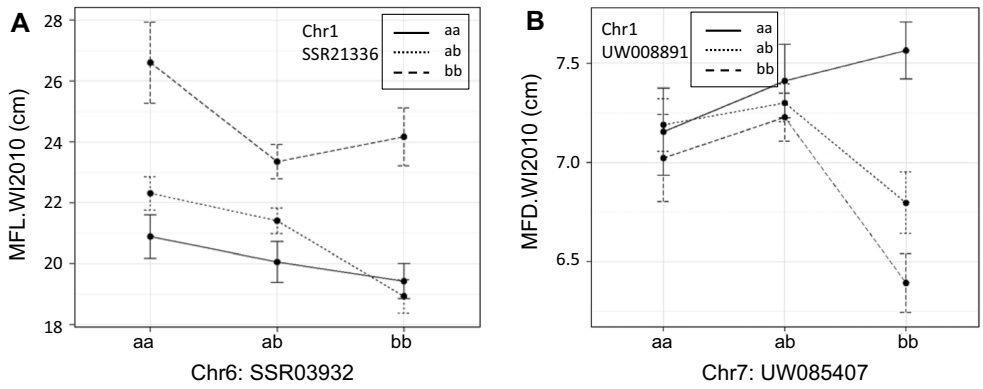

Fig. 3 Interaction plots of QTL pairs detected in F_3 population for mature fruit length (MFL, **a**) and diameter (MFD, **b**). Gy14 and 9930 carries *a* and *b* allele, respectively. In **a**, *qMFL1.1* at SSR21336 locus and *qMFL6.1* at SSR03932 increases (positive additive effect) and decreases (negative additive effect) MFL, respectively. In **b**, QTL *qMFD1.1* at SSR08891 and *qMFD7.1* at UW085407 both reduce MFD (negative additive effects). There is epistasis between each pair of QTLs

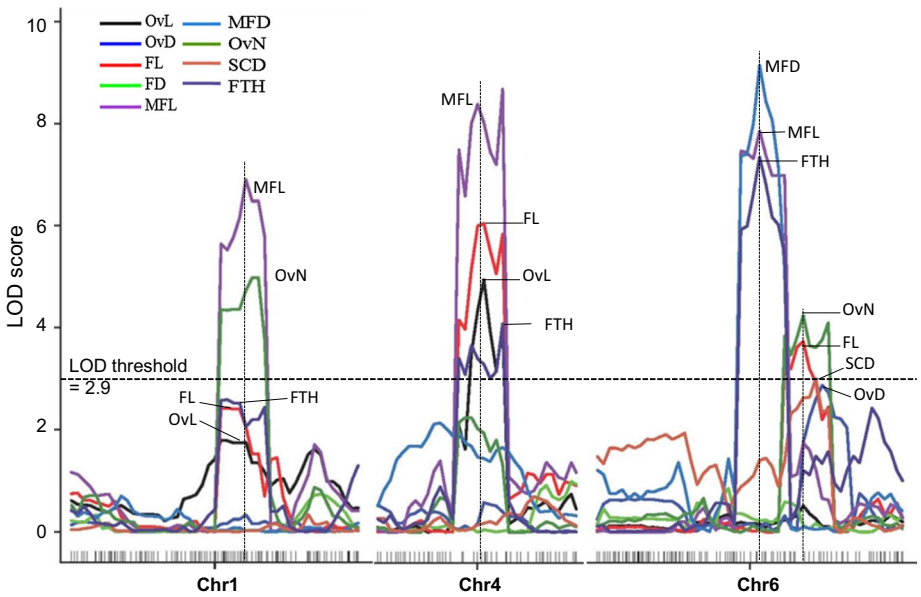

Fig. 4 LOD profiles of fruit size-related QTLs detected with MQM model in the RIL population and high-density SNP maps in cucumber chromosomes 1 (*left*), 4 (*middle*), and 6 (*right*). The *dashed horizontal line* is LOD threshold for all QTLs (LOD = 2.9). *OvL* ovary length, *OvD* ovary diameter, *FL* immature fruit length, *FD* immature fruit diameter, *MFL* mature fruit length, *MFD* mature fruit diameter. *FTH* flesh thickness, *SCD* seed cavity size and *OvN* ovule number in mature fruit

