Theor Appl Genet (2015) 128:2483–2493
DOI 10.1007/s00122-015-2603-0

ORIGINAL ARTICLE

A novel allele of *monoecious* (*m*) locus is responsible for elongated fruit shape and perfect flowers in cucumber (*Cucumis sativus* L.)

Junyi Tan[1] · Qianyi Tao[1] · Huanhuan Niu[1] · Zhen Zhang[1] · Dandan Li[2] · Zhenhui Gong[1] · Yiqun Weng[3] · Zheng Li[1]

Received: 2 June 2015 / Accepted: 28 August 2015 / Published online: 8 September 2015

Abstract

Key message **A 14 bp deletion in *CsACS2* gene encoding a truncated loss-of-function protein is responsible for elongated fruit shape and perfect flowers in cucumber.**

Abstract In cucumber (*Cucumis sativus* L.), sex expression and fruit shape are important components of biological and marketable yield. The association of fruit shape and sex expression is a very interesting phenomenon. The sex determination is controlled primarily by the *F* (female) and *M* (monoecy) loci. Homozygous recessive *mm* plants bear bisexual (perfect) flowers, and the fruits are often round shaped. *CsACS2* encoding the 1-aminocyclopropane-1-carboxylic acid synthase has been shown to be the candidate gene for the *m* locus. We recently identified an andromonoecious cucumber line H38 that has bisexual flowers but elongated fruits. To rapidly clone this *monoecious* gene in H38, we developed a tri-parent mapping strategy, which took advantage of the high-density Gy14 × 9930 cucumber genetic map and the powder of bulk segregant analysis. Microsatellite markers from the Gy14 × 9930 map were used to screen two pairs of unisexual and bisexual bulks constructed from H38 × Gy14 and H38 × 9930 F_2 populations. Polymorphic markers were identified and used to quickly develop a framework map and place the *monoecious* locus of H38 in cucumber chromosome 1. Further fine mapping allowed identification of a novel allele, *m-1*, at the *monoecious* locus to control the bisexual flower in H38, which was due to a 14 bp deletion in the third exon of the *CsACS2* gene encoding a truncated loss-of-function protein of the cucumber 1-aminocyclopropane-1-carboxylic acid synthase. This new allele provides a valuable tool in understanding the molecular mechanisms of *CsACS2* in the relationships of sex determination, fruit shape, and CsACS activities in cucumber.

Introduction

Cucumber (*Cucumis sativus* L.) has long been served as a model species for the study of plant sex determination (Galun 1961; Kubicki 1969). Three types of flowers can be present in a cucumber plant: staminate (male), pistillate (female), and hermaphrodite (bisexual/perfect). The *F/f* and *M/m* loci play the major role in sex expression in cucumber. The *F* (femaleness) is a partially dominant gene that controls femaleness, while the *M* locus determines the monoecy of cucumber plants: the dominant allele *M* will only allow the formation of stamen less female flowers, as well as male flowers, while in homozygous recessive *mm* plants, bisexual (perfect) flowers form, in addition to male flowers. An elusive third gene, *a* (androecious) increases maleness. Thus, a cucumber plant may be monoecious (*MMffAA*, with both male and female flowers), gynoecious

Communicated by S. Huang.

J. Tan and Q. Tao have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2603-0) contains supplementary material, which is available to authorized users.

✉ Zheng Li
lizheng82@nwsuaf.edu.cn

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

[2] Agricultural College, Heilongjiang Bayi Agriculture University, Daqing 163319, Heilongjiang, China

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

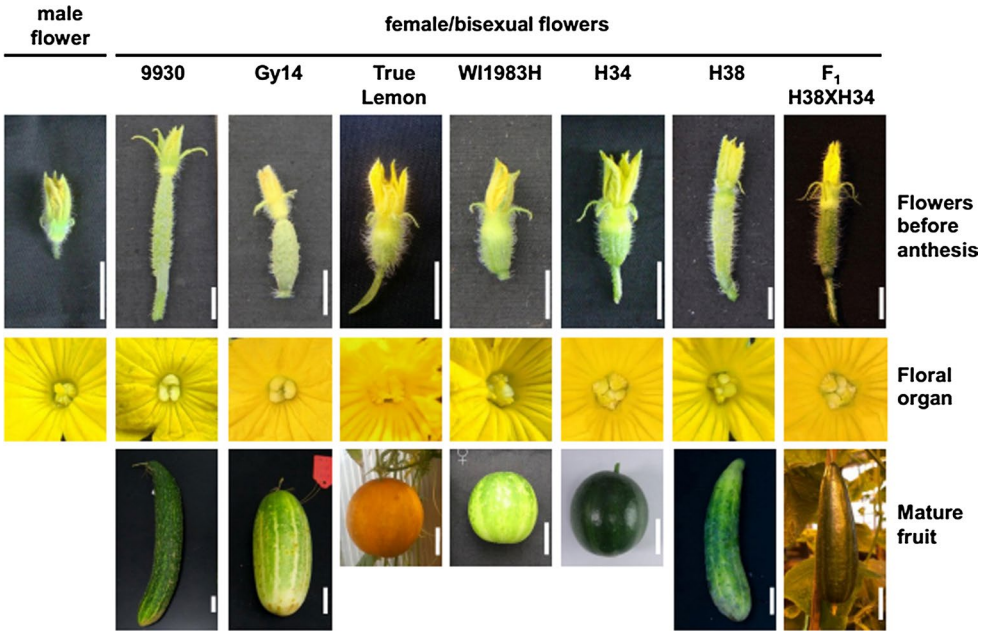

Fig. 1 Sex expression, flower, and fruit morphology of six cucumber inbred lines and an F_1 plant. *Bars* 1 cm for unopened flower, and 3 cm for mature fruit, respectively

(*MMFFAA* or *MMFFaa* with only female flowers), andromonoecious (*mmffAA* with bisexual flowers and male flowers), hermaphroditic (*mmFFAA*, or *mmFFaa*, with only perfect flowers), or androecious (*MMffaa* or *mmffaa* with only male flowers) (Mibus and Tatlioglu 2004). Additional modifier genes and environmental factors such as photoperiod, temperature can also influence sex expression in cucumber plants.

