Theoretical and Applied Genetics (2020) 133:2271–2290
https://doi.org/10.1007/s00122-020-03596-y

ORIGINAL ARTICLE

QTL for horticulturally important traits associated with pleiotropic *andromonoecy* and *carpel number* loci, and a paracentric inversion in cucumber

Yupeng Pan[1,2] · Changlong Wen[2,3] · Yonghua Han[4] · Yuhui Wang[2] · Yuhong Li[1,2] · Sen Li[2,5] · Xiaomao Cheng[2] · Yiqun Weng[2,6]

Received: 6 February 2020 / Accepted: 5 April 2020 / Published online: 18 April 2020

Abstract

The legendary cucumber inbred line WI2757 possesses a rare combination of resistances against nine pathogens, which is an important germplasm for cucumber breeding. However, WI2757 flowers late and does not perform well under field conditions. The genetic basis for horticulturally important traits other than disease resistances in WI2757 is largely unknown. In this study, we conducted QTL mapping using F_2 and recombinant inbred line (RIL) populations from the WI2757 × True Lemon cross that were segregating for multiple traits. Phenotypic data were collected in replicated field trials across multiple years for seven traits including fruit carpel number (CN) and sex expression. A high-density SNP-based genetic map was developed with genotyping by sequencing of the RIL population, which revealed a region on chromosome 1 with strong recombination suppression. The reduced recombination in this region was due to a ~ 10-Mbp paracentric inversion in WI2757 that was confirmed with additional segregation and cytological (FISH) analyses. Thirty-six QTL were detected for flowering time, fruit length (FL), fruit diameter (FD), fruit shape (LD), fruit number (FN), CN, and powdery mildew resistance. Five moderate- or major-effect QTL for FL, FD, LD, and FN inside the inversion are likely the pleiotropic effects of the *andromonoecy* (*m*), or the *cn* locus. The major-effect flowering time QTL *ft1.1* was also mapped inside the inversion, which seems to be different from the previously assigned *delayed flowering* in WI2757. Implications of these findings on the use of WI2757 in cucumber breeding are discussed.

Communicated by Amnon Levi.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-020-03596-y) contains supplementary material, which is available to authorized users.

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

[1] College of Horticulture, Northwest A&F University, Yangling 712100, Shaanxi, China

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

[3] Beijing Academy of Agricultural and Forestry Sciences, Beijing 100097, China

[4] Institute of Integrative Plant Biology, School of Life Sciences, Jiangsu Normal University, Xuzhou 221116, China

[5] Horticulture College, Shanxi Agricultural University, Taigu 030801, China

[6] USDA-ARS Vegetable Crops Research Unit, 1575 Linden Dr., Madison, WI 53706, USA

Introduction

Cucumber, *Cucumis sativus* L. ($2n = 2x = 14$), is an important fruit vegetable crop worldwide, and disease resistance is always among top priorities in cucumber breeding. The legendary cucumber inbred line WI2757 (Wisconsin 2757) has a rare combination of resistances against nine pathogens that cause significant economic losses in cucumber production. In growth chamber and greenhouse tests, WI2757 was highly resistant to the fungal pathogens *Podosphaera fusca* (formerly *Sphaerotheca fuliginea*) (powdery mildew, PM), *Cladosporium cucumerinum* (scab), *Corynespora cassiicola* (target leaf spot, TLS), *Colletotrichum orbiculare* (syn. *Colletotrichum lagenarium*) (anthracnose, AR), and *Fusarium oxysporum* f. sp. *cucumerinum* (Fusarium wilt, FW). It also confers resistance to the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis* (downy mildew, DM), the bacterial *Pseudomonas syringae* pv. *lachrymans* (angular

leaf spot, ALS), and *Erwinia trucheiphila* (bacterial wilt, BW), as well as the viral pathogen *Cucumber Mosaic Virus* (CMV) (Peterson et al. 1982). WI2757 was released almost 40 years ago, but it still holds high level of resistances to all of these diseases except for DM to which it has moderate resistance due to the appearance of new virulent field strain(s) in 2004 in the USA (Thomas et al. 2017; Wang et al. 2019a). Over the years, WI2757 has been a popular and important source of disease resistances in cucumber breeding.

WI2757 has a complicated pedigree derived from many varieties or plant introduction (PI) lines, which are illustrated in supplemental Fig. S1A. Main sources of disease resistance in WI2757 include the US pickling inbred line Gy14 and the Dutch Beit alpha-type cucumber variety 'EXPO'. Gy14 was a selection from the progeny of cross between SMR 18 (with resistance to scab, FW, and CMV) and Gy3 that has various donors of resistance genes such as PI 197087, PI 220860, PI 212233, PI 220860, and PI 234517 (Peterson et al. 1982; Wehner and Robinson 1991; Chung et al. 2003) (Fig. S1A). The genetic bases of disease resistances in Gy14 and WI2757, or their relevant lines, have been extensively studied (reviewed in Wang et al. 2020). For example, the AR, DM, and ALS resistances in WI2757 and Gy14 were derived primarily from PI 197087, which were all controlled by single, recessive genes (*cla* for AR, *dm1* for DM, and *psl* for ALS) (Fanourakis and Simon 1987; Wyszogrodzka et al. 1987; Kennard et al. 1994; Horejsi et al. 2000; Olczak-Woltman et al. 2007; Słomnicka et al. 2016). Recently, it was found that the cucumber *STAYGREEN* (*CsSGR*) gene is underlying the triple disease resistance loci *dm1*/*psl*/*cla* in WI2757 and Gy14 (Pan et al. 2018; Wang et al. 2019a). However, as compared with Gy14, WI2757 harbors addition minor-effect QTL for DM (*dm5.2*), AR (*cla3.1*), and ALS (*psl1.1* and *psl3.1*). Using segregating populations derived from the cross between WI2757 and True Lemon (TL), He et al. (2013) identified six QTL for seedling-stage PM resistance with *pm5.2* having the largest effect, which was later shown to be a loss-of-susceptibility gene encoding the MLO protein (Nie et al. 2015a, b; Berg et al. 2015).

Two single dominant genes control resistances to FW and scab in WI2757, which are tightly linked (Vakalounakis 1993, 1995; Mao et al. 2008). In the cucumber line 9110Gt, the major-effect QTL, *qFoc2.1* for FW resistance, and the *Ccu* locus for scab resistance have been mapped within a cluster of NB-LRR resistance gene homologs on Chr2 (Kang et al. 2010; Zhang et al. 2010, 2014). WI2757 probably shares the same FW/scab resistance loci as 9110Gt. Molecular mapping for BW, CMV, and TLS in WI2757 has not been reported. Wen et al. (2015) fine mapped a TLS resistance locus *cca-3* in the inbred line D31 on Chr6 with a CC-NB-ARC type *R* gene as the candidate. The TLS resistance in D31 was probably derived from WI2757.

WI2757 is gynoecious and parthenocarpic with bitterfree fruit, fine and white spines and smooth, and tender and dark green skin, which is morphologically similar to the Mediterranean-type (Beit alpha, or mini) cucumber (Fig. S1B–E). WI2757 lacks the coarse spines demanded by the American market (US pickling and slicing cucumbers); thus, it has little direct value as a processing cultivar for production in the USA, but may be more useful for the fresh consuming cucumbers targeting the Mediterranean or European market (Peterson et al. 1982). In commercial production, in addition to a good disease resistance package, each cucumber market class has its own specific requirements for many horticulturally important traits such as flowering time, fruit size (fruit length and diameter) and shape, fruit number, as well as fruit epidermal features (Weng et al. 2015; Wang et al. 2020). WI2757 was selected under greenhouse and growth chamber environments (Peterson et al. 1982), and it does not perform well under field conditions both home and abroad. In particular, WI2757 flowers late despite of its gynoecious sex expression, which was proposed to be controlled by a single recessive gene, *delayed flowering* (*df*) (Fanourakis and Simon 1987; Walters et al. 2001). Little is known about the genetic basis of these horticultural traits in this important germplasm line.

In cucumber and other cucurbit crops, fruit size, shape, and number are known to be affected by the *andromonoecy* (*m*) and *carpel number* (*Cn*) loci (reviewed by Pan et al. 2020). While Li et al. (2016) showed pleiotropic effect of the *Cn* locus on fruit size and weight, the association of fruit size and shape with the *m* locus has never been explicitly demonstrated in cucumber. The 'True Lemon' (TL) cucumber is an heirloom that is very popular among home gardeners for its lemon-shaped fruit. True Lemon is andromonoecious bearing male and perfect flowers and sets characteristic lemon size, nearly round fruit with five carpels (Fig. 1). It has normal flowering time and is susceptible to almost all diseases. Major contrasting traits between WI2757 and True Lemon are summarized in supplemental Table S1. Thus, the main objectives of the present study are to (1) investigate the genetic basis of several horticultural traits associated with WI2757 including flowering time, fruit size/shape, fruit number, and PM resistance and (2) clarify possible pleiotropic effects of the *m* and *cn* loci on these traits. The overall goal of this study was to gain a more comprehensive understanding of various horticultural traits in WI2757 for more efficient use of this germplasm in cucumber breeding.

In this study, we developed segregation F_2 and recombinant inbred line (RIL) populations from the cross between WI2757 and TL, which were used for phenotyping seven traits (fruit length, diameter, shape, fruit number, carpel number, flowering time, and PM resistance) in both greenhouse and field trials. A linkage map was developed with 129 microsatellite (simple sequence repeats, SSRs) markers and 139

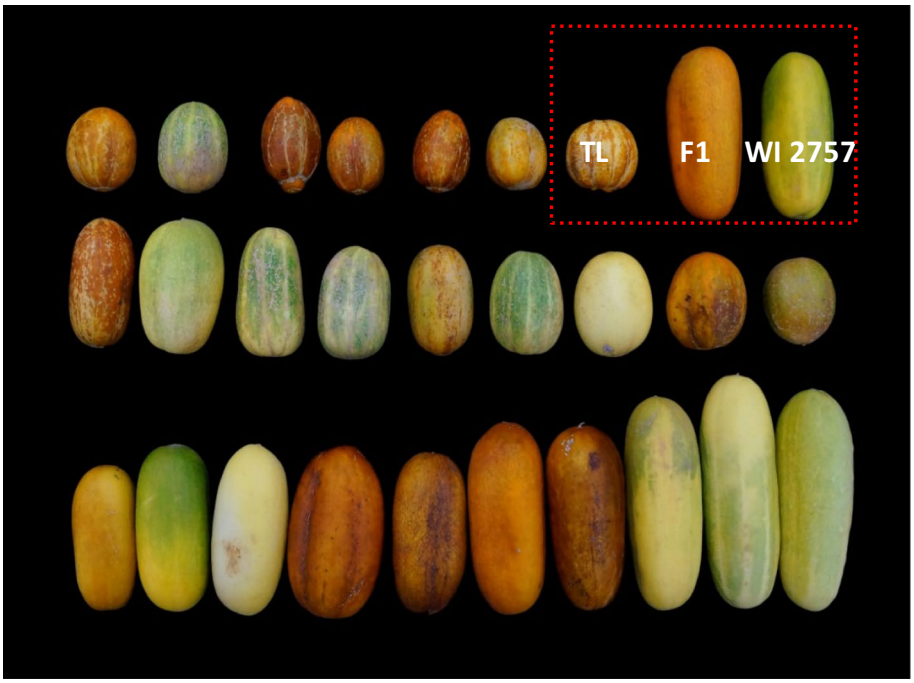

Fig. 1 Representative fruit images of WI2757, TL (TL), their F_1, and derived RILs

RILs. A high-density genetic map was also developed with 1845 SNP loci with genotyping-by-sequencing (GBS)-based SLAF-Seq. From linkage mapping, we identified a 10-Mbp large segmental inversion on Chr1 of WI2757, which was confirmed in further cytological analysis and linkage analysis in additional segregating populations. QTL mapping in both F_2 and RIL populations identified 36 QTL for seven traits. The results from these experiments revealed the important effects of the large inversion in WI2757 and the pleiotropic effects of the *m* and *cn* loci on QTL detection of the seven traits in the WI2757 × TL segregating populations.

Materials and methods

Plant materials

Two cucumber inbred lines, WI2757 and True Lemmon (TL), were the main subjects in the present study. WI2757 is a gynoecious, Beit alpha-type germplasm line known for its multiple disease resistances (Fig. S1; Peterson et al. 1982). TL is an andromonoecious heirloom with the characteristic lemon size, round, and yellow-striped fruits widely used by home gardeners. Two parental lines show a number of morphological and disease resistance differences (Table S1). Traits under investigation in the present study included sex expression, flowering time, fruit size, and shape. For linkage map development and QTL analysis, two populations were developed from the cross between WI2757 and TL including F_2 populations of different sizes and 139 F_6 RILs. These RILs were derived from the F_2 plants used by He et al. (2013) through single-seed descent (SSD).

Since high-density genetic mapping suggested a large inversion on Chr1, ad hoc F_2 populations from WI2757 × WI7200 (89 plants) and TL × WI7200 (96 plants) were developed and used to construct regional genetic maps on Chr1. To track the origin of the inversion in WI2757, four germplasm lines, Gy14, PI 197087, PI 212233, and PI 220860 that are known to be in the pedigree of WI2757 (Fig. S1) were used in fluorescence in situ hybridization (FISH).