of the three chromosomes. Three hundred and twenty-one SNP markers were added, and the resulting high-density map for the three chromosomes is presented in supplemental Table S8. QTL analysis results with the MQM model are shown in Table S9. LOD curves of detected QTLs in the three chromosomes are illustrated in Fig. 4. Compared with the results obtained from the frame work map (Table S2; Table 4), the two QTLs for OvN (*OvN1.1* and *OvN6.1*) were above the LOD threshold with the high-density map. The results from the low-density map are largely the same in terms of the numbers and magnitudes of effect of detected QTLs; however, physically, the 1.5-LOD intervals for these QTLs were significantly narrowed (cf. Table S2 and Table S8). One interesting observation in Fig. 4 was that LOD curve peaks for some traits (for example, FL, OvL and FTH in Chr1), despite below the LOD threshold, were obvious suggesting these QTLs were probably also true if examined in adequate environments.

Establishment of a consensus set of fruit size QTLs during cucumber fruit development

For each QTL, there was a 1.5-LOD interval defined by flanking markers. Since the location of most markers in the Gy14 and 9930 draft genome assemblies are known (Table 4; Tables S7 and S9), it was possible to determine the physical locations of mapped QTLs in the cucumber genome, which are presented in Table S2. The chromosomal locations of all QTLs are summarized in Table 5. In chromosomes 1, 2, and 6, detected QTLs were putatively

located in two blocks, and three blocks could be recognized for chromosome 3 (see Table S2 for details). QTLs detected by single QTL model with either grand mean or BLUPs were also included in Table 5. Despite weak LOD support, these QTL were probably true since they were co-localized with other well supported QTLs. As such, 39 QTLs for 9 traits could be recognized with 7, 4, 7, 5, 5, 8, and 3 QTLs on Chr1 to Chr7, respectively (Table 5),

When QTLs for different traits were mapped in the same chromosomal block, they were treated as the same QTL. As such, from the synopsis of QTL information presented in Table 5, 12 consensus QTLs for cucumber fruit size (FS) could be inferred, which were *FS1.1*, *FS1.2*, *FS2.1*, *FS2.2*, *FS3.1*, *FS3.2*, FS3.3, *FS4.1*, *FS5.1*, *FS6.1*, *FS6.2* and *FS7.1* (last row of Table 5). Based on their roles in fruit development, each QTL could act alone or in combination with other QTLs for fruit elongation, radial growth (increase of diameter), or both at one or multiple stages of fruit development. Based on the growth stages detected, the putative roles of these 12 QTLs in fruit development are illustrated in Fig. 5, in which only QTLs for fruit elongation and widening were considered. It is clear that both *FS4.1* and *FS5.1* played roles in fruit elongation, and *FS5.1* was important for the increase of fruit diameter throughout the whole fruit development. Five QTLs (boldface typed), *FS1.1*, *FS4.1*, *FS5.1*, *FS6.1*, and *FS6.2* acted in both processes. Taken together, as shown in Table 5, except for *FS1.2*, each of the remaining 11 consensus QTLs seems to affect multiple traits examined in the present study.

The co-localization of QTLs, which may have the same genetic basis, can explain correlations among different traits. For example, there were significant and strong correlations among OvL, FL and MFL (Table 3) because they shared QTLs *FS1.1* and *FS4.1*. For the same reason, there were significant correlations between OvL and OvD ($r_s = 0.4584$, $P < 0.001$; sharing *FS1.1* and *FS5.1*), MFL and MFD ($r_s = 0.3510$, $P < 0.001$; sharing *FS4.1* and *FS6.1*). No significant correlation between FL and FD was found; this is consistent with the fact that no common QTLs were identified for FL and FD in the present study (Fig. 5).