The plant hormone ethylene, and in particular, the enzyme 1-aminocyclopropane-1 -carboxylic acid synthase (ACS) in ethylene biosynthesis plays a critical regulatory role in cucumber sex determination. Previous studies (Kamachi et al. 1997, 2000; Trebitsh et al. 1997; Mibus and Tatlioglu 2004; Knopf and Trebitsh 2006) have identified *CsACS1G* in chromosomes 6 as the most possible candidate gene of the *F/f* locus. More recently, the *m* locus in cucumber chromosome 1 was shown may encode *CsACS2* (Boualem et al. 2009; Li et al. 2009).

For breeding application, cucumber is an important economical vegetable crop (Li et al. 2013). Gynoecious sex expression (gynoecy) and fruit shape (in particular, length to diameter ratio) are important yield components in cucumber (Cramer and Wehner 2000). Gynoecy has been shown to be positively correlated with the production, and elongated fruit shape is an important determinant of marketable yield (Serquen et al. 1997). The association of fruit shape and sex expression is a very interesting phenomenon in cucumber. The majority of cucumber cultivars are monoecious or gynoecious, and their fruits usually show elongated shape. However, the fruits developed from perfect flowers on andromonoecious (for example, True Lemon cucumber, Fig. 1) or hermaphroditic plants (for example H34 and WI1983H, Fig. 1) are round. Similar observations have also well documented in melon (for example, Loy 2006; Abdelmohsin and Pitrat 2008; Sakata et al. 2013). It is known that this association of bisexual flower and round fruit shape is due to the pleiotropic effects of the monoecy locus (*m* in cucumber and *a* in melon) rather than due to its close linkage with a fruit shape gene (Robinson 1978; Abdelmohsin and Pitrat 2008). However, the genetic and physiological mechanism of this association is not well understood.

We recently identified a novel cucumber line (H38) which is andromonoecious but bears elongated fruits (Fig. 1). The shape of the mature fruit of H38 was similar to normal fruit which usually developed from the female flower (unisexual flower). Moreover, just like the female flower, the development of the pistil on H38 was more pronounced than in H34 and WI1983H, which may contribute to higher efficiency of pollination and seed setting than the latter two lines. Therefore, we designated the gene for bisexual flower in H38 *m-1*. Cloning of this bisexual flower gene in H38 may clarify its allelic relationship with the *m* locus cloned previously (Boualem et al. 2009; Li et al. 2009), and it may also be helpful to understand the effects of the *m* locus on fruit elongation.

In recent years, many genetic and applied genomics resources in cucumber have been developed including draft genome assemblies (Huang et al. 2009; Yang et al. 2012), molecular markers, and genetic maps (e.g., Ren et al. 2009; Cavagnaro et al. 2010). These resources have greatly facilitated genetic mapping and gene cloning in cucumber. In this study, we developed a tri-parent mapping strategy (TMPS) for rapid gene cloning using two crosses derived

from three parental lines. Two such lines (A and B) were used previously to develop a high-resolution genetic map in cucumber, and the third line (C) carries the target gene to be mapped. Two F_2 segregating populations (A × C and B × C) were developed. Molecular markers were selected from the A × B genetic map for polymorphism screening and genetic mapping using bulk segregant analysis (BSA) (Michelmore et al. 1991). This strategy was exemplified with the rapid cloning of the *m-1* gene in H38.

Materials and methods

Plant materials

Three cucumber inbred lines, 9930 (monoecious, genotype *MMff*), Gy14 (gynoecious, genotype *MMFF*) and H38 (andromonoecious, genotype *m-1m-1FF*) were used for various investigations in the present study. For quick framework mapping of the *m-1* locus in H38 using TPMS, two F_2 populations with 129 and 140 individuals were developed from a cross of H38 with 9930 and Gy14, respectively. To affirm the inheritance of the *m-1* gene, two BC_1 populations were also developed using H38 as the recurrent parent to backcross with the Gy14 × H38 and 9930 × H38 F_1's. Fine genetic mapping was conducted with over 2000 F_2 plants from each of the two crosses. Genetic test was also carried with crossing H34 (a hermaphroditic line used for *m* cloning previously, genotype *mmFF*) to H38. The relative F_1 progeny and F_2 population were planted to screen the sex expression and fruit shapes.