Phenotypic data collection

Phenotypic data from F_2 individuals and RILs were collected in six experiments over 6 years, namely WI2009F2,

WI2013F2, WI2014RIL, WI2015RIL, WI2016RIL, and WI2017F2. The details of each experiment are provided in supplemental Table S2. The WI2009F2 experiment was conducted in the Walnut Street Greenhouses (WSGH) of the University of Wisconsin at Madison, and the rest were in open fields at the Hancock Agriculture Research Station, Hancock, Wisconsin (HARS). The two parental lines and their F_1 were included in all experiments for data collection.

Target traits for phenotypic data collection in the six experiments included flowering time (FT), mature fruit length (FL) and diameter (FD), fruit number (FN), fruit carpel number (CN), and powdery mildew (PM) resistances. Sex expression of each plant in all populations was also recorded. Not all traits were phenotyped in all experiments (Table S2). Data collection for the F_2 population was on the individual plant basis. For RILs, data were collected from each plant per family per replication and the RIL means were used in QTL analysis. Flower time was recorded as the dates from sowing to anthesis of the first female or perfect flower on the plant. Measurement of FL, FD, LD, and FN at the mature fruit stage (> 35 days after pollination) followed Weng et al. (2015) and Bo et al. (2015). Counting of carpel numbers of each fruit was based on Li et al. (2016). For sex expression, each plant was recorded as andromonoecious (with both male and bisexual flowers), gynoecious (with only female flowers), and monoecious (with both male and female flowers). The rating of PM disease scores on RIL plants under natural infection in the field was based on percentage of diseased areas on leaves with a 1–9 scale where 1 = no symptom and 9 = > 90% diseased area or dead (He et al. 2013; Wang et al. 2019a). In each season, the RIL population was scored twice (1–2 weeks apart), and the mean disease scores of each RIL were used in QTL analysis.

Genotyping and linkage map development

A linkage map with 236 SSR loci and 132 WI2757 × TL F_2 plants was developed previously for QTL mapping of PM resistance (He et al. 2013), which was also used in the present study for QTL analysis of FT, FL, FN, LD, and CN collected in the WI2009F2 experiment. From the early map, 129 polymorphic SSR markers evenly distributed across seven chromosomes (Chr) were used to develop a genetic map with 139 WI2757 × TL F_6 RILs. A subset of 87 RILs were genotyped with SLAF-Seq based on high-throughput genotyping by sequencing (GBS) (Sun et al. 2013). SLAF-Seq library, Illumina Hi-Seq 2500 sequencing, raw reads processing, filtering, SNP calling, and linkage map construction were all the same as described early (Wang et al. 2016, 2018). A combined linkage map was developed with 87 RILs, 1845 SNP, and 129 SSR markers using the function '*mstmap*' in the *R/ASMap* package (Wang et al. 2018).

To confirm a possible inversion identified from linkage analysis in the RIL population, additional linkage mapping in the target region was performed in a larger WI2757 × TL F_2 population with 340 plants with 14 polymorphic SSRs from He et al. (2013). This population was also phenotyped for various traits in the WI2017F2 field experiment with the intention to refine the locations of QTL detected in the RIL population. After the confirmation of the inversion, two regional maps on the long arm of Chr1 were further developed with 5–9 polymorphic SSR markers in the WI2757 × WI7200 (89 individuals) and WI7200 × TL (96 plants) F_2 populations, which revealed WI2757 harboring the ~ 10-Mbp inversion.

DNA extraction, molecular marker analysis, and linkage analysis all followed standard protocols described in our previous studies (e.g., Pan et al. 2017a, b).

Fluorescence in situ hybridization (FISH) for visualizing segmental inversion

To better visualize the segmental inversion on Chr1, fosmid clones located in the target region were selected from Yang et al. (2014) and used as probes in FISH analysis. Chromosome preparation and FISH procedures followed Yang et al. (2014). Briefly, root tips were harvested from germinated seeds, pretreated in 4 °C water for 2–4 h to capture pro-metaphase and metaphase cells, and fixed in Carnoy's solution (3 ethanol: 1 glacial acetic acid). Root tips were then macerated in 2% cellulose and 1% pectolyase at 37 °C for 2 h, and squashes were prepared using the same fixative. DNA probes were labeled with digoxigenin-dUTP or biotin-dUTP via nick translation and detected with antidigoxigenin antibody coupled with rhodamine (Roche) or avidin conjugated with FITC (Vector Laboratories), respectively. Chromosomes were counterstained by 4,6-diamidino-2-phenylindole (DAPI) in a Vecta Shield antifade solution (Vector Laboratories). Images were captured digitally using a CCD camera (QIMAGING, RETIGA-SRV, FAST 1394) attached to an Olympus BX63 epifluorescence microscope.

QTL analysis

QTL analysis was performed with the R/qtl software package (http://www.rqtl.org/) (Broman et al. 2003). The initial whole genome scan for QTL was conducted with a window size of 25 cM. Refinement of QTL number and location was then performed with both composite interval mapping (CIM) and multiple QTL mapping (MQM) methods using a 10 cM window size following Weng et al. (2015) and Pan

et al. (2017b). Genome-wide LOD thresholds ($P < 0.05$) for declaring the presence of QTL were determined using 1000 permutations. For each detected QTL, a 1.5-LOD-support interval was calculated and defined by left and right markers. Naming of QTL followed the nomenclature recommendations by Pan et al. (2020) and Wang et al. (2020). For example, *ft1.1* and *fl6.2* designated the first QTL for flowering time and the second QTL for mature fruit length on cucumber Chr1 and Chr6, respectively.

Results

Phenotypic variation of flowering time, fruit size/shape-related traits, and PM resistance in WI2757 × TL segregating populations

We collected phenotypic data for seven traits (FL, FD, LD, FN, FT, CN, and PM) from the F_2 and RIL populations in six experiments (Table S2). The phenotypic means, standard deviation, the range, and estimated heritability (from RILs) of these traits in five of the six experiments are presented in Table 1 (WI2013F2 was not included because only FT data were collected). The frequency distribution of these traits in all experiments is illustrated in supplemental Fig. S2. Representative fruit images of the two parents, their F_1, and RIL plants are shown in Fig. 1.

All seven traits showed high heritability (Table 1). Among them, FL, LD, and FT (earlier flowering) exhibited heterosis (i.e., F_1 plants had larger fruits and earlier flowering time than either parent). CN and FN of F_1 were similar to WI2757, whereas FD and PM disease scores of F_1 were closer to TL (Table 1). All traits except CN showed largely continuous distribution in the F_2 and RIL populations suggesting their quantitative nature. Transgressive segregation was observed for all traits (Fig. S2) indicating different genetic architectures of each trait between the two parental lines. For example, in all trials, WI2757 consistently flowered later than that of TL, and the F_1 plants flowered earlier than either parent, suggesting that earlier flowering time is dominant to later flowering. In the F_2 and RIL populations, there were many plants with much earlier flowering dates than the F_1 or later than WI2757 (Fig. S2).

The WI2757 × TL F_2 and RIL populations were segregating for the andromonoecious (*m*) and carpel number (*cn*) loci, both of which are located on Chr1, and are physically ~ 15 Mbp away from each other. Monoecious/gynoecious (*M_*) sex expression is dominant to andromonoecious (*mm* in TL); three carpels (*Cn_*) are dominant to five carpels (*cncn*, in TL) (Tan et al. 2015; Li et al. 2016). In cucurbits, both loci are known to exert pleiotropic effects on fruit size, shape, and fruit number (reviewed

Table 1 Phenotypic means and range of horticulturally important traits in WI2757, TL, their F_1, F_2 and RIL populations in six experiments

Traits	WI2757	TL	F_1	WI2009F2 ($n = 132$)		WI2017F2 ($n = 340$)		WI2014RIL ($n = 91$)		WI2015RIL (i = 139)		WI2016RIL ($n = 139$)		Heritability (h^2)[c]
	Mean[a]	Mean[a]	Mean[a]	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	
FL (cm)	14.7	8.8	17.8	13.9 ± 3.2	7.0–20.8	13.7 ± 3.6	5.6–20.6	12.6 ± 3.3	6.7–20.1	12.3 ± 3.0	7.8–18.8	12.8 ± 3.6	7.9–20.7	0.97
FD (cm)	7.7	8.7	8.6	6.6 ± 0.7	4.8–8.0	7.8 ± 1.0	5.1–11.1	7.6 ± 0.9	5.6–11.8	7.9 ± 0.7	6.4–9.7	8.5 ± 0.8	6.7–11.3	0.77
LD	1.9	1.0	2.1	2.1 ± 0.4	1.1–3.2	1.8 ± 0.4	0.9–2.6	1.6 ± 0.4	1.0–2.4	1.6 ± 0.4	1.0–2.3	1.5 ± 0.4	1.0–2.3	0.97
FN	6.7	12.3	6.7	NA	NA	6.6 ± 4.4	1–31	9.3 ± 6.0	2–31	8.1 ± 5.3	1–32	12.5 ± 7.6	1–45	0.88
FT (day)	61.2	51.4	46.6	73.2 ± 12.1	55–98	51.1 ± 5.9	38–74	57.5 ± 7.2	37–78	50.5 ± 6.6	40–76	45.2 ± 7.0	27–63	0.90
CN	3.1	4.9	3.1	3.3 ± 0.5	3.0–5.0	NA	NA	3.9 ± 0.7	2.9–5.0	4.0 ± 0.7	2.9–5.0	NA	NA	0.98
PM[b]	2.2	7.4	6.8	NA	NA	NA	NA	4.7 ± 2.2	1.0–9.0	5.8 ± 2.0	2.0–9.0	6.0 ± 2.0	2.0–9.0	0.84

FL/FD fruit length/diameter, *LD* FL/FD, *FN* fruit number, *FT* flowering time, *CN* carpel number, *PM* powdery mildew disease scores, *NA* not applicable or no data

[a]Based on data across all field experiments

[b]Mean disease scores of two ratings in each experiment (when applicable)

[c]The heritability is calculated based on the phenotypic data collected in the RIL populations

in Pan et al. 2020; also see "Discussion" section), which could explain most of the phenotypic variations observed among the F_2 and RIL populations in this study. TL sets many (mean FN = 12.3 per plant), nearly round (LD ≈ 1.0) fruits with five carpels (mean CN = 4.9), while WI2757 bears fewer (mean FN = 6.7 per plant), oblong (LD > 2.0) fruits with mean CN = 3.1. The F_1 had fruits with similar FL, LD, and FN with WI2757 (Table 1; Fig. 1). To better understand the pleiotropic effects of the *m* locus, the means of FL, FD, LD, FN, and FT in the six experiments were calculated by sex (*mm* vs. *M_*) in each population, and the reorganized data are presented in supplemental Table S3. Boxplots and significance tests of population means between the two sex groups in each experiment are illustrated in Fig. 2. In each experiment, as compared with *M_* or *MM* (gynoecious or monoecious) plants, the mean values for FL, FD, and LD were lower in andromonoecious (*mm*) plants, but higher for FN and CN. The mean values for each trait in the F_2 and RIL populations were somewhat different, though, which could be explained by the different percentages of andromonoecious plants in the F_2 (1/4, or ~ 25%) and RIL (1/2, or ~ 50%) (Table S3) populations.

The pleiotropic effect of the *m* locus could also be seen from trait correlations. We calculated the Spearman's rank correlation coefficients (r_s) among morphological traits in different experiments (Table 2). FL exhibited strong and positive correlations with both FD and LD in all experiments, while no or weak correlation was found between FD and LD. This may suggest that FL and FD share some genetic basis, and FL dictates LD, which is a compound trait