Discussion

Power of QTL detection in different mapping populations using different QTL models

We conducted QTL mapping of fruit size-related traits in F_2, F_3 and RIL populations derived from the same cross (Gy14 × 9930) with three QTL models (CIM, BIM and MQM). This gave a good opportunity to assess the power of QTL mapping in different populations. One, 6, and 8

Table 5 Summary of chromosomal locations of fruit size QTLs detected from present study and consensus fruit size QTLs. Chromosomes 1, 2, 3 and 6 were each divided into two or three blocks each. QTLs placed in the same chromosomal block are putatively co-localized*

Descriptors	Traits	Chr1-B1	Chr1-B2	Chr2-B1	Ch2-B2	Chr3-B1	Chr3-B2	Chr3-B3	Chr4	Chr5	Chr6-B1	Chr6-B2	Chr7
Longitudinal growth (fruit elongation)	OvL	***qOvL1.1*****					***qOvL3.1***		***qOvL4.1*****	*qOvL5.1*			
	FL	*qFL1.1*				*qFL3.1*	***qFL3.2***	*qFL3.3*	***qFL4.1***			*−qFL6.1*	
	MFL	*qMFL1.1***				***qMFL3.1*****			***qMFL4.1*****		***−qMFL6.1*****		
	OvN	***qOvN1.1***	–						*qOvN4.1*			***−qOvN6.1***	
Radial growth (increase diameter)	OvD	*qOvD1.1*		***−qOvD2.1*****						*qOvD5.1*		***qOvD6.1*****	***−qOvD7.1***
	FD			*qFD2.1*						***−qFD5.1*****	***qFD6.1*****		
	MFD		*−qMFD1.1***		*−qMFD2.1*				*qMFD4.1*	*−qMFD5.1*	***−qMFD6.1*****		*−qMFD7.1***
	SCD				*−qSCD2.1*			***−qSCD3.1***				***qSCD6.1***	***−qSCD7.1***
	FTH	*−qFTH1.1*						*qFTH3.1*		*−qFTH5.1*		***−qFTH6.1***	
Consensus QTL		*FS1.1*	*FS1.2*	*FS2.1*	*FS2.2*	*FS3.1*	*FS3.2*	*FS3.3*	*FS4.1*	*FS5.1*	*FS6.1*	*FS6.2*	*FS7.1*

* Based on integrated information from QTL analysis with three QTL models (BIM, CIM, MQM) in F_2, F_3 and RILs with both grand means and BLUPs. Boldface typed are major-effect ($R^2 > 10$ %) QTLs; underlined QTLs are supported with multiple lines of evidence; double asterisks indicate interacting QTLs for the same trait. The minus sign preceding a QTL indicates negative additive effect

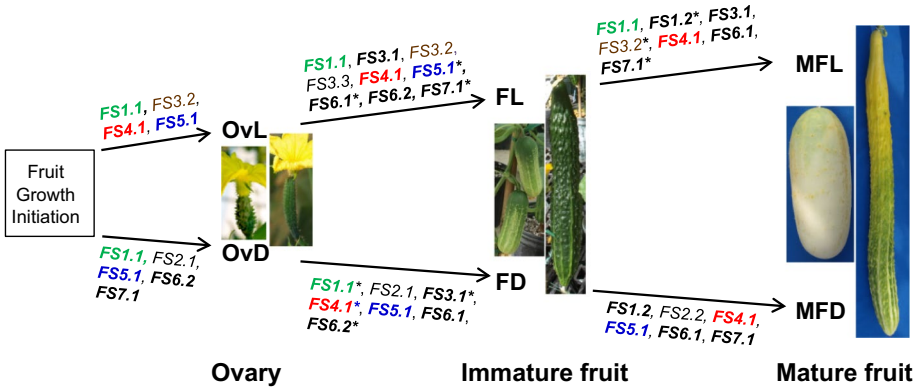

Fig. 5 Dynamic roles of 12 consensus fruit shape (FS) QTLs in control of fruit elongation and increase of diameter during different stages of cucumber fruit development. QTLs with colors act throughout whole development stage; *boldface*-typed QTLs contribute to both longitudinal and radial growth. QTLs also identified in prior studies are marked with an *asterisk*