Principle of TPMS: marker development, analysis, and genetic mapping of *m-1* locus

The TPMS was built on pre-selection of markers that are polymorphic between two of the three parental lines which were then applied to tri-parental mapping populations with the bulk segregatant analysis. Two bulks, the unisexual bulk (plants that are either monoecious or gynoecious) and the bisexual bulk (plants with perfect flowers) were constructed from each of the two F_2 populations (Gy14 × H38 and 9930 × H38). For polymorphism screening of the two pairs of bulks, we selected SSR markers from the high-density genetic map using Gy14 × 9930 F_2 population (Yang et al. 2012). Assume a marker has allele A in Gy14, allele B in 9930, and allele C in H38. A marker selected from the Gy14 × 9930 map must be polymorphic between Gy14 and 9930 (A ≠ B). For allele C, there are three possibilities: C ≠ A, C ≠ B; C = A, C ≠ B; and C ≠ A, C = B. Therefore, all markers selected from A × B map will be informative in at least one of the two populations of A × C and B × C. While this pre-selection of markers will significantly increase the chance of detection of polymorphic markers in the two F_2 populations.

Each unisexual bulk consisted of 10 monoecious or gynoecious plants; and each bisexual bulk was composed of 10 andromonoecious or hermaphroditic plants. In general, the chance for identifying markers that are linked to a target gene depends on the number of individuals in the bulk (Wang and Paterson 1994). In the each F_2 population, there were 10 plants in each bulk. The high-density Gy14 × 9930 genetic map we employed contains 783 SSR loci in 708 cM (Yang et al. 2012). If we want to identify at least one polymorphic marker linked with the target gene (*m-1* locus) with confidence interval of 5 cM (1 recombinant in 20 individuals), at least 72 markers are needed. We eventually selected 81 SSR markers (details in Supplementary Table 1) at approximately 10 cM interval on the map (5 cM apart from each side of a target gene) for polymorphic screening of the bulks. Positive markers (polymorphic ones in either pair of the bulks) were then applied to the two F_2 populations (129 plants from 9930 × H38 and 140 from Gy14 × H38) to develop a local genetic map for the target gene.

A draft genome scaffold-based chromosome walking was then initiated for fine mapping of the *m-1* gene. The scaffold sequences of the 9930 (Huang et al. 2009; http://www.icugi.org) and Gy14 (Yang et al. 2012; http://cucumber.vcru.wisc.edu/wenglab/gy14-9930/) draft genomes that harbored flanking SSR markers were downloaded for developing new markers. New SSRs in the target interval were selected, and putative polymorphisms between 9930 and Gy14 were analyzed using in silico PCR strategy following Cavagnaro et al. (2010). Single nucleotide polymorphism (SNP) markers were also identified in the target region based on the sequence of the Gy14 and 9930 genome. The PRIMER PREMIER 6.0 software (http://www.premierbiosoft.com/) was used to design all primers, which were synthesized by Sangon Biological Engineering Technology & Service Co. (Shanghai, China).

The populations used for fine mapping of the *m-1* locus included over 2000 F_2 plants each of Gy14 × H38 and 9930 × H38 crosses same as the primary population. F_2 seedling plants at the cotyledon stage were first screened with two SSR markers flanking the *m-1* locus; putative recombinants between the flanking markers were grown to mature to collect phenotypic data for sex expression to validate the linkage relationship between markers and the *m-1* locus.

The CTAB method was used to extract genomic DNA from young leaves and cotyledon tissue; the PCR procedure followed Li et al. (2008). PCR products were separated on a 6 % denatured polyacrylamide gel (PAGE) with 1 × TBE and stained with $AgNO_3$ solution.

MAPMAKER/EXP3.0 (Lander et al. 1987) was used for linkage analysis with a logarithm of odds (LOD) threshold of 3.0 or more. The Kosambi mapping function (Kosambi 1944) was used in estimation of genetic distances.

Candidate gene sequence analysis

The candidate genes in a 43.6 kb region were analyzed using the Cucumber Genome Database (http://cucumber.genomics.org.cn). The primers designed for the whole-genomic sequence of wild-type *CsACS2* gene (Li et al. 2009) were used to amplify the corresponding CDS and promoter fragment in the H38 line. The PCR product was sub-cloned and sequenced by Sangon Biological Engineering Technology & Service Co. (Shanghai, China). DNAMAN v6.0 software (http://dnaman.software.informer.com/6.0/) was used to compare the DNA and its deduced protein sequences. Multiple sequence alignment of full-length protein sequences was performed using the ClustalW (http://www.ebi.ac.uk/Tools/clustalw2).

Prokaryotic expression and immunological analysis

Prokaryotic expression and immunological analysis of the candidate gene were conducted following Li et al. (2012). Briefly, cucumber apical shoots were excised under a stereomicroscope (Olympus, Japan). Total RNA was isolated from Gy14, H34, and H38 lines from the apical shoots of adult plants (with ~30 nodes on the main stem). Reagents for RNA extraction were purchased from TIANGEN (China). DNase I (Takara, China) was used to remove contaminating DNA, and first-strand cDNA was synthesized with the RevertAid^{TM} first-Strand cDNA Synthesis Kit (Fermentas, China). The coding region of the target gene was amplified using a Long Distance PCR Kit (Takara, China) with appropriate primers (Supplementary Table 1), and the positive results were confirmed by sequencing of the sub-clones. The gene sequences of *CsACS1/G* and wild-type *CsACS2* (named with *CsACS2-M* in this study) were amplified from Gy14 cDNA, and mutant *CsACS2* alleles were prepared from H34 (*CsACS2-m*) and H38 (*CsACS2-m-1*) cDNA, respectively.