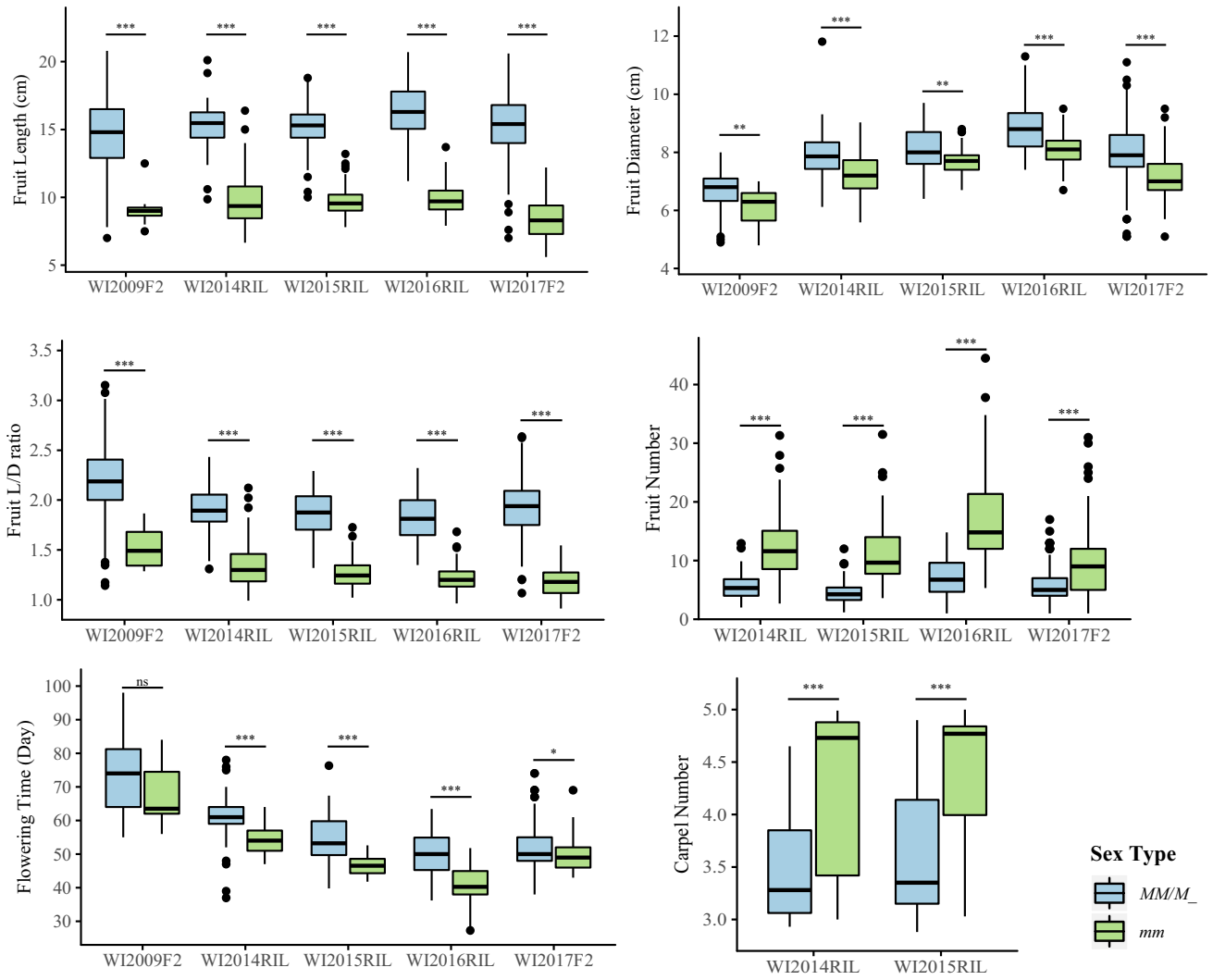

Fig. 2 Boxplots of population means of flowering time (FT), fruit length (FL), fruit diameter (FD), length-to-diameter ratio (LD), fruit number (FN) per plant, and carpel number (CN) of gynoecious or monoecious (*M_*) and andromonoecious (*mm*) plants in different experiments. Comparison in each group pair is conducted with the Wilcoxon rank sum test implemented in *R*. Significance level: $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; *ns* not significant

Table 2 Spearman's rank correlation coefficients among morphological traits in WI2757 × TL populations

	Environments	FT	FN	FL	FD	LD
FN	WI2009F2	n/a				
	WI2014RIL	−0.3814***				
	WI2015RIL	−0.5752***				
	WI2016RIL	−0.5106***				
	WI2017F2	−0.3912***				
FL	WI2009F2	ns	n/a			
	WI2014RIL	0.3583***	−0.4908***			
	WI2015RIL	0.4700***	−0.6113***			
	WI2016RIL	0.5094***	−0.6969***			
	WI2017F2	ns	ns			
FD	WI2009F2	ns	n/a	0.4813***		
	WI2014RIL	ns	ns	0.5747***		
	WI2015RIL	ns	−0.2263**	0.3669***		
	WI2016RIL	0.3772***	−0.4766***	0.5684***		
	WI2017F2	−0.2896***	ns	0.6107***		
LD	WI2009F2	ns	n/a	0.8698***	ns	
	WI2014RIL	0.3485***	−0.4823***	0.9050***	ns	
	WI2015RIL	0.5067***	−0.5889***	0.9309***	ns	
	WI2016RIL	0.4850***	−0.6376***	0.9301***	0.2661**	
	WI2017F2	ns	ns	0.8314***	ns	
CN	WI2009F2	ns	n/a	−0.3292***	ns	−0.4245***
	WI2014RIL	−0.3230**	0.2727**	−0.6192***	ns	−0.7302***
	WI2015RIL	−0.3273***	0.3002***	−0.6415***	ns	−0.7610***

ns not significant, *n/a* data not available, *FT* flowering time, *FN* fruit number, *FL* fruit length, *FD* fruit diameter, *LD* fruit length by diameter, *CN* carpel number

***$P < 0.001$; **$P < 0.01$

depending on both FL and FD (Bo et al. 2015). We observed a weak positive correlation of CN with FN (in RILs only) and a moderate negative correlation with FL/LD (both F_2 and RIL populations). Li et al. (2016) suggested possible pleiotropic effect of the *Cn* locus on fruit size and weight: in monoecious background, more fruits with more carpels tend to be larger and heavier. We reorganized the FL, FD, LD, FN, and FT phenotypic data from three experiments (WI2009F2, WI2014RIL, and WI2015RIL) by carpel number (use CN = 3.5 as the cutoff) and sex expression, which are presented in Supplemental Table S4. The boxplots and significance tests of population means between the high (CN ≥ 3.5) and low (CN < 3.5) plants in two sex groups in each experiment are presented in supplemental Fig. S3. When only monoecious and gynoecious plant (genotype *M*_) were considered (blocking of *m* effects), no differences were observed in FT and FN means between CN < 3.5 (WI2757 type) and CN ≥ 3.5 (TL type) plants. In general, plants with higher CN showed slightly lower FL and LD, but higher FD than those with CN < 3.5 (Table S4). Therefore, the observed phenotypic variation of FD could at least partially be attributed to the pleiotropic effect of the *cn* locus (see "Discussion" section).

Not all traits were associated with pleiotropic effects of the *m* and *cn* loci. In all populations, andromonoecious (*mm*) plants in general flowered earlier than gynoecious or monoecious plants (Table S3). There were negative correlations between FT and FN, as well as between FT and CN (RIL population only), implying that early flowering plants set more fruits with more carpels. In the RIL population, FT positively correlated with FL and LD, but such correlation was not significant in the F_2. The correlation between FT and FD was dependent on the environments and populations: In WI2016RIL and WI2017F2, they showed positive and negative correlations, respectively, whereas no correlation was found between FT and FD in the WI2009F2, WI2014RIL, and WI2015RIL experiments (Table 2). The correlations of FT with other traits are likely due to the recombination suppression inside a large paracentric inversion in WI2757 in which both the *M* and *df* (delayed flowering) loci are located.

Under field natural infections, WI2757 was highly resistant to powdery mildew. TL and F_1 were both susceptible to the PM pathogen, and the mean disease scores in the RIL population showed largely normal distribution (Table 1; Fig. S2), which was consistent with the quantitative nature of PM resistance in WI2757 (He et al. 2013). No significant

correlation was found between mean PM disease scores and any other traits in the RIL population (data not shown).

High-density linkage mapping in WI2757 × TL F_2 and RIL populations reveals recombination suppression in the lower distal end of Chr1

Using the WI2757 × TL F_2 population, we previously developed a linkage map with 249 SSR marker loci for QTL mapping of PM resistance in WI2757 (He et al. 2013). We advanced these F_2 plants through SSD and developed 139 F_6 RILs. A linage map was developed using the 139 RILs and 129 polymorphic SSR markers selected from the early map by He et al. (2013). Detailed information of this map is presented in supplemental Table S5, which spanned 871.4 cM in seven linkage groups (LG = chromosome). Four cloned genes were added onto the map based on their locations in the genome including *cn* (Li et al. 2016), *m*, *de* (for *determinate growth habit*) (Wen et al. 2019), and *F* (for *femaleness*, or gynoecious sex expression) (Li et al. 2020). Each marker was also aligned against the Gy14v2.0 draft genome. The marker order along each LG was highly consistent with their physical positions, but there were some large gaps (> 10 cM) on each chromosome. In addition, in the distal end of the long arm of Chr1, there were five markers (UW084366, UW083725, UW083752, UW084651, and UW084538) that did not show any recombination among the 139 RILs, which physically spanned at least 10 Mbp including the *m* locus (Table S5).

To increase the marker density on the map, a subset of 87 RILs were genotyped with SLAF-seq. In total, 6.48 Gb sequences from 40,504,760 150-bp-paired-end reads were generated from high-throughput Illumina sequencing. Among them, 80.1% reads had > Q30 quality score, from which 50,256 high-quality SLAF tags were obtained, and 6491 (12.9%) were polymorphic between WI2757 and TL. Distribution of these polymorphic SLAFs across seven chromosomes is presented in supplemental Table S6. After filtering with the same criteria used by Wang et al. (2018), 1845 SNP markers were kept for linkage analysis. A genetic map combining the 1845 SNPs and 129 SSRs (total 1974) was constructed with 87 RILs. Main statistics of the map are summarized in Table S6, and the complete information of the map is presented in supplemental Table S7.

We evaluated the quality of the high-density linkage map by alignment of mapped markers against the Gy14v2.0 draft genome, which are graphically presented in Fig. 3a (for Chr1) and supplemental Fig. S4 (for Chr2 to Chr7). Physically, this map covered the majority of the cucumber genome, and the marker orders on the genetic map were highly congruent with their physical positions in the draft Gy14v2.0 genome. On Chr2 to Chr7, the genetic-to-physical distance ratios across the whole chromosomes were largely linear except in the centromeric regions (Fig. S4). However, in a ~ 10-Mbp region at the bottom of Chr1 (23.4–33.2 Mbp), there were 224 marker loci that spanned only ~ 15 cM (from 245.65 to the end at 261.78 cM); 212 of the 224 (95%) markers were in four clusters (Table S7) suggesting significantly reduced recombination in this region (Table S7, Fig. S4), which was consistent with the finding from the SSR-based RIL map in this region (Table S5).

Additional evidence supports a ~ 10 Mbp segmental inversion in WI2757

To confirm the recombination suppression in the Chr1 region observed in the RIL population, we developed a local linkage map in a larger WI2757 × TL F_2 population with 340 plants (WI2017F2). Linkage analysis was performed in this population with 14 SSR markers from the long arm of Chr1. The resulting genetic map is shown in Fig. 3b. Consistent with the RIL map, there was clear marker clustering in this region despite the much larger F_2 population used.

Recombination suppression often occurs in chromosomal regions heterozygous for large structural variations (SV) such as inversions (e.g., Ren et al. 2009; Yang et al. 2012). The above observations suggested the SV must be present in either WI2757 or TL. To identify the origin of this SV, linkage maps were developed for the Chr1 region with two segregating populations including 96 F_2 plants from WI7200 × TL and 89 F_2 plants from WI2757 × WI7200. We tested 14 SSR markers in the suspected inversion region, and 5 and 9 were polymorphic between the two pairs of parental lines, respectively. Linkage maps from the two F_2 populations are illustrated in Fig. 3b. The five markers on the WI7200 × TL map were well separated indicating their normal recombination and segregation. However, all nine markers in the WI2757 × WI7200 F_2 population were clustered, suggesting that WI2757 carries the SV (Fig. 3b).

To confirm the presence and the nature of the SV in WI2757, we conducted fluorescence in situ hybridization (FISH) on metaphase chromosomes in WI2757, Gy14 and 9930 cucumber lines. Three fosmid clones located in the ~ 10-Mbp target region (Yang et al. 2014; details in supplemental Table S8) were used as probes in FISH analysis. The results are shown in Fig. 3c. The physical locations of signals for the probes 255H13 (at ~ 24.6 Mbp on Gy14_v2.0), 255I22 (at ~ 27.4 Mbp), and 32O20 (at distal end of Chr1; ~ 32.4 Mbp) in 9930 and Gy14 chromosomes were consistent with their draft genome positions, but the order was reversed in WI2757. The centromere of Chr1 is located at approximately 12.0–13.0-Mbp region. This FISH work provided convincing evidence that a large paracentric inversion (~ 10 Mbp, from 23.4 Mbp to the end) is present in the long arm of Chr1 of WI2757 cucumber.