QTLs were detected for MFL and/or MFD in F_2, F_3, and RILs, respectively (Table 4). The RIL population appears to be able to detect more QTLs than F_3, and F_2 was the least powerful for QTL detection. Similar results were obtained in some previous studies in cucumber. For example, Yuan et al. (2007, 2008) conducted QTL mapping for horticultural traits including FL, FD, SCD, FTH using F_2, F_3 and RIL populations from the cross between S94 and S06; 4, 15 and 28 QTLs were detected in F_2, F_3, and RIL populations, respectively, for the four traits. In the F_2 and F_3 populations derived from PI 183967 × 981, 2 and 3 QTLs were detected for FL (one shared) (Cheng et al. 2010), but 4 were picked up in the RIL population (one shared with F_3) (Wang et al. 2014). On the other hand, many important interactions and the dominance effects could only be detected in F_2 or F_3 populations which sometimes may explain a significant portion of phenotypic variations (for example, Table 4). It is also common that the effects of QTLs detected in F_2/F_3 and RILs could be very different in both magnitude and direction (Table 4; Cheng et al. 2010; Wang et al. 2014). Therefore, in a self-pollinated crop, while F_2 is in general not an ideal population for QTL mapping, if F_3 families are available, QTL mapping power could be increased by phenotyping more plants per family with adequate number of test environments and replications (e.g., He et al. 2013).

Using grand means of the RIL population, we compared and contrasted the power of QTL detection with three QTL models. In total 17, 21 and 26 QTLs were detected by MQM, CIM and BIM, respectively; 1 and 4 interactions were also identified with MQM and BIM (none by CIM) (Table S7). All 17 QTLs detected by MQM were also detected by BIM and CIM, and the magnitude and direction of QTL effects were largely comparable. While BIM seems to be able to detect more QTLs than MQM or CIM, since BIM does not provide a LOD threshold or LOD support interval for each QTL, it is difficult to assess the confidence of each QTL. CIM has some limitations in estimating the joint contribution to the genetic variance of multiple linked QTLs (Zeng et al. 1999). MQM model seems to be a good choice for QTL analysis. However, analysis with all three models may be beneficial to uncover additional true QTLs which otherwise would not be identified by MQM alone (for example, Table 4).

To summarize, in QTL mapping, it is helpful to run different QTL models to obtain the numbers, locations, the magnitude and direction of effects of underlying QTLs for the target traits. While it is preferable to conduct QTL mapping with RILs, an F_3 population could achieve similar power if phenotyping is conducted in replicated trials in multiple environments with added benefits to identify QTL interactions and dominance effects.

Correlation of fruit length and diameter in cucumber

Each cucumber market class has its unique requirements for fruit length and diameter. For North American pickling cucumber, the length-by-diameter (*L/D*) ratio at harvest should be around 2.8–3.1. Several studies have conducted QTL mapping for *L/D* (e.g., Kennard and Havey 1995; Yuan et al. 2007, 2008; Miao et al. 2011) aiming for marker-assisted selection for *L/D* in cucumber breeding. However, previous studies have revealed no or weak correlation between fruit length and diameter in cucumber at different development stages, or the correlation sometimes was dependent on the environments (Yuan et al. 2007, 2008; Miao et al. 2011; Bo et al. 2015). In the present study, we found significant correlation between OvL and OvD, between MFL and MFD, but not between FL and FD (Table 3), which could be explained by common QTLs at the ovary and mature fruit stages but not at the immature fruit stage (Fig. 5). This may suggest elongation of fruit (FL) and increase of diameter (FD) at the exponential growth stage might be under different genetic mechanisms. These observations may also suggest that the correlation between fruit length and diameter may vary in different development stages of fruit development and different genetic backgrounds. Therefore, as suggested in Bo et al. (2015), since *L/D* is a composite trait (calculated from

FL and FD) with unknown genetic basis, caution should be taken in using L/D as a selection criterion in marker-assisted selection. This is especially true for L/D at the immature fruit stage.

QTLs for fruit size in cucumber: a dynamic view

By integrating information of all of the QTLs for the 9 fruit size-related traits detected in the present study (Table S7), 12 consensus QTLs underlying cucumber fruit development were inferred (Table 5); each of the 12 QTLs may play roles in one or multiple stages for fruit elongation or increase of diameter or both (Fig. 5). A number of fruit length and diameter QTLs have been identified in several earlier studies (e.g., Kennard and Havey 1995; Dijkhuizen and Staub 2002; Yuan et al. 2008; Cheng et al. 2010; Bo et al. 2015; Wang et al. 2014; Wei et al. 2014). The approximate locations of these QTLs were placed onto the SNP-based map developed in the present study (shaded blocks in Table S2). We assumed that QTLs at the same or close physical locations across different studies belonged to the same QTL locus, thus may share common genetic mechanisms underlying the fruit growth.