All these gene sequences were sub-cloned into the vector pFLAG-MAC using a NovoRec® PCR Kit (Novoprotein, China) to create four translational fusions of the genes to a Flag tag at the N-terminus. To express the genes, all the positive plasmids were transformed into *E. coli* BL21 (DE3). Recombinant strains were grown in Luria–Bertani medium (LB) at 37 °C to an optical OD_{600} of 0.5 to 0.6, and then induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) for another 3 h and harvested. Cells were collected by centrifugation, resuspended in 2 × SDS loading buffer and boiled for 3 min. Cell debris was removed by centrifugation, and the supernatant was stored at −20 °C. For immunoblot analysis, the frozen cells were separated by 10 % SDS-PAGE. Protein bands were electrotransferred from another unstained gel onto PVDF membranes. Immunoprecipitated proteins were detected by immunoblot analysis with an anti-Flag antibody (Sigma) at 1:5000 dilution (Wang et al. 2005).

Enzymatic analysis

The CsACS enzyme assay in vivo was conducted according to Li et al. (2009). In brief, the *E. coli* JAde 6 system was used to analyze the enzymatic function of various alleles of the *m* locus. Tarun et al. (1998) integrated the ACC deaminase gene into the genome of the mutant strain JHM544, which is an Ile auxotroph, and developed the *E. coli* strain JAde 6, which is effective for testing putative ACC synthases. The plasmids carrying target DNA sequences were transformed into JAde 6 using a Competent Cell Preparation Kit (Takara, China). Transformed cells were plated on LB medium and incubated at 37 °C overnight. Colonies from the LB plates were patched onto minimal media (M9) with or without 1 mM ACC (1-aminocyclopropane-1-carboxylic acid). Negative results were identified by their inability to grow on M9 plates after 3 days of incubation at 37 °C (Tarun et al. 1998).

Purified recombinant CsACS-M and CsACS-m-1 protein were used to analyze ACS activity in vitro. The frozen cells were resuspended and lysed with CelLytic B (Sigma) solution. The fusion proteins were purified using ANTI-FLAG® M2 affinity gel (Sigma) as described by the manufacturer, and the proteins were released by competitive elution with FLAG peptide in TBS (50 mM Tris HCl, 150 mM NaCl, pH 7.4). Protein concentrations were determined using the Bradford assay as described by the manufacturer (Bio-Rad). ACS activity was analyzed according to Chae et al. (2003) and modified as Bulens et al. (2011). Five microliters of the purified protein was placed into 22-mL GC vials containing 2 mL of reaction solution [250 mM phosphate buffer, pH 8.0, 10 μM pyridoxal phosphate (PLP), 1 mM EDTA, 5 mM DTT, and protease inhibitors]. After adding 100 μL of 5 mM S-adenosyl-Met, the mixtures were incubated for 30 min at room temperature. The 1-aminocyclopropane-1-carboxylic acid was converted to ethylene using HgCl (20 mM), followed by a 1:1 mix of saturated NaOH:bleach (Lizada and Yang 1979). The tubes were capped immediately and incubated on ice for 10 min. Ten milliliters of headspace was removed, and the ethylene was measured as described by Vogel et al. (1998). All reactions were performed in triplicate and compared with controls of which S-adenosyl-Met was not added.

Table 1 Segregation of sex phenotypes among F_1, F_2, and BC_1 populations derived from cucumber inbred lines H38, Gy14, and 9930

Populations	# Plants examed	# Unisexual plants	# Bisexual plants	Excepted unisexual to bisexual ratio	χ^2 values	*P* values
Gy14 × H38 F_1	12	12	0	1:0		
9930 × H38 F_1	12	12	0	1:0		
Gy14 × H38 F_2	129	88	41	3:1	3.165	0.075
9930 × H38 F_2	140	110	30	3:1	0.952	0.329
(Gy14 × H38) × H38 BC_1	134	66	68	1:1	0.030	0.863
(9930 × H38) × H38 BC_1	129	59	70	1:1	0.938	0.333

Results

Inheritance of monoecy locus in H38

Sex expression of three cucumber inbred lines Gy14, 9930, and H38 and their F_2 and BC_1 derivatives was recorded. For comparison purpose, two other lines, WI1983H and H34 were also included in the observation. The flower and fruit morphology of these five lines is shown in Fig. 1. Segregation data in F_1 and F_2, BC_1 population are presented in Table 1. As expected, Gy14 was gynoecious, 9930 was monoecious, H34 and WI1983H were hermaphroditic, and H38 was andromonoecious. In general, perfect (bisexual) flowers in cucumber bear round fruits (for example, H34 and WI1983H), but H38 was unusual in that it had relatively long fruits just as 9930.

The Gy14 × H38 F_1 plants were gynoecious, and in F_2 and BC_1, the segregation of plants with unisexual (monoecious or gynoecious) flowers to those with bisexual (andromonoecious or hermaphroditic) flowers fit the expected 3:1 and 1:1 ratio, respectively (Table 1). The same was true for the data from the F_1, F_2, and BC_1 populations derived from 9930 × H38, suggesting the monoecy locus for bisexual flowering in H38 was controlled by a simply inherited recessive locus as found in H34 and WI1983H (Li et al. 2009).