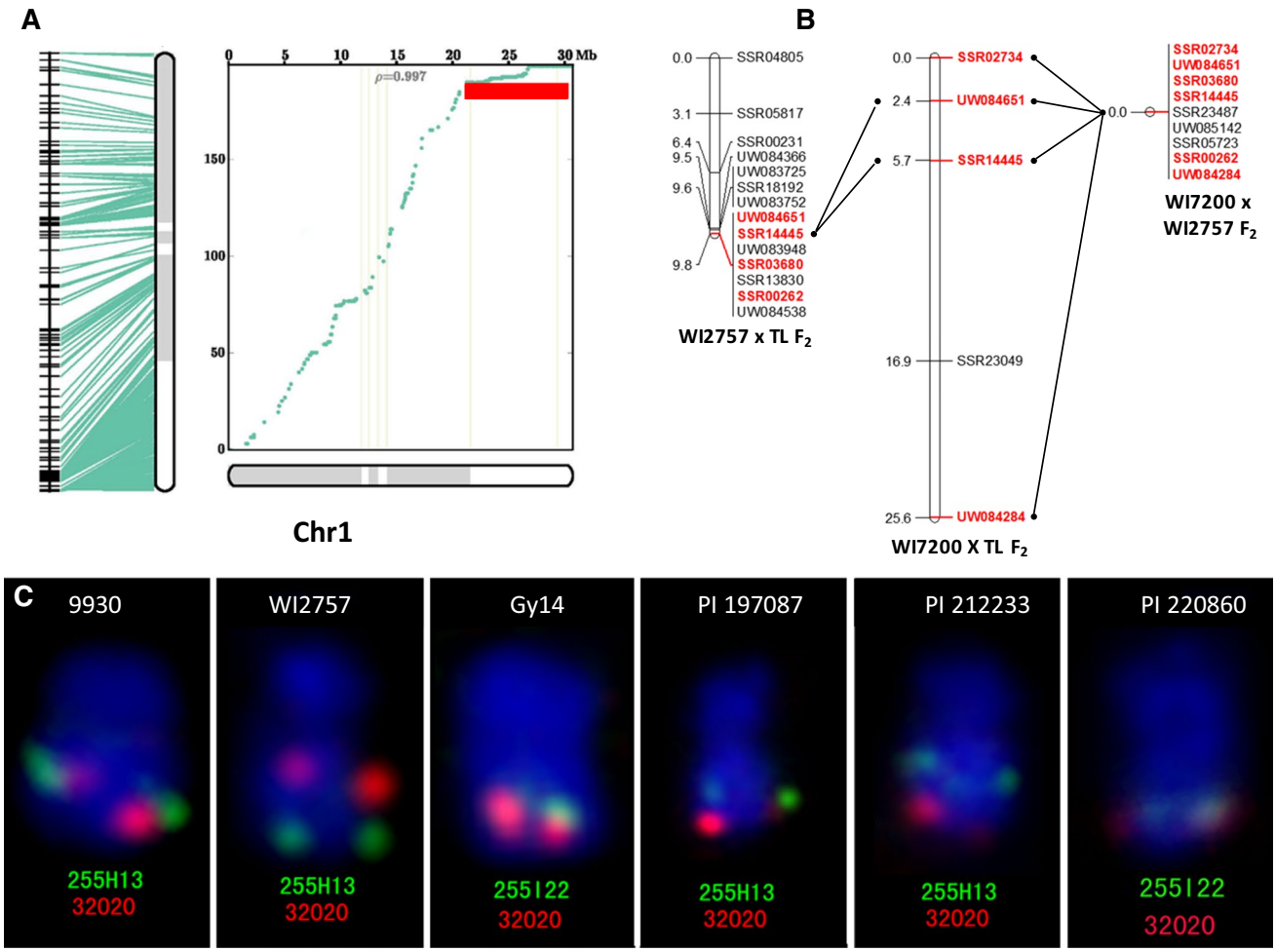

Fig. 3 Large inversion (~ 10 Mbp) on Chr1 of WI2757 cucumber. **a** Distribution of mapped SNP loci against physical map of Chr1 (Gy14v2.0). The left panel is side-by-side alignment between genetic (left) and physical (right) maps. The right panel is a scatter plot comparing the genetic distance (Y axis, in cM) in relation to physical distance (X axis, in Mbp). Gray and white regions represent different scaffolds in the assembly. The ρ value is Pearson correlation coefficient. (Values closer to − 1 and 1 indicate near-perfect colinearity.) Horizontal red bar shows the marker clustering and no-recombination region. **b** Linkage mapping in three F_2 populations confirmed the inversion in WI2757 but not TL. **c** Fosmid clone-based FISH among six cucumber inbred lines further validated the inversion in WI2757

To confirm the presence and the nature of the SV in WI2757, we conducted fluorescence in situ hybridization (FISH) on metaphase chromosomes in WI2757, Gy14 and 9930 cucumber lines. Three fosmid clones located in the ~ 10-Mbp target region (Yang et al. 2014; details in supplemental Table S8) were used as probes in FISH analysis. The results are shown in Fig. 3c. The physical locations of signals for the probes 255H13 (at ~ 24.6 Mbp on Gy14_v2.0), 255I22 (at ~ 27.4 Mbp), and 32O20 (at distal end of Chr1; ~ 32.4 Mbp) in 9930 and Gy14 chromosomes were consistent with their draft genome positions, but the order was reversed in WI2757. The centromere of Chr1 is located at approximately 12.0–13.0 Mbp region. This FISH work provided convincing evidence that a large paracentric inversion (~ 10 Mbp, from 23.4 Mbp to the end) is present in the long arm of Chr1 of WI2757 cucumber.

WI2757 has a complicated pedigree in which many lines were involved during its development including PI 220860 (Korea; source of gynoecy), PI 197087 (India; DM and anthracnose resistance source), and PI 212233 (Japan; PM resistance source) (Fig. S1; Peterson et al. 1982). To check if any of the donor lines may carry this inversion, we also included the three PI lines in FISH analysis. The three PI lines and the immediate donor Gy14 all had the same FISH signal pattern as 9930 (Fig. 3c). This suggests that the inversion in WI2757 may be generated de novo. It is also possible that this inversion was from some other donors in its pedigree, which are not available to test in this study (for example, EXPO or PM66, Fig. S1).

The m locus and the inversion affect QTL detection for fruit size, shape, and FT in the WI2757 × TL segregating populations

QTL analysis for FT, FL, FD, LD, FN, CN, and PM was performed using phenotypic data from WI2009F2, WI2014RIL, WI2015RIL, and WI2016RIL experiments. The linkage maps for the F_2 and RIL populations were the one developed by He et al. (2013), and the 1974-locus SNP-SSR map developed herein (Table S7), respectively. For each trait, the LOD threshold to declare significance of QTL was determined with 1000 permutation tests ($P = 0.05$), which ranged from 3.1 to 3.5. A global view of all QTL detected across the seven chromosomes for the WI2009F2 experiment is shown in supplemental Fig. S5. Those detected with the SSR/SNP map (87 RILs), and the 129-SSR map (139 RILs) are provided in Fig. 4 and supplemental Fig. S6, respectively. Details of all QTL for the morphological traits and FT from the four experiments including map locations, LOD support values, percentage of observed phenotypic variance explained (PVE, or R^2), additive effects, as well as 1.5-LOD support intervals are provided in Table 3. QTL information for PM resistance detected with the RIL population is listed separately in Table 4. Each QTL was assigned a name. If multiple QTL for the same trait detected by different experiments were located at the same or nearby locations, the same name was assigned (Weng et al. 2015; Pan et al. 2020). In total, 36 QTL were detected for the seven traits including 2 for FT, 6 for FL, 6 for FD, 5 for LD, 6 for FN, 2 for CN, and 9 for PM, which are briefly described below.

QTL for FT

Major-effect QTLs for FT were detected on Ch1 in all four experiments (PVE = 13.9–44.9%; Table 3, Fig. 4). This likely reflected the presence of a single QTL with consistent effect across environments, which thus was named *ft1.1*. A second FT QTL, *ft6.4* with moderate effect (PVE = 14.8%), was only detected in WI2009F2 (with CIM). Interestingly, *ft1.1* had positive additive effect (i.e., the WI2757 allele delays anthesis of female flowers), whereas the WI2757 allele of *ft6.4* promoted early flowering (negative additive effect) (Table 3). This could explain the earlier flowering time of F_1 of WI2757 × TL than either parent (Table 1; Fig. S2).

QTL of FL, FD, LD, FN, and CN

Six FL QTL, *fl1.1*, *fl1.2*, *fl3.1*, *fl4.1*, *fl4.2*, and *fl6.1*, were detected from the four experiments (Table 3). Two QTL, the minor-effect *fl1.1* (PVE = 6.1–7.6%) and the major-effect *fl1.2* (PVE = 34.3–68.7%), were detected in WI2014RIL, WI2015RIL, and WI2016RIL experiments. (*fl1.2* was also detected in WI2009F2.) Each of the four remaining minor-effect QTL, *fl3.1*, *fl4.1*, *fl4.2*, and *fl6.1* (PVE = 2.5–5.8%), was only detected in one experiment. Four of the six QTL (*fl1.1*, *fl1.2*, *fl3.1*, and *fl6.1*) had positive additive effects

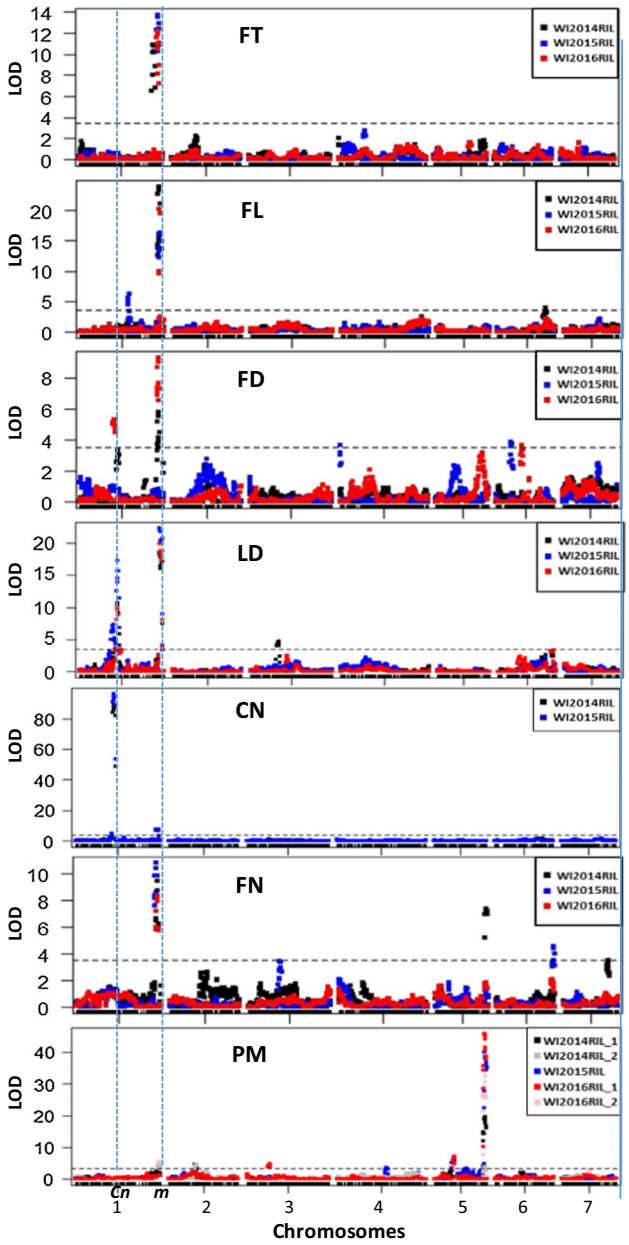

Fig. 4 Genome-wide view of QTL locations for flowering time (FT), mature fruit length (FL) and diameter (FD), mature fruit LD, carpel number (CN), fruit number (FN), and powdery mildew (PM) resistance detected with the RIL population in three experiments (WI2014RIL, WI2015RIL and WI2016RIL) and the 1974_SSR/SNP_locus genetic map. Horizontal dashed lines indicate LOD thresholds for significant QTL. Vertical dashed lines indicate QTL co-localized with the *carpel number* (*cn*), and *andromonoecious* (*m*) loci on Chr1

Table 3 QTL for flowering time and fruit-related traits detected with WI2757 × TL F_2 and RIL populations in four experiments