QTLs for fruit size from various studies are distributed in all seven chromosomes with QTLs in chromosome 2 being detected only in the present study (Table S2). In some cases, the co-localized QTLs for the same trait from different studies were highly consistent in the magnitude and direction of QTL effects. For example, in the Gy14 × 9930 RILs, two major-effect QTLs, *qMFL3.1* (*FS3.1*) and *qMFL4.1* (*FS4.*1) (R^2 > 20 % each) were detected, which seem to correspond well, respectively, to *mfl3.1* (R^2 = 12.3 %), and *mfl4.1* (R^2 = 10.3 %) identified in Bo et al. (2015) (Table S2). In other cases, QTLs mapped in the same or nearby location for the same trait were different in their magnitudes of effects, or roles in different development stages (Table S2). Four early studies (Yuan et al. 2008; Cheng et al. 2010; Bo et al. 2015; Wei et al. 2014) identified major-effect QTLs for FL, FD, MFL or MFD in the *FS1.1* or *FS1.2* region in chromosome 1 where only minor-effect QTLs were detected in the present study. At *FS3.2* locus, we only detected a minor-effect QTL (*qFL3.2*, R^2 = 2.6–8.2 %, Table S7), but Wei et al. (2014) detected major-effect QTLs for both FL and MFL with R^2 as high as 45 %. Some of the 12 consensus QTL identified herein seem to play different roles in other studies. For example, in our study, *FS7.1* was found to only affect fruit radial growth (Table 5), but Bo et al. (2015) identified a major-effect QTL, *mfl7.1* (Table S2) at this location that affects fruit longitudinal growth. It is possible that *FS7.1* may play roles in both longitudinal and radial growth during cucumber fruit development. Integrating QTL mapping results from the earlier and present may give us a more complete picture on the roles of the 12 consensus QTLs established herein. As such, additional roles for 9 of the 12 consensus QTLs could be recognized, which are shown in Fig. 5 (QTLs with asterisks). Thus, among the 12 consensus QTLs, *FS2.1* and *FS2.2* are involved in radial growth; *FS3.2* and *FS3.3* act in fruit elongation, and the remaining 8 QTLs play roles in both processes. For their roles in different growth stages, *FS3.3* and *FS6.2* seem to be inactive in mature fruit growth whereas *FS1.2* is the opposite that is active only during this stage (Fig. 5).

If the 12 consensus QTLs are underlying fruit growth and development in cucumber, how could the discrepancies in the number, location and magnitude of effect of QTLs for the same trait in different studies (Table S2) be explained? It is easy to see that different mapping populations (F_2, F_3 and RIL), growth stages at time of data collection (e.g., commercial harvest stage vs. mature fruits), the criteria of trait phenotyping, the seasons (spring, summer or fall) of the trials, the experimental design, and the growth environment (open field, greenhouse or protected plastic houses, number of replications) may all contribute to the varying results in these QTL mapping studies. But the most reasonable explanation is the cucumber lines used in these studies that belong to different taxonomic groups or market classes. These studies used QTL mapping populations developed from crosses between cucumber lines of different market classes (Yuan et al. 2007, 2008; Miao 2011; Wei et al. 2014; this study), between cultivated (North China type) and wild (*C. s.* var. *hardwickii* PI183967) cucumbers (Kennard and Havey 1995; Dijkhuizen and Staub 2002; Cheng et al. 2010; Wang et al. 2014), or between cultivated (North China type) and semi-wild (*C. s.* var. *xishuangbannesis*) cucumbers (Bo et al. 2015). It is possible that these traits have undergone domestication or diversifying selection for specialized market classes. The underlying genes or QTLs may be different targets of the selection. Therefore, the same gene/QTL responsible for the same trait may be differentially expressed in different market classes of cucumbers, consequently rendering its effect detectable in some lines but undetectable in other genetic backgrounds.