Finally, genetic test was carried with crossing H34 to H38, and the F_1 progeny showed a similar bisexual flower structure to the two parental lines (Fig. 1). Since H38 bore long fruit which was different from H34 that bears round fruits, we designated the gene for bisexual flower in H38 *m-1*, and it might be a new allele of the *m* locus for determining the bisexual flower. Nevertheless, when observing the shape of the mature fruit from the F_1 progeny, the long pear-shaped fruit was different from its two parental lines (Fig. 1). All the observations challenge the hypothesis of the pleiotropic effects of the monoecy locus on the association of bisexual flower and round fruit shape in cucumber and melon. To clarify the relationship between fruit shape and sex expression, it was necessary to clone the *m-1* gene in H38.

Application of TPMS: a case study for genetic mapping of the *m-1* locus

The TMPS takes advantage of the high-density genetic map already developed as well as the powder of the BSA (bulk segregating analysis). In this study, to quickly map the *m-1* locus in H38, we used three parental lines (H38, Gy14, and 9930) and developed two F_2 populations, Gy14 × H38 F_2 and 9930 × H38 F_2. Two bulks, bisexual and unisexual, were constructed from each of the two populations. Eighty-one SSRs makers on the high-density cucumber genetic map (Yang et al. 2012) were selected to screen in the two pairs of bulks.

Among the 81 SSR, no polymorphic markers were detected between the two bulks of 9930 × H38, which was due probably to the fact that both 9930 and H38 belong to North China market class cucumbers that were genetically very close to each other. In contrast, four of the 81 markers, SSR01366, SSR23487, UW083752, and UW085142 showed polymorphisms between the two bulks constructed from Gy14 × H38 F_2. Some examples of polymorphism screening in the two pairs of bulks are illustrated in Fig. 2.

The four polymorphic markers between the two bulks were applied 140 Gy14 × H38 F_2 plants, and a linkage

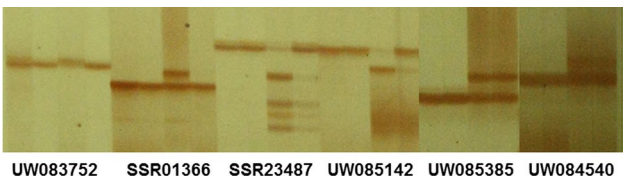

Fig. 2 Examples of PAGE gel profiles of six SSR markers in four bulks for demonstration of TPMS. For each marker, the four *lanes* from *left* to *right* were bulks of unisexual, bisexual bulks of 9930 × H38, unisexual and bisexual bulks of Gy14 × H38, respectively. All the pre-selected markers showed polymorphism between the first two bulks with 9930 background and the latter two bulks with Gy14 background, and the common bands in each bulks were derived from H38. The *m-1* locus-linked markers (SSR01366, SSR23487, UW083752, and UW085142) indicated specific bands in the two bulks from Gy14 × H38, compared with the results of the non-linked markers (UW085385 and UW084540)

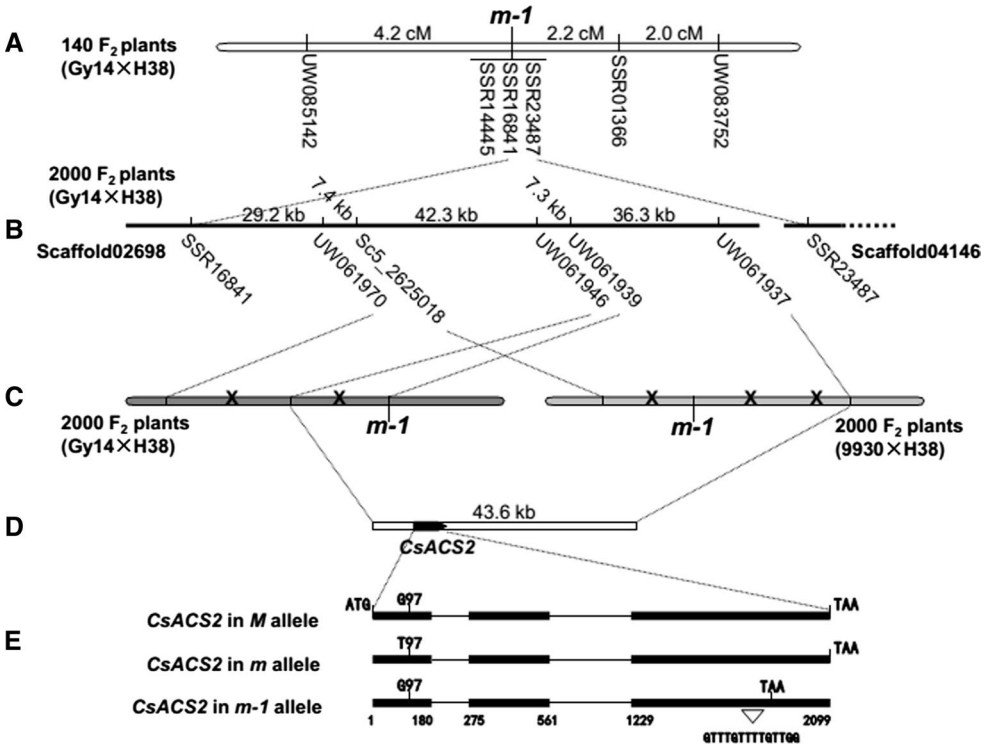

Fig. 3 Map-based cloning of the *m-1* allele. **a** TPMS allowed quick framework map development to locate *m-1* locus in cucumber chromosome 1 with three co-segregating SSR markers. **b–d**, fine mapping delimited *m-1* locus into a 43.6 kb region in which only one gene, *CsACS2* was annotated. **e** alignment of the wild-type *M* allele, the *m* allele of H34 (Li et al. 2009), and the allele in H38 (this study) identified a 14 bp deletion in the third exon of the *CsACS2* gene in H38 that resulted in premature stop codon (TAA) in this gene. "X" in C indicates recombinant events

map was developed for the *m-1* locus which is presented in Fig. 3a. It turned out that SSR23487 was co-segregating with the *m-1* locus; SSR01366 and UW083752 were on one side of *m-1* that was 2.2 cM and 4.2 cM away from *m-1*, whereas UW085142 was on the other side of *m-1* at 4.2 cM away. These markers were all located in cucumber chromosome 1 in a region the *m* locus resides (Li et al. 2009).