Traits	QTL	QTL Model	Chr	LOD	PVE %	Additive Effect[a]	Peak		1.5-LOD support interval			
							Marker	Position[b]	Left marker	Position[b]	Right marker	Position[b]
Flowering time (FT)												
WI2009F2	*ft1.1*	CIM	1	4.6	13.9	4.0	SSR04805	20,481,399	SSR00231	22,234,860	SSR04805	20,481,399
WI2009F2	*ft6.3*	CIM	6	5.2	14.8	−6.3	SSR15482	27,270,372	SSR21885	24,261,658	SSR17408	31,132,697
WI2014RIL	*ft1.1*	CIM, MQM	1	10.9	40.8	4.6	Marker1_18664549	22,175,676	Marker1_18295129	21,802,066	Marker1_19918318	23,483,836
WI2015RIL	*ft1.1*	CIM, MQM	1	13.8	44.9	4.7	Marker1_23086761	26,731,655	Marker1_19303313	22,836,447	Marker1_23190081	26,835,218
WI2016RIL	*ft1.1*	CIM, MQM	1	12.2	40.2	4.3	Marker1_19939037	23,471,205	Marker1_19303313	22,836,447	Marker1_23190081	26,835,218
Fruit length (FL)												
WI2014RIL	*fl1.1*	CIM, MQM	1	5.0	7.6	1.0	UW084377	16,711,449	UW083720	14,349,167	UW050055	18,112,141
WI2015RIL	*fl1.1*	CIM, MQM	1	6.3	6.1	0.9	012F_926781	17,293,627	Marker1_13517889	16,651,910	012F_1024812	17,391,728
WI2016RIL	*fl1.1*	CIM, MQM	1	19.4	6.5	1.0	UW050055	31,422,515	UW083720	14,349,167	SSR11340	18,577,265
WI2009F2	*fl1.2*	CIM, MQM	1	22.1	59.9	2.0	UW084538	31,422,515	UW084366	23,877,072	SSR21747	24,873,748
WI2014RIL	*fl1.2*	CIM, MQM	1	24.0	66.0	2.7	Marker1_19918318	23,483,836	Marker1_19303313	22,836,447	Marker1_19939037	23,471,205
WI2015RIL	*fl1.2*	CIM, MQM	1	16.3	34.3	2.1	Marker1_23166425	26,811,505	Marker1_19918318	23,483,836	Marker1_26995587	30,743,710
WI2016RIL	*fl1.2*	CIM, MQM	1	20.1	68.7	3.0	Marker1_19918318	23,483,836	Marker1_19303313	22,836,447	Marker1_23166425	26,811,505
WI2009F2	*fl3.1*	MQM	3	3.4	2.5	0.7	c3.loc70	n/a	SSR03918	10,317,233	SSR14649	24,422,117
WI2009F2	*fl4.1*	CIM	4	4.2	4.1	−0.6	c4.loc46	n/a	UW084382	4,715,516	UW084454	5,390,140
WI2009F2	*fl4.2*	MQM	4	4.2	5.3	−0.7	c4.loc64	n/a	SSR06225	10,211,019	SSR15420	19,981,772
WI2014RIL	*fl6.1*	CIM	6	3.9	5.8	0.8	Marker6_25501791	28,829,673	SSR15516	27,799,589	Marker6_26690309	30,010,544
Fruit diameter (FD)												
WI2009F2	*fd1.1*	CIM, MQM	1	4.9	16.6	−0.4	UW084540	8,915,615	UW084475	8,205,816	UW029623	9,073,308
WI2014RIL	*fd1.1*	CIM	1	3.3	10.6	−0.3	034F_323106	13,862,514	Marker1_11948097	14,819,392	Marker1_12787829	15,595,850
WI2015RIL	*fd1.1*	CIM, MQM	1	3.1	10.3	−0.2	SSR01816	10,462,162	UW084540	8,915,615	UW083728	24,083,996
WI2016RIL	*fd1.1*	CIM, MQM	1	5.3	12.3	−0.3	Marker1_12242079	13,651,158	SSR01816	10,462,162	Marker1_12350750	13,761,659
WI2009F2	*fd1.2*	CIM, MQM	1	4.9	15.8	0.3	SSR13830	28,065,667	UW083948	26,739,928	SSR14991	22,365,734
WI2014RIL	*fd1.2*	CIM, MQM	1	5.8	24.2	0.5	Marker1_19918318	23,483,836	Marker1_19303313	22,836,447	Marker1_23166425	26,811,505
WI2015RIL	*fd1.2*	CIM, MQM	1	5.3	15.2	0.3	UW084538	31,422,515	SSR05817	21,339,881	UW084538	31,422,515
WI2016RIL	*fd1.2*	CIM, MQM	1	9.3	27.4	0.4	Marker1_19918318	23,483,836	Marker1_19259121	22,791,715	Marker1_23086761	26,731,655
WI2015RIL	*fd4.1*	CIM	4	3.7	13.3	0.3	UW084071	237,807	Marker4_160277	173,550	UW084234	289,358
WI2009F2	*fd5.1*	MQM	5	3.0	7.3	−0.2	c5.loc92	n/a	SSR07531	22,838,798	UW084533	29,087,064
WI2016RIL	*fd5.1*	CIM, MQM	5	3.2	7.5	−0.2	Marker5_22084017	27,761,877	Marker5_19022355	24,657,020	Marker5_23736665	29,407,936
WI2015RIL	*fd6.1*	CIM	6	3.9	13.5	−0.3	SSR01331	10,658,631	Marker6_9509058	9,607,568	Marker6_11328269	11,483,687
WI2016RIL	*fd6.2*	CIM, MQM	6	3.7	11.6	−0.3	Marker6_17243061	20,199,470	Marker6_16798128	19,739,575	UW039897	20,587,347
Fruit shape (LD)												
WI2009F2	*ld1.1*	CIM, MQM	1	5.7	9.9	0.2	UW029623	9,073,308	UW029643	9,042,740	SSR01816	10,462,162
WI2014RIL	*ld1.1*	CIM, MQM	1	10.7	9.9	0.1	c1.loc118	n/a	Marker1_12350750	13,761,659	Marker1_12787829	15,595,850
WI2015RIL	*ld1.1*	CIM, MQM	1	10.2	19.9	0.2	X034F_323106	13,862,514	Marker1_12350750	13,761,659	Marker1_12787829	15,595,850

Table 3 (continued)

Traits	QTL	QTL Model	Chr	LOD	PVE %	Additive Effect[a]	Peak		1.5-LOD support interval			
							Marker	Position[b]	Left marker	Position[b]	Right marker	Position[b]
WI2016RIL	*ld1.1*	CIM, MQM	1	9.4	9.5	0.1	UW083720	14,349,386	Marker1_12350750	13,761,659	Marker1_12787829	15,595,850
WI2009F2	*ld1.2*	CIM, MQM	1	18.9	46.1	0.2	UW084538	31,422,515	UW083948	26,739,928	SSR21747	24,873,748
WI2014RIL	*ld1.2*	CIM, MQM	1	13.9	26.0	0.2	Marker1_23166425	26,811,505	Marker1_19918318	23,483,836	Marker1_23190081	26,835,218
WI2015RIL	*ld1.2*	CIM, MQM	1	18.5	24.4	0.2	Marker1_26830768	30,581,529	Marker1_19918318	23,483,836	Marker1_26995587	30,743,710
WI2016RIL	*ld1.2*	CIM, MQM	1	16.1	34.8	0.2	Marker1_23166425	26,811,505	Marker1_23086761	26,731,655	Marker1_23190081	26,835,218
WI2014RIL	*ld3.1*	CIM, MQM	3	3.4	4.5	0.1	Marker3_20651272	21,521,831	Marker3_11823594	11,855,957	Marker3_23721966	25,479,141
WI2015RIL	*ld3.1*	MQM	3	2.9	3.0	0.1	SSR14649	24,422,117	SSR00525	16,912,717	SSR10357	26,651,364
WI2016RIL	*ld3.1*	CIM, MQM	3	2.6	5.8	0.1	c3.loc112	n/a	Marker3_19064177	19,846,560	SSR14649	24,422,117
WI2014RIL	*ld6.1*	MQM	6	2.4	1.8	0.1	SSR15492	18,317,278	SSR01331	10,658,631	UW083909	18,920,412
WI2015RIL	*ld6.1*	MQM	6	2.8	2.7	0.1	UW083909	18,920,412	SSR17023	13,883,910	SSR01566	21,115,057
WI2016RIL	*ld6.1*	MQM	6	4.1	4.3	0.1	UW083909	18,920,412	SSR17023	13,883,910	UW039897	20,587,347
WI2014RIL	*ld6.2*	CIM, MQM	6	3.3	5.0	0.1	Marker6_24000756	27,269,854	Marker6_23191407	26,467,480	Marker6_24784229	28,111,525
WI2015RIL	*ld6.2*	CIM, MQM	6	3.1	4.0	0.1	X005F_6240376	24,050,505	X005F_5100842	22,910,439	Marker6_24779474	28,106,787
WI2016RIL	*ld6.2*	CIM, MQM	6	3.3	7.3	0.1	Marker6_25501791	28,829,673	Marker6_19048849	22,076,882	Marker6_28790991	32,100,317
Fruit number (FN)												
WI2014RIL	*fn1.1*	CIM, MQM	1	9.9	27.9	−3.2	Marker1_23166425	26,811,505	Marker1_23086761	26,731,655	Marker1_23190081	26,835,218
WI2015RIL	*fn1.1*	CIM, MQM	1	10.9	35.5	−3.2	Marker1_19918318	23,483,836	Marker1_19303313	22,836,447	Marker1_23166425	26,811,505
WI2016RIL	*fn1.1*	CIM, MQM	1	8.3	28.2	−3.9	Marker1_23166425	26,811,505	Marker1_19918318	23,483,836	Marker1_23190081	26,835,218
WI2015RIL	*fn3.1*	CIM	3	3.5	9.2	−1.7	Marker3_16904206	17,347,174	Marker3_16186669	16,653,341	Marker3_17244763	17,707,567
WI2014RIL	*fn4.1*	CIM, MQM	4	3.6	8.6	−1.8	UW024693	4,659,682	SSR05783	3,573,774	UW084454	5,390,140
WI2014RIL	*fn5.1*	CIM, MQM	5	7.4	19.7	2.7	Marker5_23736665	29,407,936	Marker5_23426207	29,098,943	Marker5_27562369	33,285,710
WI2015RIL	*fn5.1*	CIM, MQM	5	7.2	14.7	2.0	UW060738	29,364,751	UW013256	27,157,184	SSR10224	32,140,739
WI2015RIL	*fn6.1*	CIM, MQM	6	4.5	12.4	−1.9	Marker6_28574238	31,888,053	Marker6_28257773	31,571,745	Marker6_28790991	32,100,317
WI2014RIL	*fn7.1*	CIM, MQM	7	3.5	9.2	−1.8	SSR18133	19,662,205	003F_8413128	18,746,418	Marker7_15863145	20,293,106
Carpel number (CN)												
WI2009F2	*cn1.1*	CIM, MQM	1	20.6	57.0	−0.4	c1.loc62	n/a	SSR01816	10,462,162	UW084542	16,701,880
WI2014RIL	*cn1.1*	CIM, MQM	1	36.2	57.6	−0.6	c1.loc118	n/a	Marker1_12242079	13,651,158	Marker1_12787829	15,595,850
WI2015RIL	*cn1.1*	CIM, MQM	1	35.6	57.9	−0.6	c1.loc118	n/a	Marker1_12242079	13,651,158	Marker1_12787829	15,595,850
WI2014RIL	*cn1.2*	CIM, MQM	1	6.9	6.7	−0.2	Marker1_19018252	22,544,635	Marker1_18664549	22,175,676	Marker1_19918318	23,483,836
WI2015RIL	*cn1.2*	CIM, MQM	1	6.4	5.9	−0.2	Marker1_19018252	22,544,854	Marker1_18664549	22,175,676	Marker1_19918318	23,483,836

CIM composite interval mapping, *MQM* multiple QTL mapping, *Chr* chromosome, *LOD* logarithm of odds, *PVE* phenotypic variation explained by QTL

[a]The positive or negative additive effect indicates that the WI2757 allele increases or decreases means of each trait, respectively

[b]On Gy14 V2.0 draft genome assembly

Table 4 QTL for powdery mildew (PM) resistance detected with 139 WI2757 × TL RILs in three environments

Environments	QTL	QTL model	Chr	LOD	PVE %	Additive effect[a]	Peak		1.5-LOD support interval			
							Marker	Position[b]	Left marker	Position[b]	Right marker	Position[b]
WI2014RIL_1	*pm1.2**	MQM	1	2.6	2.5	−0.4	c1.loc252	n/a	Marker1_18936894	22,458,316	Marker1_29106871	33,265,916
WI2014RIL_2	*pm1.2**	CIM, MQM	1	5.1	15.8	−0.9	Marker1_18664549	22,175,676	Marker1_18295129	21,802,066	Marker1_19018252	22,544,635
WI2015RIL	*pm1.2**	CIM	1	3.8	2.8	−0.4	UW084366	23,877,072	SSR05817	21,339,881	UW084538	31,422,515
WI2016RIL_2	*pm1.2**	CIM, MQM	1	4.3	8.7	−0.5	Marker1_19939037	23,471,205	Marker1_19303313	22,836,447	Marker1_23190081	26,835,218
WI2014RIL_1	*pm2.2*	CIM, MQM	2	5.3	12.6	0.8	c2.loc86	n/a	Marker2_10265237	10,425,405	X001F_11952227	11,958,572
WI2014RIL_2	*pm2.2*	CIM, MQM	2	4.6	15.0	0.9	c2.loc88	n/a	Marker2_10191180	10,346,063	X001F_12835012	12,841,898
WI2016RIL_2	*pm2.2*	CIM, MQM	2	3.9	6.4	0.5	Marker2_10890264	11,062,750	Marker2_8488189	8,605,387	SSR12083	16,020,788
WI2014RIL_1	*pm3.2**	MQM	3	3.3	2.5	−0.4	SSR06031	34,767,643	UW084839	33,536,014	SSR10370	39,544,062
WI2016RIL_1	*pm3.3*	CIM, MQM	3	4.4	7.5	−0.6	Marker3_11723798	11,755,591	Marker3_11714116	11,745,907	X016F_1446755	14,410,086
WI2014RIL_2	*pm4.2**	MQM	4	3.9	9.9	0.7	SSR15420	19,981,772	SSR11043	17,993,400	UW084381	19,945,019
WI2015RIL	*pm5.1*	MQM	5	6.2	8.6	−0.6	c5.loc60	n/a	Marker5_12313140	8,627,310	Marker5_2111575	12,870,842
WI2016RIL_1	*pm5.1*	CIM, MQM	5	5.4	8.4	−0.6	Marker5_1143496	11,848,659	UW084597	12,172,875	Marker5_12878807	17,268,409
WI2016RIL_2	*pm5.1*	CIM, MQM	5	3.6	7.5	−0.5	Marker5_13042774	17,421,976	Marker5_2111575	12,870,842	Marker5_14250394	19,674,782
WI2014RIL_1	*pm5.2**	CIM, MQM	5	2.5	5.1	−0.5	UW084570	21,307,329	UW084597	12,172,875	SSR10725	22,531,563
WI2015RIL	*pm5.2**	CIM, MQM	5	6.2	12.6	−0.7	UW084570	21,307,329	Marker5_15286281	20,750,167	Marker5_16371112	21,934,854
WI2016RIL_1	*pm5.2**	CIM, MQM	5	7.3	6.2	−0.5	UW084844	22,786,622	UW084570	21,307,329	UW083851	23,511,903
WI2016RIL_2	*pm5.2**	CIM, MQM	5	3.8	4.9	−0.4	UW084844	22,786,622	UW084570	21,307,329	UW083851	23,511,903
WI2014RIL_1	*pm5.3**	CIM, MQM	5	13.2	41.6	−1.4	UW084553	30,194,516	Marker5_23426207	29,098,943	Marker5_26085430	31,791,223
WI2014RIL_2	*pm5.3**	CIM, MQM	5	3.6	11.3	−0.7	c5.loc156	n/a	X027F_10815205	29,098,039	Marker5_27562369	33,285,710
WI2015RIL	*pm5.3**	CIM, MQM	5	21.7	49.9	−1.4	UW084553	30,194,516	Marker5_23736665	29,407,936	Marker5_26179502	31,885,328
WI2016RIL_1	*pm5.3**	CIM, MQM	5	21.9	59.7	−1.6	Marker5_23736665	29,407,936	Marker5_23426207	29,098,943	Marker5_26085430	31,791,223
WI2016RIL_2	*pm5.3**	CIM, MQM	5	16.7	52.4	−1.4	c5.loc156	n/a	Marker5_23426207	29,098,943	Marker5_26085430	31,791,223
WI2014RIL_1	*pm6.1*	CIM, MQM	6	3.1	7.1	0.6	UW083909	18,920,412	SSR17023	13,883,910	UW039897	20,587,347
WI2015RIL	*pm6.1*	MQM	6	2.9	1.6	0.3	UW039897	20,587,347	SSR15492	18,317,278	SSR01566	21,115,057

QTL with asterisks are also detected in He et al. (2013)

CIM composite interval mapping, *MQM* multiple QTL mapping, *Chr* chromosome, *LOD* logarithm of odds, *PVE* phenotypic variation explained by QTL. WI2014RIL_1, WI2014RIL_2, WI2016RIL_1, and WI2016_2 represent the first and second scoring of PM resistance in WI2014RIL and WI2016RIL experiments, respectively

[a]The positive or negative additive effect indicates that the WI2757 allele increases or decreases PM resistance, respectively

[b]On Gy14v2.0 draft genome assembly

suggesting alleles from WI2757 at these loci promoting fruit elongation.