Cultivated cucumber was domesticated from its wild progenitor *C. sativus* var. *hardwickii* (Yang et al. 2012). Using populations derived from the cross between this wild progenitor (PI 183967) and a cultivated cucumber (line 981), Cheng et al. (2010) and Wang et al. (2014) identified several QTLs for FL and MFL which may potentially be under selection during domestication (Qi et al. 2013). The QTL *FL4.1* in Cheng et al. (2010) and Wang et al. (2014) has largely consistent map locations with *FS4.1* from the present study (Table S2), which seems to be a candidate locus that has undergone selection during domestication. Clearly, additional work is needed to improve the map resolution for those QTLs to make any firm conclusion.

Candidate genes for cucumber fruit size/shape QTLs

The genetic mechanisms of fruit development in tomato (*Solanum lycopersicum*) have been extensively studied. Several genes for fruit size or shape in tomato have been cloned (reviewed by Rodriguez et al. 2011; van der Knaap et al. 2014; Azzi et al. 2015). It seems that four genes are responsible for the majority of fruit shape diversity found in tomato, which include *SUN* and *OVATE* regulating fruit elongation; and *LOCULE NUMBER* (*LC*) and *FASCIATED* (*FAS*) for the regulation of locule number and flat shape. *SUN* encodes a protein that is a member of the IQ domain family; *OVATE* encodes a protein in the ovate family protein (OFP); *FAS* encodes a protein that is a member of the YABBY family, whereas *LC* is probably encoded by the ortholog of the *Arabidopsis thaliana* gene *WUSCHEL*, which is a member of the WOX family. In addition, two fruit weight (FW) genes, *CNR/FW2.2* and *SlKLUH/FW3.2*, are also related with tomato fruit size. *CNR/FW2.2* encodes a member of the cell number regulator (CNR); *SlKLUH/FW3.2* encodes a member of a subfamily of cytochrome P450 A78 class (*CYP78A*) and the ortholog of KLUH. More recently, a tomato mutant *Slelf1* was identified which exhibits an elongated fruit shape caused by increased cell layers (Chusreeaeom et al. 2014).

With the assumption that similar mechanisms may exist in melon (*Cucumis melo* L.) for fruit shape and size control, Monforte et al. (2014) identified 74 homologs of the CNR, CYP78A, OFP, SUN, WOX, and YABBY gene families in the melon genome and associate fruit weight or shape QTLs with some of these homologs. All the 74 melon homologs are also present in the cucumber genome (Bo et al. 2015). Yang et al. (2013a) compared the expression pattern of seven kinesin genes (*CsKF1* to *CsKF7*) in the cucumber genome among cucumber lines with varying fruit sizes, and found that *CsKF2–CsKF6* were positively correlated with rapid cell production; whereas, *CsKF1* and *CsKF7* showed a strongly positive correlation with rapid cell expansion. The approximate locations of the 74 fruit size/shape homologs as well as the seven kinesin genes on the cucumber SNP map are shown in Table S2. Many tomato fruit shape gene homologs, or kinesin gene family members seem to be located within the QTL regions detected from the present and other studies. However, in many cases, multiple homologs are present in the same region. Additional work is needed to improve the resolution of QTL loci to associate these QTLs with any of those candidate genes.

Author contribution statement Y. Weng designed this experiment, analyzed the data and wrote the manuscript with input from RG and MC. MC and RG collected data in Michigan trials; YL and Y. Weng collected data in Wisconsin trials. Y. Wang performed QTL analysis of data. MR, AS and RO conducted SNP genotyping of the RIL population. All authors reviewed and approved this submission.

Acknowledgments The authors thank Linda Crubaugh and Kristin Haider for technical assistance. This research was supported by US–Israel Binational Agricultural Research and Development (BARD) fund (Grant number IS-4341-10). Relevant work in Y. Weng's lab was also partially supported by a US Department of Agriculture Specialty Crop Research Initiative grant (project number 2011-51181-30661).

Conflict of interest The authors declare that they have no conflict of interest.