Fine genetic mapping of *m-1* locus

The TPMS mapping allowed us to quickly place the *m-1* locus into a relatively small region in cucumber chromosome 1. We then initiated fine genetic mapping of the *m-1* gene. We first tested additional markers already mapped on the Gy14 × 9930 map (Yang et al. 2012) in the 140 Gy14 × H38 F_2 population, and identified two more markers, SSR16841 and SSR14445 that were co-segregating in this population (Fig. 3a). Then, large segregate populations were developed for fine mapping. We first screened 2012 additional F_2 individuals from Gy14 × H38 using two flanking markers SSR01366 and UW085142. To reduce the workload for growing large amount of plants, only plants showing recombination between the two flanking markers were allowed to grow to maturity, which were phenotyped for sex expression. Among the 2012 plants, 112 were recombinants between SSR01366 and UW085142. Linkage analysis established the order of the three markers in relation to the *m-1* locus, and *m-1* was mapped between SSR23487 and SSR16481 at 0.2 cM and 0.3 cM away, respectively. Physically, the region delimited by the two flanking marker was ~180 kb in the Gy14 draft genome scaffold02698 and scaffold04146 (Fig. 3b).

Chromosome walking was conducted in the 180 kb region. To increase the resolving power, polymorphism screening was conducted among the three parental lines (Gy14, 9930 and H38). Four new SSRs markers and one SNP marker were identified; of which, SSR markers UW061970, UW061946, and UW061939 were polymorphic between Gy14 and H38; other two (SSR markers UW061937 and, SNP marker Sc5_2625018) were polymorphic between 9930 and H38. Linkage analysis in both F_2 populations revealed the map locations of these new markers in relation to the *m-1* gene (Fig. 3c), which allowed to delimit the *m-1* locus into a 43.6 kb region (Fig. 3d).

Annotation of this 43.6 kb region identified only one gene, *CsACS2*, that encodes a 1-aminocyclopropane-1-carboxylic acid synthase which has been previously shown to be the *M/m* gene for unisexual/bisexual flowers in cucumber (Boualem et al. 2009; Li et al. 2009). However, alignment of the DNA sequences of the *M* allele in Gy14 and 9930, the *m* allele in H34, and the *m-1* locus in H38 revealed different mutations at the bisexual alleles in H34 and H38 (Fig. 3e). There are three exons in the cucumber *CsACS2* gene. Compared with the CDS of the wild-type *M* allele and the bisexual *m* allele in H34, there was a 14 bp deletion in exon 3 (GenBank accession no. KM272632) in H38, which was different from the previously identified mutation in

H34 and WI1983H. The mutations in the latter two lines were due to one nucleotide transversion in exon 1. The CDS regions of *CsACS2* gene in all recombinant individuals were sequenced, and indicated that the deletion was co-segregating with the *m-1* locus. Additionally, the 1958 bp promoter fragments showed no DNA polymorphisms among the three alleles (GenBank accession no. FJ529216).

Functional validation of *m-1* candidate gene

Fine genetic mapping suggested *CsACS2* was the only candidate gene for the *m-1* locus in H38. Sequence analysis predicted that the 14 bp deletion would cause a frame shift mutation and generate a truncated protein. The deduced truncated protein lost 74 amino acid residues from the carboxyl terminal. When the well-characterized ACS homologous sequences from different plants (*Lycopersicon esculentum*, *Arabidopsis thaliana*, *Cucumis melo*, and *C. sativus*) were aligned (Fig. 5a), we observed that the mutant protein from *m-1* allele missed the conserved box 7 (Rottmann et al. 1991). To verify the truncated product, prokaryotic expression and immunological analysis of the protein were conducted, and the results are shown in Fig. 4. As expected, the results revealed *m-1* allele encoded a truncated protein compared with wild-type *M* gene, and the predicted molecular weight of the protein decreased from 50.4 KD to 41.7 KD (calculated without FLAG tag).

Crystallographic studies have detected that residues in the box 7 are involved in binding of the enzyme substrate S-adenosyl methionine (SAM) (Huai et al. 2001). The critical role of R^{412} for substrate binding is supported by the result that the R407K mutation of apple ACS increased the *K*m value of SAM at least 20-fold (White et al. 1994). In this study, *E. coli* system was used to test the ACS enzymatic activity, and the constructs of the *F*, *M*, *m*, and *m-1* CDS were placed in the *E. coli* mutant strain JAde 6. After 3-day incubation, the enzyme activity assay in vivo with JAde 6 suggested ACS activity in wild-type CsACS1/G and CsACS2, and both H38 (*CsACS2-m-1*) and H34 (*CsACS2-m*) did not show normal enzymatic functions (Fig. 5b). Meanwhile, we expressed CsACS-M and CsACS-m-1 as FLAG tag (DYKDDDDK) fusion proteins in *E. coli* strain BL21 (DE3). Purified recombinant CsACS-M and CsACS-m-1 were assessed for enzymatic activity in vitro by monitoring ethylene formation. The results indicated that CsACS-m-1 isoform was totally inactive (Fig. 5c), which was the same as the observation from JAde 6.