Among the six FD QTL, *fd1.1* (PVE = 10.3–16.6%) and *fd1.2* (PVE = 15.2–27.4%) were detected in all four environments. The QTL *fd4.1* (PVE = 13.3%), *fd6.1* (PVE = 13.5%), and *fd6.2* (PVE = 11.6%) were only identified in one experiment (WI2015RIL or WI2016RIL), whereas *fd5.1* (R^2 = 7.3–7.5%) was detected in both WI2009F2 and WI2016RIL. Both *fd1.2* and *fd4.1* showed positive additive effect on FD, while the rest had negative additive effect on FD (that is, WI2757 alleles reduce radial growth).

Of the five fruit shape (LD) QTL, the two major-effect ones, *ld1.1* (PVE = 9.5–19.9%) and *ld1.2* (PVE = 24.4–46.1%) were detected in both F_2 and the RIL populations (Table 3).The three minor-effect QTL, *ld3.1* (PVE = 3.0–5.8%), *ld6.1* (PVE = 1.8–4.3%), and *ld6.2* (PVE = 4.0–7.3%) showed up only in WI2014RIL, WI2015RIL, and WI2016RIL experiments, respectively. All QTL had positive additive effects on LD suggesting alleles from TL at these loci contributed to a rounder fruit shape.

Phenotype data for FN were only collected from the RIL population in WI2014RIL, WI2015RIL, and WI2016RIL. Among the six FN QTL detected, *fn1.1* (PVE = 27.9–35.5%) was consistently identified in all three experiments; *fn5.1* (PVE = 14.7–19.7%) was detectable in WI2014RIL and WI2015RIL experiment, whereas *fn3.1* (PVE = 9.2%), *fn4.1* (PVE = 8.6%), *fn6.1* (PVE = 12.4%), and *fn7.1* (PVE = 9.2%) were each detected in only one environment. TL alleles at all loci except for *fn5.1* contributed to increase of FN (negative additive effect for WI2757).

One major-effect QTL *cn1.1* (PVE = 57.0–57.9%) and one minor-effect QTL *cn1.2* (PVE = 5.9–6.7%) were identified with phenotypic data collected from WI2009F2, WI2014RIL, and WI2015RIL experiments. Both *cn1.1* and *cn1.2* had negative additive effect on carpel number, suggesting that the WI2757 alleles reduce CN.

QTL for PM resistance

Phenotypic data of PM disease scores upon natural infection were collected once for WI2015RIL and twice in WI2014RIL and WI2016RIL with a 2-week interval. Genome-wide PM QTL scan was conducted with data from individual ratings, and the results are summarized in Table 4. QTL naming followed Wang et al. (2020). In total, nine PM QTL were detected. Except for *pm3.2* (PVE = 2.5%), *pm3.3* (PVE = 7.5%), and *pm4.2* (PVE = 9.9%) that were detected in a single experiment, the rest six were detected in at least two environments including *pm1.2* (PVE = 2.8–15.8%), *pm2.2* (PVE = 6.4–15.0%), *pm5.1* (PVE = 7.5–8.6%), *pm5.2* (PVE= 4.9–12.6%), *pm5.3* (PVE = 11.3–59.7%), and *pm6.1* (PVE = 1.6–7.1%). Among the nine QTL, *pm5.3* was detected in all experiments and rating times with the strongest effect on conferring PM resistance. Three QTL, *pm2.2*, *pm4.2,* and *pm6.1*, showed positive additive effect, indicating that WI2757 alleles of these QTL increase disease scores. The rest six had negative additive effect. (WI2757 alleles reduce disease scores.)

Clustering of QTL at the andromonoecious (*m*) and carpel number (*cn*) loci

A glimpse of the genome-wide locations of all QTL detected in the present study showed that major-effect QTL for fruit morphology (FL, FD, LD, and CN) and flowering time (FT) were clustered on three regions on Chr1 (Fig. 4, Figs. S5 and S6). A close look of the 1.5-LOD intervals of these QTL suggested that the FL, FD, LD, FN, and FT QTL are co-localized with the *andromonoecy* (*m*) locus, and FD QTL was co-localized with the *cn* locus (*cn1.1*) (Fig. 4; Table 3). The *m* locus was inside the 10-Mbp inversion detected in WI2757. These observations suggested that at least some of those QTL are due to pleiotropic effects of the *m* or *cn* loci, or due to recombination suppression of the large paracentric inversion (see "Discussion" section). However, PM resistance did not seem to have any correlation with either locus.

Discussion

Phenotypic variation of morphological traits in the WI2757 × TL populations due to pleiotropic effects of the *m* and *cn* loci

In this study, we phenotyped flowering time (FT), fruit size, shape and number (FL, FD, LD, CN, and FN), as well as PM resistance in six environments in the WI2757 × TL F_2 and RIL populations. Trait correlations and QTL analysis clearly suggested that, except for the true QTL underlying these traits, additional factors contributed to the observed phenotypic variation in these populations including the pleiotropic effects of the *m* and *cn* loci and the large chromosomal inversion in WI2757 (Fig. 4; Table 3).

TL used in the present study exhibits andromonoecious sex expression (*mm*), and its fruit has five carpels (*cncn*) (Fig. 1). In cucurbits, both loci have pleiotropic effects on fruit size and shape (reviewed in Pan et al. 2020). Indeed, in segregation populations, andromonoecious plants had significantly more fruits (larger FN) that were shorter/rounder (smaller FL and LD), and with higher carpel numbers than gynoecious or monoecious plants (Fig. 2; Table S3). Consistent with this, QTL mapping identified QTL for FL, FD, LD, and FN that are co-localized with the *m* locus (Fig. 4).

At the whole population level, the mean FD values of andromonoecious plants were slightly lower than those in gynoecious and monoecious ones (Fig. 2; Table S3), which

probably suggests that the *m* locus restricts radial growth. On the other hand, andromonoecious plants tended to have slightly higher CN which is likely due to the weak linkage of the *m* and *cn* loci which are ~ 15 Mbp away on Chr1 (Fig. 4; Table S5). However, when only non-andromonoecious plants were considered, fruits with more carpels tend to have slightly lower FL, but higher FD (Table S4). This suggests that, in the absence of the effects of the *m* locus, the *cn* locus seems to promote radial fruit growth, which is consistent with observations in natural cucumber populations (Li et al. 2016), and other crops such as melon, and tomato (e.g., Perin et al. 2002; Monforte et al. 2004; Eduardo et al. 2007; Barrero and Tanksley 2004; Muños et al. 2011; Rodriguez et al. 2011). In this study, the CN QTL was mapped to a region where the candidate gene (*CsCLV3*) of *cn* locus is located (Li et al. 2016), and the intermediate-effect QTL for FD (*fd1.1*) was co-localized with CN QTL (and the *cn* locus) (Fig. 4) supporting the pleiotropic effect of CN on FD.

It should be pointed out that the magnitude of the observed phenotypic correlations was dependent on environments. In addition to the *m* and *cn* loci, the large inversion in WI2757, and the population structure (F_2 vs. RIL), environmental factors (field vs. greenhouse, and culture practices) were also some other factors that may affect the performance of these traits.

The 10-Mbp segmental inversion in WI2757

In this study, comparison of the SSR- and SNP-based linkage maps with the physical map identified a region of strong recombination suppression on the long arm of cucumber Chr1 in the WI2757 × TL F_2 and RIL populations (Tables S5 and S7). Linkage analysis in additional segregating populations and FISH revealed a large paracentric inversion in WI2757, which spanned ~ 10 Mbp (on Gy14_v2.0, from Marker1_19918318 to the end of Chr1; Fig. 4; Table S3). Inversions are the most common structural variations widely present in plant and animal genomes (e.g., Kirkpatrick 2010; Wellenreuther and Bernatchez 2018). In cucumber, Yang et al. (2012) identified six large inversions on three chromosomes (4, 5, and 7) between the cultivated (*C. s.* var. *sativus*) and wild (*C. s.* var. *hardwickii*) cucumbers. The ~ 10-Mbp inversion in WI2757 identified in this study seems to be the first report of large inversions within cultivated cucumbers.

A major consequence of large heterozygous inversions is recombination suppression in the inverted region resulting in transmission of genes within the inversion as a whole called 'supergene' (Thompson and Jiggins 2014). A supergene often consists of multiple coadapted loci associated with speciation, local adaptation, or fitness of plants (Dobzhansky 1947; Kirkpatrick and Barton 2006; Kirkpatrick 2010; Schwander et al. 2014; Wellenreuther and Bernatchez 2018). For example, in the yellow monkey flower (*Mimulus guttatus*), an inversion is associated with flowering time and other morphological traits in annual and perennial ecotypes (Lowry and Willis 2010; Lee et al. 2016). In maize (*Zea mays*), an inversion on Chr1 shows a strong altitudinal signature (Fang et al. 2012), and in *Arabidopsis thaliana*, a 1.17-Mbp inversion on Chr4 shows a strong association with fecundity under drought stress (Fransz et al. 2016). However, considering the very recent history of WI2757, any claim of adaptive significance of this inversion in WI2757 would be dubious.

The origin and the roles of the large paracentric inversion in WI2757 are unknown. WI2757 was released nearly 40 years ago, and five lines were involved in the development including Gy14, PM66, RS, 817, and EXPO (Fig. S1). FISH excluded three PI lines as possible donor of the inversion (Fig. 3c). Gy14 is the immediate donor of most disease resistances in WI2757. Gy14 flowers earlier than WI2757. From the draft genomes of Gy14 and 9930, no major structural changes have been observed in the inversion region. Thus, it is unlikely that Gy14 carries this inversion. As shown in Table S6, the average polymorphic level of the 6491 SLAF tags (SNPs) between WI2757 and TL was 12.9%, which was 26.3% for those on Chr1. We aligned WI2757 resequencing reads against the Gy14v2.0 draft genome for SNP calling. The genome-wide distribution of SNPs per 500Kbp window was plotted against physical length on each Chr (supplemental Fig. S7). The SNP polymorphisms in the 10-Mbp inversion region were more than twice of any other regions in the genome. This was probably due to the recombination suppression in this region and the high-level polymorphisms from the original donor of this region. This observation also hints that the inversion was not from Gy14, because, if this is true, we should see very low level of polymorphism in this region between WI2757 and Gy14. Conversely, we can also infer that regions with very low-level SNP polymorphism shown in Fig. S7 were probably derived from Gy14. Such regions are physically quite large, which seems to be consistent with the heavy presence of Gy14 in the pedigree of WI2757. One good example is the region harboring the cucumber *staygreen* gene (*CsSGR*) which is located ~ 5.0 Mbp position on Chr5 that is responsible for disease resistances against the downy mildew, angular leaf spot, and anthracnose pathogens, which has been shown to be derived from the original donor PI 197087 through Gy14 (Wang et al. 2019a). The low polymorphism level of SNPs between WI2757 and Gy14 in this region was in agreement of its Gy14 origin (Fig. S7).