References

Ando K, Grumet R (2010) Transcriptional profiling of rapidly growing cucumber fruit by 454-pyrosequencing analysis. J Amer Soc Hort Sci 135:291–302

Ando K, Carr KM, Grumet R (2012) Transcriptome analyses of early cucumber fruit growth identifies distinct gene modules associated with phases of development. BMC Genom 13:518

Arends D, Prins P, Jansen RC, Broman KW (2010) R/qtl: high-throughput multiple QTL mapping. Bioinformatics 26:2990–2992

Azzi L, Deluche C, Gévaudant F, Frangne NF, Delmas F, Hernould M, Chevalier C (2015) Fruit growth-related genes in tomato. J Expt Bot 66:1075–1086

Bates DM (2010) *lme4*: mixed-effects modeling with R. http://lme4.r-forge.r-project.org/book

Bo KL, Ma Z, Chen JF, Weng Y (2015) Molecular mapping reveals structural rearrangements and quantitative trait loci underlying traits with local adaptation in semi-wild Xishuangbanna cucumber (*Cucumis sativus* L. *var. xishuangbannanesis* Qi et Yuan). Theor Appl Genet 128:25–39

Boonkorkaew P, Hikosaka S, Sugiyama N (2008) Effect of pollination on cell division, cell enlargement, and endogenous hormones in fruit development in a gynoecious cucumber. Scientia Hort 116:1–7

Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890

Candolle AD (1959) Origin of cultivated plants. Hafner Pub. Co., New York

Cavagnaro PF, Senalik DA, Yang L, Simon PW, Harkins TT, Kodira CD, Huang S, Weng Y (2010) Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). BMC Genom 11:569

Cheng ZC, Gu XF, Zhang SP, Mail H, Zhang RW, Liu MM, Yang SJ (2010) QTL Mapping of fruit length in cucumber. China Veg Issue 12:20–25 **(in Chinese)**

Chusreeaeom K, Ariizumi T, Asamizu E, Okabe Y, Shirasawa K, Ezura H (2014) A novel tomato mutant, *Solanum lycopersicum elongated fruit1* (*Slelf1*), exhibits an elongated fruit shape caused by increased cell layers in the proximal region of the ovary. Mol Genet Genomics 289:399–409

Dijkhuizen A, Staub JE (2002) QTL conditioning yield and fruit quality traits in cucumber (*Cucumis sativus* L.): effects of environment and genetic background. J New Seeds 4:1–30

Fu F, Mao W, Shi K, Zhou Y, Yu J (2010) Spatio-temporal changes in cell division, endoreduplication and expression of cell cycle-related genes in pollinated and plant growth substances-treated ovaries of cucumber. Plant Biol 12:98–107

Gillaspy G, Ben-David H, Gruissem W (1993) Fruits: a developmental perspective. Plant Cell 5:1439–1451

He XM, Li Y, Pandey S, Yandell B, Pathak M, Weng Y (2013) QTL mapping of powdery mildew resistance in WI 2757 cucumber. Theor Appl Genet 126:2149–2161

Huang S, Li R, Zhang Z et al (2009) The genome of the cucumber, *Cucumis sativus* L. Nat Genet 41:1275–1281

Jiang L, Yan SS, Yang WC, Li YQ, Xia MX, Chen ZJ, Wang Q, Yan LY, Song XF, Liu RY, Zhang XL (2015) Transcriptomic analysis reveals the roles of microtubule-related genes and transcription factors in fruit length regulation in cucumber (*Cucumis sativus* L.). Sci Rep 5:8031. doi:10.1038/srep08031

Keng H (1974) Economic plants of ancient North-China as mentioned in Shih-Ching (Book of Poetry). Econ Bot 28:391–410

Kennard WC, Havey MJ (1995) Quantitative trait analysis of fruit quality in cucumber: QTL detection, confirmation, and comparison with mating design variation. Theor Appl Genet 91:53–61

Kuznetsova A, Brockhoff PB, Christensen RHB (2013) lmerTest: Tests for random and fixed effects for linear mixed effect models (*lmer* objects of *lme4* package). R package V2.0-3

Marwede V, Schierholt A, Mollers C, Becker HC (2004) Genotype × environment interactions and heritability of tocopherol contents in canola. Crop Sci 44:728–731