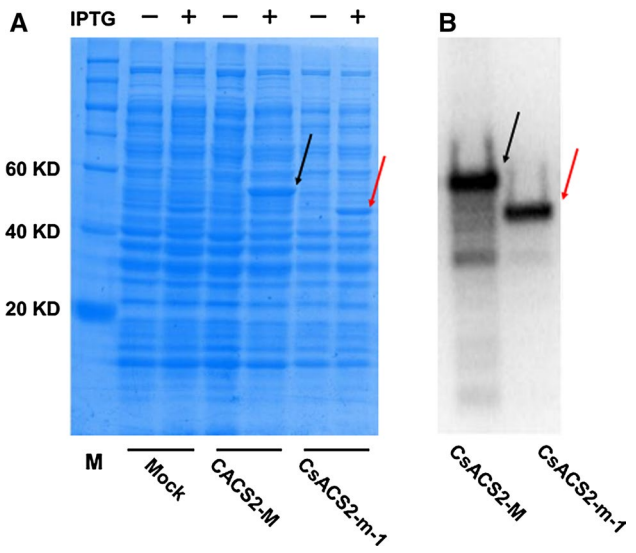

Fig. 4 Analysis of the truncated protein encoded by the *m-1* allele. **a** Flag-tagged wild-type *M* (*black arrowheads*) and *m-1* mutant (*red arrowheads*) constructs were expressed in *E. coli*. **b** Immunoblot analysis of the two proteins using the anti-Flag antibody. *M* molecular mass markers; *Mock* cells transferred with empty pFLAG-MAC vector; *IPTG−* and *IPTG+* cells un-treated and incubated with isopropyl-D-thiogalactopyranoside, respectively

Discussion

A new allele at the *monoecious* (*m*) locus in cucumber

In cucumber, the monoecious sex expression is controlled by the identity of the alleles at the *M* locus. Through the map-based (Li et al. 2009) and homology-based (Boualem et al. 2009) cloning strategies, the *CsACS2* gene was shown to be the candidate gene of the *m* locus in cucumber. Although the cucumber materials used were from different sources, both studies identified the same conserved nucleotide mutation in the first exon (G97T) of the *CsACS2* gene, which produced a missense mutation (G33C) in deduced amino acid sequence leading to andromonoecy. Besides this mutation, two other mutations (P209S and S399L) also seemed to be associated with andromonoecy in some other cucumber cultivars (Boualem et al. 2009). More recently, in an EMS-induced TILLinG population, Boualem et al. (2014) identified six mutations in *CsACS2* coding sequence, three of which resulted in changes in amino acid sequence. Interestingly, phenotypic analysis in backcross population revealed only one of the three CsACS2 mutants showed sexual transition from monoecy to andromonoecy, which happened to be the same as observed previously (G33C), suggesting the critical role of G^{33} amino acid in ACS function for sex determination (Boualem et al. 2014).