The df (delayed flowering) locus in WI2757 and flowering time (FT) QTL in cucumber

Flowering time is an important trait for cucumber. Shifriss and George (1965) were probably the first to investigate the

inheritance of flowering time and seed dormancy in cucumber. The wild cucumber (*C. s.* var. *hardwickii*, HARD hereinafter) line 'Baroda' (PI 212896) exhibited strong seed dormancy and required short day for flowering. They found that the delayed flowering was conditioned by a single-recessive locus *df*, which seemed to be linked with seed dormancy that was controlled by probably three dominant genes. Della Vecchia et al. (1982, 1984) suggested that the day-length sensitive flowering in the HARD accession PI 215589 is also controlled by *df*. Fanourakis and Simon (1987) studied the delayed flowering habit in WI2757 and proposed a single-recessive gene *df* under this trait, which was linked with the femaleness (*F*) locus at ~ 34.7 cM. This linkage of *df* with *F* in WI2757 was concurred by Walters et al. (2001). In particular, in the F_2 populations from the cross between WI2757 and LJ90430 that is a selection from the wild cucumber line PI 183967, the segregation of early flowering and delayed flowering plants was consistent with 3:1 ratio, and the genetic distances between *df* and *F* or *de* (determinate) were 0–35.4 cM. This implies that *df* in WI2757 and PI 183967 was probably allelic and was located in a region of Chr6 that is very close to the *F/de* loci (*de* at 24.6 Mbp and *F* at ~ 27.6 Mbp of Chr6 on Gy14v2.0) (Table S7; Wen et al. 2019; Li et al. 2020).

In the present study, we detected two FT QTL: the major-effect *ft1.1* (PVE = 13.9–44.9%) and minor-effect *ft6.4* (PVE = 14.8%) (Table 4). While the exact region harboring *ft1.1* in WI2757 is not known due to recombination suppression inside the inversion, its location is largely consistent with major-effect FT QTL detected in several previous studies (e.g., Miao et al. 2012; Lv et al. 2014; Bo et al. 2015; Sheng et al. 2019; Wang et al. 2019b; reviewed by Wang et al. 2020). In particular, Sheng et al. (2019) conducted QTL mapping of flowering time using segregation populations derived from the cross between Gy14 and the monoecious, later flowering HARD accession PI 183967. Interestingly, Sheng et al. (2019) also identified two FT QTL, *FT1.1* (PVE = 16.2–42.8%) and *FT6.3* (PVE = 6.0–23.8%), which were located in the same 1.5-LOD intervals as *ft1.1* and *ft6.4* detected in this study, respectively. However, while the alleles of both *ft1.1* in PI 183967 and *ft1.1* in WI2757 contributed to late flowering, the allele effect of *ft6.3* in HARD and *ft6.4* in WI2757 had the opposite effect, which delayed and promoted flowering, respectively (Sheng et al. 2019; Table 4). The most possible location of *FT1.1* QTL is in an interval from 27.6 to 30.8 Mbp (Gy14v2.0, Sheng et al. 2019), which overlaps with the region of the 10-Mbp inversion (Fig. 3a), suggesting that *FT1.1* in HARD and *ft1.1* in WI2757 are closely linked. However, there are two important differences in the late flowering habit between WI2757 and PI 183967. First, as mentioned above, the effect of *ft6.3* and *ft6.4* on flowering in the two lines was the opposite. Second, flowering time in HARD is day length sensitive (short day promotes flowering) (Della Vecchia et al. 1982, 1984). It is not known if the flowering time in WI2757 is affected by photoperiod. Based on our observations, such day length requirement, if any, is not as strong as PI 183967 because WI2757 usually flowers earlier than PI 183967 in Wisconsin greenhouse or field conditions. There is also no indication that any of those lines presented in the pedigree of WI2757 (Fig. S1) are day length sensitive for flowering. Meanwhile, we cannot eliminate the possibility that the late flowering in WI2757 is caused by the paracentric inversion because large inversions may alter expressions of involved genes (e.g., Huang et al. 2018).

Considering that *ft1.1* on Chr1 in both HARD and WI2757 is the major-effect QTL for late flowering, it is puzzling that the *df* locus was mapped on Chr6 and linked with the *F* or *de* loci (Fanourakis and Simon 1987; Walters et al. 2001). This raises the question if the *delayed flowering* (*df*) loci in HARD and WI2757 are the same as the flowering time QTL detected herein. One possible explanation is the phenotyping method of 'delay flowering' used by Fanourakis and Simon (1987) and Walters et al. (2001), which was defined as plants that had no flowers in the first five nodes. This definition was similar to the 'first female node' (FFN) used in QTL mapping studies for this trait by Yuan et al. (2008), and Miao et al. (2012). The mapping populations used in the four aforementioned studies (Fanourakis and Simon 1987; Walters et al. 2001; Yuan et al. 2008; Miao et al. 2012) were from crosses between a gynoecious and a monoecious parental lines that are segregating at the *F* locus. The phenotyping method for FFN may be affected by the gynoecious sex expression controlled by the *F* locus. Among the nine FFN QTL detected by Yuan et al. (2008) and Miao et al. (2012), *ffn1.3*, *ffn3.2*, and *ffn6.2* had the largest effects on FFN (PVE > 20% each), which are located in the regions where *FT1.1*, the subgynoecious QTL *sg3.1*, and the *F* locus were placed, respectively (Wang et al. 2020). This may explain the linkage of the *df* locus with *F/de* loci on Chr6 in WI2757 and PI 183967 (Fanourakis and Simon 1987; Walters et al. 2001). In this context, the *df* locus first proposed by Shifriss and George (1965) to define the delayed flowering in the wild cucumber line Baroda is consistent with *ft1.1* detected in the present study and Sheng et al. (2019). An indirect evidence to support this is the linkage of 'delayed flowering' and seed dormancy in Baroda because a major QTL for seed dormancy in wild cucumber seems to be located on Chr1 (unpublished data).

QTL for fruit size/shape/number, and PM resistance in the WI2757 × TL populations

Nearly, 200 QTL for fruit size, shape, and weight have been identified in cucumber. Pan et al. (2020) reviewed the literature and proposed 19 consensus fruit size (FS), and 11 fruit shape (FSI or LD) QTL that could explain the majority of fruit size and shape variation observed so far in cucumber. In this study, we identified six FL QTL (*fl1.1*, *fl1.2*, *fl3.1*, *fl4.1*, *fl4.2*, and *fl6.1*), six FD QTL (*fd1.1*, *fd1.2*, *fd4.1*, *fd5.1*, *fd6.1*, and *fd6.2*), and five LD QTL (*ld1.1*, *ld1.2*, *ld3.1*, *ld6.1*, and *ld6.2*) (Table 4). Based on their chromosomal locations, the relationships of these QTL with previously detected consensus QTL are listed in supplemental Table S9. All these QTL seem to overlap with the consensus FS or FSI QTL proposed by Pan et al. (2020) except for *fd4.1*. Since *fd4.1* was detected only in one season (WI2015), additional work is needed to validate this QTL.

In this study, five major-effect QTL (*ft1.1*, *fl1.2*, *fd1.2*, *ld1.2*, and *fn1.1*) and two minor-effect QTL (*cn1.2*, and *pm1.1*) were mapped inside the inversion that also harbors the *M* locus (Tables 3, 4; Fig. 4). Previous studies have shown pleiotropic effects of the andromonoecious *m* locus on fruit size and shape (Pan et al. 2020). Thus, as discussed early, the four QTL, *fl1.2, fd1.2, ld1.2,*, and *fn1.1*, detected in the WI2757 × TL populations are likely due to pleiotropy of the *m* locus, which could be evidenced from the significant correlations observed in the segregating populations discussed early (Table 2). However, we cannot eliminate the possibility that true FL, FD, and LD QTL are located in this region, which are closely linked with the *m* locus. For example, the *CsSUN* is a candidate gene for the fruit size QTL *FS1.2*, which is only 200 kb away from the *m* locus (Pan et al. 2017a, 2020). Due to the recombination suppression, we will not be able to prove either possibility in the populations we used in this study.

Two previous studies investigated the genetic basis of fruit number (FN) in cucumber. Pan et al. (2017b) detected four QTL (*fn1.1*, *fn3.1*, *fn6.1*, and *fn7.1*) with the $F_{2:3}$ population derived from the cross between WI7200 and WI7167. Sheng et al. (2019) also identified four FN QTL (*fn1.1*, *fn2.1*, *fn4.1*, and *fn6.1*) used Gy14 × HARD (PI 183967) segregating populations, but none of these FN QTL from the two early studies physically overlap. In the present study, in additional to the major-effect QTL *fn1.1*, which is likely the pleiotropic effect of the *m* locus, we detected five additional FN QTL with intermediate or minor effects (*fn3.1*, *fn4.1*, *fn5.1*, *fn6.1*, and *fn7.1*) (Table 4). The TL alleles of all except *fn5.1* contributed to increase of FN, suggesting that the higher fruit number in TL could also be contributed by other genetic loci. Among the six FN QTL, only *fn7.1* seems to co-localize with *fn7.1* detected by Pan et al. (2017b) (Table S9). While additional work is needed to confirm/validate these QTL, these observations revealed a complex genetic basis of FN variation in this population, which is likely affected by many factors such as plant architecture, sex expression, fruit size/weight, as well as environmental conditions.

The genetic basis of PM resistance (PMR) at the seedling stage in WI2757 has been previously investigated in a WI2757 × TL $F_{2:3}$ population with artificial inoculation under controlled environments (greenhouses) (He et al. 2013). In the early study, six PMR QTL, *pm1.1*, *pm1.2*, *pm3.2*, *pm4.2*, *pm5.2*, and *pm5.3* (*pm-h*, for hypocotyl resistance) were identified with the WI2757 alleles of *pm3.2* and *pm4.2* contributing to susceptibility (He et al. 2013). In this study, QTL mapping of PMR was conducted on adult plants of RILs from the same cross under natural infection in open fields. Five of the six QTL shared between the two studies, suggesting that these QTL are effective on both seeding and adult-plant stages. Besides, more minor-effect QTL were detected in the RIL population, which might be due to the different environments or development stages of plants in these two studies. In both studies, *pm5.3* had the strongest effect for PM resistance, which was co-localized with the well-characterized loss-of-susceptibility *R* gene,and the *mlo* locus for PM resistance in cucumber (Berg et al. 2015; Nie et al. 2015a, b).

Wang et al. (2020) reviewed the literature in cucumber on QTL mapping studies and summarized 19 PM resistance QTL so far identified in cucumber. Among the nine PM QTL identified in the present study, the 1.5-LOD support interval of *pm3.3* did not overlap with any reported QTL (Table S9). However, *pm3.3* was only detectable in the first rating of WI2016RIL, and further work is needed to verify its existence.

Use of WI2757 in cucumber breeding: perspectives

The multiple disease-resistant WI2757 has been an important germplasm for cucumber breeding. In the present study, we identified a large paracentric inversion of ~ 10 Mbp in size in Chr1 in WI2757 (Fig. 3). No known single genes or major-effect QTL for disease resistances in WI2757 were mapped in this inversion region (Wang et al. 2020). Thus, if WI2757 is to be used as the donor for these disease resistances in cucumber breeding, this inversion should not constitute as a major obstacle. However, Iezzoni and Peterson (1980) found linkage between bacterial wilt (BW) resistance and andromonoecious sex expression (*m* locus) in cucumber. If the BW resistance in WI2757 is also located in this region, it may be difficult to use this resistance in WI2757. Overall, the late flowering and poor field performance of WI2757 remain a nuisance for its use. The later flowering of WI2757 is likely associated with the major-effect QTL *ft1.1* inside the large inversion. It is not known if the poor field performance of WI2757 is also

linked with this inversion or is due to a fitness cost associated with the multiple disease resistances. To overcome these shortcomings, one possible solution is to replace the long arm of Chr1 in WI2757 (thus the inversion) through marker-assisted selection.

Large inversions bear extensive, long-range linkage disequilibrium due to recombination suppression, which may influence genome-wide association analysis (e.g., Nordborg et al. 2002; Fang et al. 2012; Fransz et al. 2016). In the inversion region, there is a very high level of SNP polymorphisms between WI2757 and other cucumber lines (Fig. S7), which may influence on estimation of population structure and GWAS analysis if WI2757 is present in the association panel. Therefore, caution should be exercised in using the germplasms that have large chromosome inversions such as the wild HARD cucumbers (Yang et al. 2012) and WI2757 in GWAS analysis of cucumber.

Acknowledgements This research was supported by the Agriculture and Food Research Initiative Competitive Grants under award Numbers 2015-51181-24285 and 2017-67013-26195 from the US Department of Agriculture National Institute of Food and Agriculture (to Y. Weng). USDA is an equal opportunity provider and employer.

Author Contribution statement YP conducted majority of the reported research. CW genotyped the RIL population with SLAF-Seq. YHW conducted GBS data analysis and linkage analysis with SNP markers. YH performed FISH analysis. XC, SL, and YL participated in phenotypic data collection in different experiments. YW conceived and supervised the research and wrote the manuscript with YP.

Availability of data and materials All data pertinent to the reported work have been provided in the manuscript or in the supplemental online materials.