Miao H, Gu XF, Zhang SP, Zhang ZH, Huang SW, Wang Y, Cheng ZC, Zhang RW, Mu S, Li M, Zhang ZX, Fang ZY (2011) Mapping QTLs for fruit-associated traits in *Cucumis sativus* L. Sci Agri Sinica 44:5031–5040

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

Owens KW, Bliss FA, Peterson CE (1985) Genetic analysis of fruit length and weight in two cucumber populations using the inbred backcross lines. J Am Soc Hort Sci 110:431–436

Paris HS, Daunay MC, Janick J (2012) Occidental diffusion of cucumber (*Cucumis sativus*) 500–1300 CE, two routes to Europe. Ann Bot 109:117–126

Qi JJ, Liu X, Shen D, Miao H, Xie BY, Li XX, Zeng P, Wang SH, Shang Y, Gu XF et al (2013) A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. Nat Genet 45:1510–1515

Rodriguez GR, Munos S, Anderson C, Sim SC, 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

Rubinstein M, Katzenellenbogen M, Eshed R, Rozen A, Katzir N, Colle M, Yang L, Grumet R, Weng Y, Sherman A, Ophir R (2015) Ultra high-density linkage map for cultivated cucumber using a single nucleotide polymorphism genotyping array. PLoS One 10:e0124101

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

Smith OS, Lower RL, Moll RH (1978) Estimates of heritabilities and variance components in pickling cucumber. J Am Soc Hort Sci 103:222–225

Strefeler MS, Wehner TC (1986) Estimates of heritabilities and genetic variances of three yield- and five quality traits in three fresh-market cucumber populations. J Am Soc Hort Sci 111:599–605

Van der Knaap E, Chakrabarti M, Chu YH, Clevenger JP, Illa-Berenguer E, Huang ZJ, Keyhaninejad N, Mu Q, Sun L, Wang YP, Wu S (2014) What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. Front Plant Sci 5(article 227):1–13

Wang M, Liu SL, Zhang SP, Miao H, Wang Y, Tian GL, Lu HW, Gu XF (2014) Quantitative trait loci associated with fruit length and stalk length in cucumber using RIL population. Act Bot Boreal Occident Sin 34:1764–1770 **(in Chinese with English abstract)**

Wei QZ, Wang YZ, Qin XD, Zhang YX, Zhang ZT, Wang J, Li J, Lou QF, Chen JF (2014) An SNP-based saturated genetic map and QTL analysis of fruit-related traits in cucumber using specific-length amplified fragment (SLAF) sequencing. BMC Genom 15:1158

Whitaker TW, Davis GN (1962) Cucurbits. InterScience Publishers, London

Yandell BS, Mehta T, Banerjee S, Shriner D, Venkataraman R, Moon JY, Neely WW, Wu H, Smith R, Yi N (2007) R/qtlbim: QTL with Bayesian interval mapping in experimental crosses. Bioinformatics 23:641–643

Yang LM, Koo DH, Li Y, Zhang X, Luan F, Havey MJ, Jiang J, Weng Y (2012) Chromosome rearrangements during domestication of cucumber as revealed from high-density genetic mapping and draft genome assembly. Plant J 71:895–906

Yang XY, Wang Y, Jiang WJ, Liu XL, Zhang XM, Yu HJ, Huang SW, Liu GQ (2013a) Characterization and expression profiling of cucumber kinesin genes during early fruit development: revealing the roles of kinesins in exponential cell production and enlargement in cucumber fruit. J Exp Bot 64:4541–4557

Yang LM, Li DW, Li YH, Gu XF, Huang SW, Garcia-Mas J, Weng Y (2013b) A 1,681-locus consensus genetic map of cultivated cucumber including 67 NB-LRR resistance gene homolog and ten gene loci. BMC Plant Biol 13:53

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 (2007) Genetic linkage map construction and location of QTLs for fruit-related traits in cucumber. Plant Breed 127:180–188

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

Zeng ZB, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. Genet Res 74:279–289

Zhou XY, Qin ZW, Wang XG (2005) Study of the market characters of cucumber (*Cucumis sativus* L.) germplasm. J Northeast Agri Univ 36:707–713 **(In Chinese with English abstract)**