Compliance with ethical standards

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

References

Barrero LS, Tanksley SD (2004) Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars. Theor Appl Genet 109:669–679

Berg JA, Appiano M, Santillán Martínez M, Hermans FW, Vriezen WH, Visser RG, Bai Y, Schouten HJ (2015) A transposable element insertion in the susceptibility gene *CsaMLO8* results in hypocotyl resistance to powdery mildew in cucumber. BMC Plant Biol 15:243

Bo K, Ma Z, Chen J, 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

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

Chung S-M, Staub JE, Fazio G (2003) Inheritance of chilling injury: a maternally inherited trait in cucumber. J Am Soc Hortic Sci 128:526–530

Della Vecchia PT, Peterson CE, Staub JE (1982) Effect of the duration of short-day treatment on the flowering response of a *Cucumis sativus* var. *hardwickii* (R.) Alef. line. Cucurbit Genet Coop Rep 5:2–3

Della Vecchia PT, Peterson CE, Staub JE (1984) Inheritance of flowering responses in cucumber. J Am Soc Hortic Sci 109:761–763

Dobzhansky T (1947) Adaptive changes induced by natural selection in wild populations of *Drosophila*. Evol Int J Org Evol 1:1–16

Eduardo I, Arus P, Monforte AJ, Obando J, Fernandez-Trujillo JP, Martinez JA, Alarcon AL, Alvarez JM, van der Knaap E (2007) Estimating the genetic architecture of fruit quality traits in melon using a genomic library of near isogenic lines. J Am Soc Hortic Sci 132:80–89

Fang Z, Pyhajarvi T, Weber AL, Dawe RK, Glaubitz JC, de Gonzalez J, Ross-Ibarra C, Doebley J, Morrell PL, Ross-Ibarra J (2012) Megabase-scale inversion polymorphism in the wild ancestor of maize. Genetics 191:883–894

Fanourakis NE, Simon PW (1987) Analysis of genetic linkage in the cucumber. J Hered 78:238–242

Fransz P, Linc G, Lee CR, Aflitos SA et al (2016) Molecular, genetic and evolutionary analysis of a paracentric inversion in *Arabidopsis thaliana*. Plant J 88:159–178

He XM, Li YH, Pandey S, Yandell BS, Pathak M, Weng YQ (2013) QTL mapping of powdery mildew resistance in WI2757 cucumber (*Cucumis sativus* L.). Theor Appl Genet 126:2149–2161

Horejsi T, Staub JE, Thomas C (2000) Linkage of random amplified polymorphic DNA markers to downy mildew resistance in cucumber (*Cucumis sativus* L.). Euphytica 115:105–113

Huang Y-C, Dang VD, Chang N-C, Wang J (2018) Multiple large inversions and breakpoint rewiring of gene expression in the evolution of the fire ant social supergene. Proc R Soc B 285:20180221

Iezzoni AF, Peterson CE (1980) Linkage of bacterial wilt resistance and sex expression in cucumber. HortScience 15:257–258

Kang HX, Weng Y, Yang YH, Zhang ZH, Zhang SP et al (2010) Fine genetic mapping localizes cucumber scab resistance gene *Ccu* into an R gene cluster. Theor Appl Genet 122:795–803

Kennard WC, Poetter K, Dijkhuizen A, Meglic V, Staub JE, Havey MJ (1994) Linkages among RFLP, RAPD, isozyme, disease-resistance, and morphological markers in narrow and wide crosses of cucumber. Theor Appl Genet 89:42–48

Kirkpatrick M (2010) How and why chromosome inversions evolve? PLoS Biol 8:e1000501

Kirkpatrick M, Barton N (2006) Chromosome inversions, local adaptation and speciation. Genetics 173:419–434

Lee YW, Fishman L, Kelly JK, Willis JH (2016) A Segregating inversion generates fitness variation in yellow monkeyflower (*Mimulus guttatus*). Genetics 202:1473–1484

Li S, Pan YP, Wen CL, Li YH, Liu XF, Zhang XL, Behera TK, Xing GM, Weng YQ (2016) Integrated analysis in bi-parental and natural populations reveals *CsCLAVATA3* (*CsCLV3*) underlying carpel number variations in cucumber. Theor Appl Genet 129:1007–1022

Li Z, Han YH, Niu HH, Wang YH, Jiang B, Weng Y (2020) Gynoecy instability in cucumber (*Cucumis sativus* L.) is due to unequal crossover at the copy number variation-dependent *femaleness* (*F*) locus. Hortic Res 7:32. https://doi.org/10.1038/s41438-020-0251-2

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS Biol 8:e1000500

Lv H, Lin T, Klein J, Wang S, Zhou Q, Sun J, Weng Y, Huang S (2014) QTL-Seq identifies an early flowering QTL located near *Flowering Locus T* in cucumber. Theor Appl Genet 127:1491–1499

Mao AJ, Zhang F, Zhang LR, Wang YJ (2008) Analysis on the inheritance of resistance to *Fusarium* wilt race 4 and scab and their linkage in WI2757 cucumber. Sci Agric Sin 41:3382–3388

Miao H, Gu XF, Zhang SP, Zhang ZH, Huang SW, Wang Y, Fang ZY (2012) Mapping QTLs for seedling-associated traits in cucumber. Acta Hortic Sin 39:879–887

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

Muños S, Ranc N, Botton E et al (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near *WUSCHEL*. Plant Physiol 156:2244–2254

Nie J, He HL, Peng JL, Yang XQ, Bie BB, Zhao JL, Wang YL, Si LT, Pan JS, Cai R (2015a) Identification and fine mapping of *pm5.1*: a recessive gene for powdery mildew resistance in cucumber (*Cucumis sativus* L.). Mol Breed 35:7

Nie J, Wang YL, He HL, Guo CL, Zhu WY, Pan J, Li DD, Lian HL, Pan JS, Cai R (2015b) Loss-of-Function mutations in *CsMLO1* confer durable powdery mildew resistance in cucumber (*Cucumis sativus* L.). Front Plant Sci 6:1–14

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat Genet 30:190–193

Olczak-Woltman H, Masny A, Bartoszewski G, Płucienniczak A, Niemirowicz-Szczytt K (2007) Genetic diversity of *Pseudomonas syringae* pv. *lachrymans* strains isolated from cucumber leaves collected in Poland. Plant Pathol 56:373–382

Pan YP, Liang XJ, Gao ML, Liu HQ, Meng HW, Weng YQ, Cheng ZH (2017a) Round fruit shape in WI7239 cucumber is controlled by two interacting quantitative trait loci with one putatively encoding a tomato *SUN* homolog. Theor Appl Genet 130:573–586

Pan YP, Qu SP, Bo KL, Gao ML, Haider KR, Weng YQ (2017b) QTL mapping of domestication and diversifying selection related traits in round-fruited semi-wild Xishuangbanna cucumber (*Cucumis sativus* L. *var. xishuangbannanesis*). Theor Appl Genet 130:1531–1548

Pan JS, Tan JY, Wang YH, Zheng XY, Owens K, Li DW, Li YH, Weng Y (2018) *STAYGREEN* (*CsSGR*) is a candidate for the anthracnose (*Colletotrichum orbiculare*) resistance locus *cla* in Gy14 cucumber. Theor Appl Genet 131:1577–1587

Pan YP, Wang YH, McGregor C, Liu S, Luan FS, Gao ML, Weng Y (2020) Genetic architecture of fruit size and shape variation in cucurbits: a comparative perspective. Theor Appl Genet 133:1–21

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

Peterson CE, Williams PH, Palmer M, Louward P (1982) Wisconsin 2757 cucumber. HortScience 17:268

Ren Y, Zhang Z, Liu J, Staub JE, Han Y, Cheng Z, Li X, Lu J, Miao H, Kang H, Xie B, Gu X, Wang X, Du Y, Jin W, Huang S (2009) An integrated genetic and cytogenetic map of the cucumber genome. PLoS ONE 4:e5795

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

Schwander T, Libbrecht R, Keller L (2014) Supergenes and complex phenotypes. Curr Biol 24:R288–R294

Sheng YS, Pan YP, Li YH, Yang LM, Weng Y (2019) Quantitative trait loci for fruit size and flowering time-related traits under domestication and diversifying selection in cucumber (*Cucumis sativus* L.). Plant Breed. https://doi.org/10.1111/pbr.12754

Shifriss O, George WL Jr (1965) Delayed germination and flowering in cucumbers. Nature 206:424–425

Słomnicka R, Olczak-Woltman H, Korzeniewska A, Niemirowicz-Szczytt K, Bartoszewski G (2016) Evaluation of a cucumber RILs population for resistance to angular leaf spot. In: Kozik EU, Paris HS (eds) Proceeding of cucurbitaceae 2016, 11th EUCARPIA meeting on genetics and breeding of cucurbitaceae. Warsaw, Poland, pp 231–233

Sun XW, Liu DY, Zhang XF, Li WB, Liu H et al (2013) SLAF-Seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PLoS ONE 8:e58700

Tan J, Tao Q, Niu H, Zhang Z, Li D, Gong Z, Weng Y, Li Z (2015) A novel allele of *monoecious* (*m*) locus is responsible for elongated fruit shape and perfect flowers in cucumber (*Cucumis sativus* L.). Theor Appl Genet 128:2483–2493

Thomas A, Carbone I, Choe K, Quesada-Ocampo LM, Ojiambo PS (2017) Resurgence of cucurbit downy mildew in the United States: insights from comparative genomic analysis of *Pseudoperonospora cubensis*. Ecol Evol 7:6231–6246

Thompson MJ, Jiggins CD (2014) Supergenes and their role in evolution. Heredity 113:1–8

Vakalounakis DJ (1993) Inheritance and genetic linkage of Fusarium wilt (*Fusarium oxysporum f.sp. cucumerinum* race 1) and scab (*Cladosporiurn cucumerinum*) resistance genes in cucumber (*Cucumis sativus*). Ann Appl Biol 123:359–365

Vakalounakis DJ (1995) Inheritance and linkage of resistance in cucumber line SMR-18 to races 1 and 2 of *Fusarium oxysporum f.sp. cucumerinurn*. Plant Pathol 44:169–172

Walters A, Shetty N, Wehner TC (2001) Segregation and linkage of several genes in cucumber. J Am Soc Hortic Sci 126:442–450

Wang YH, VandenLangenberg K, Wehner TC, Kraan PA, Suelmann J, Zheng X, Owens K, Weng Y (2016) QTL mapping for downy mildew resistance in cucumber inbred line WI7120 (PI 330628). Theor Appl Genet 129:1493–1505

Wang YH, VandenLangenberg K, Wen CL, Wehner TC, Weng YQ (2018) QTL mapping of downy and powdery mildew resistances in PI 197088 cucumber with genotyping-by-sequencing in RIL population. Theor Appl Genet 131:597–611

Wang YH, Tan JY, Wu ZM, VandenLangenberg K, Wehner TC, Wen CL, Zheng XY, Owens K, Thornton A, Bang HH, Hoeft E, Kraan PAG, Suelmann J, Pan JS, Weng Y (2019a) STAYGREEN, STAY HEALTHY: a loss-of-susceptibility mutation in the *STAYGREEN* gene provides durable, broad-spectrum disease resistances for over 50 years of US cucumber production. New Phytol 221:415–430

Wang SH, Li HB, Li YY, Li Z, Qi JJ, Lin T, Yang XY, Zhang ZH, Hang SW (2019b) *FLOWERING LOCUS T* Improved cucumber adaptation to higher latitudes. Plant Physiol. https://doi.org/10.1104/pp.19.01215

Wang YH, Bo KL, Gu XF, Pan JS, Li YH, Chen JF, Wen CL, Ren ZH, Ren HZ, Chen XH, Grumet R, Weng Y (2020) Molecularly tagged genes and quantitative trait loci in cucumber—and recommendation of controlled vocabulary for QTL mapping. Hortic Res 7:3

Wehner TC, Robinson RW (1991) A brief history of the development of cucumber cultivars in the U. S. Cucurbit Genet. Coop Rpt 14:1

Wellenreuther M, Bernatchez L (2018) Eco-evolutionary genomics of chromosomal inversions. Trends Ecol Evol 33:427–440

Wen C, Mao A, Dong C, Liu H, Yu S, Guo Y, Weng Y, Xu Y (2015) Fine genetic mapping of target leaf spot resistance gene *cca-3* in cucumber, *Cucumis sativus* L. Theor Appl Genet 128:2495–2506

Wen CL, Zhao WS, Liu WL, Yang LM, Wang YH, Liu XW, Xu Y, Ren HZ, Guo YD, Li C, Li JG, Weng Y, Zhang XL (2019) CsTFL1 inhibits determinate growth through interaction with CsNOT2 and CsFDP in cucumber (*Cucumis sativus* L.). Development 146:dev180166. https://doi.org/10.1242/dev.180166

Weng YQ, Colle M, Wang YH, Yang LM, Rubinstein M, Sherman A, Ophir R, Grumet R (2015) QTL mapping in multiple populations and development stages reveals dynamic quantitative trait loci for

fruit size in cucumbers of different market classes. Theor Appl Genet 128:1747–1763

Wyszogrodzka A, Williams P, Peterson C (1987) Multiple-pathogen inoculation of cucumber (*Cucumis sativus*) seedlings. Plant Dis 71:275

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

Yang LM, Koo DH, Li DW, Zhang T, Jiang JM, Luan FS, Renner SS, Henaff E, Sanseverino W, Garcia-Mas J, Casacuberta J, Senalik DA, Simon PW, Chen JF, Weng YQ (2014) Next-generation sequencing, FISH mapping and synteny-based modeling reveal mechanisms of decreasing dysploidy in *Cucumis*. Plant J 77:16–30

Yuan XJ, Pan JS, Cai R, Guan Y, Liu LZ, Zhang WW, Li Z, He HL, Zhang C, Si LT, Zhu LH (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

Zhang S, Miao H, Gu XF, Yang YH, Xie BY et al (2010) Genetic mapping of the scab resistance gene in cucumber. J Am Soc Hortic Sci 135:53–58

Zhang S, Miao H, Yang YH, Xie BY, Wang Y, Gu X (2014) A major quantitative trait locus conferring resistance to Fusarium wilt was detected in cucumber by using recombinant inbred lines. Mol Breed 34:1805–1815