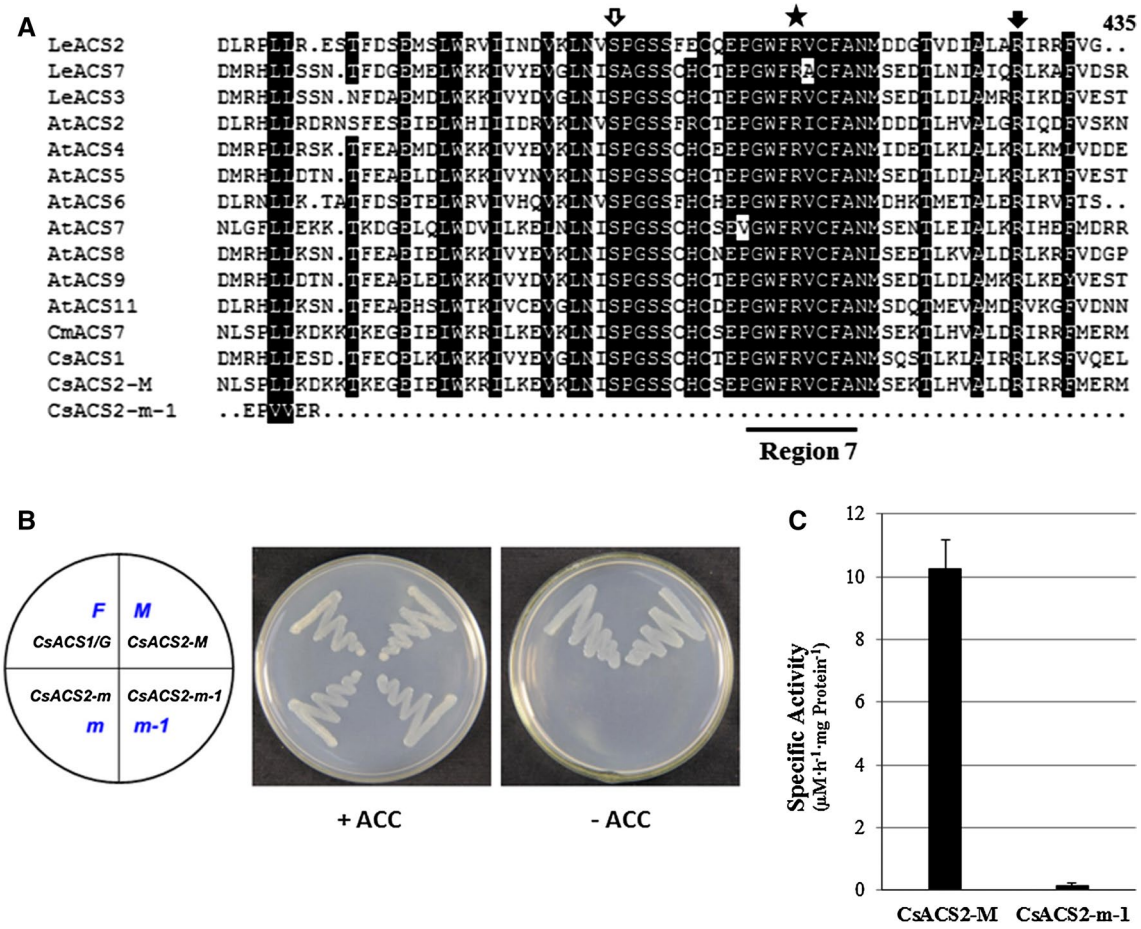

Fig. 5 Alignment of deduced amino acid sequence of ACS isoforms and enzymatic activity analysis. **a** Amino acid alignments of CsACS1 (AB133818), CsACS2-M (FJ971626), CsACS2-m-1 (KM272632) and well-characterized ACS homologous proteins from *L. esculentum* (LeACS2, LeACS3, LeACS7), *A. thaliana* (AtACS2, AtACS4, AtACS5, AtACS6, AtACS7, AtACS8, AtACS9, AtACS11), *Cucumis melo* (CmACS7, EU791279). Numbers above the alignment indicate the amino acid positions along the LeACS2 protein. *Box* 7 indicates the conserved domain in ACS. *Blank arrow* black star and *black arrow* mark the conserved residues S^{399}, R^{412} and R^{429}, respectively. **b** Complementation of the *E. coli* integrative strain JAde 6 with various *CsACS* genes. *Left* schematic representation of the *F*, *M*, *m*, *m-1* genes. The streaked sections in *middle* and *right* correspond to the genes in the pie chart. *Middle* growth of all strains of JAde 6 on minimal media with 1 mM ACC as a positive control. *Right* growth of all strains of JAde 6 on minimal media to analyze the enzymatic functions. **c** Specific activity of purified recombinant CsACS2 proteins. The fractions analyzed in Fig. 4b were assayed for ACS activity as described in Methods. The activity was normalized to the amount of protein present. *Vertical bars* indicated the standard deviation of the means for triplicate samples

In the present study, we identified a novel allele, *m-1*, of the *m* locus in the andromonoecious cucumber line H38, which was due to a 14 bp deletion in exon 3 of *CsACS2* (Fig. 3e). The deletion in CDS region caused a truncated protein, which missed a sequence of 74 amino acid residues in the carboxyl-terminal region. In this sequence, there are some residues critical for the ACS function (Fig. 5a). The conserved box 7 and R^{412} are involved in binding of SAM (Huai et al. 2001). Carboxyl-terminal deletion analysis detected that deletion encompassing the highly conserved R^{429} completely abolished the enzyme activity; thus conservation of R^{429} is important for the sustenance of enzyme activity (Li and Mattoo 1994). The enzymatic conservation of S^{399} was also confirmed, and the S399L isoform was found to be totally inactive (Boualem et al. 2014). Finally, we analyze the enzymatic activity of the mutant protein cloned in this study. The observations from in vivo and in vitro tests confirmed that the truncated product encoded by the *m-1* allele lost the enzyme activity totally. Taking the mapping data together, we concluded that the *m-1* mutation is the case of apparition of hermaphrodite flowers in cucumber line H38.

While the independent cloning of mutation in H38 provides further evidence that *CsACS2* is indeed the *m* locus, this new allele also provides a valuable tool in understanding the molecular mechanisms of *CsACS2* in sex determination in cucumber.

The relationship between fruit shape and sex expression in cucumber

The ovaries and fruits of the majority of monoecious cucumber are usually elongated, whereas those of hermaphroditic and andromonoecious plants are round shaped. Rosa (1928) observed the correlation between sex type and fruit shape in muskmelon, cucumber and watermelon, and he suggested that this may be the result of pleiotropic action of the gene determining sex expression. Kubicki (1962) also found the similar phenomenon, and interpreted his data as a case of apparent linkage of genes determining fruit form and sexuality. In our previous study (Li et al. 2009), we screened more than 5500 F_2 plants from the cross of H34 × S52 and WI1983G × WI1983H. We observed co-segregation between oval/round fruit and bisexual flowers, which might be explained by pleiotropy in the *m* locus. However, in the present study, the andromonoecious H38 bears elongated ovary and normal long mature fruit which is similar to fruits set on monoecious plants (Fig. 1). This finding makes it possible to study the mechanism of the pleiotropy in the *m* locus. Interestingly, the fruit shape of the F_1 between H34 × H38 exhibited long pear shape (Fig. 1). The various fruit shapes from different combination of *m* alleles imply a special regulation mechanism between *m* gene and the fruit shape in bisexual flower background. Unexpectedly, some primary observation showed that the fruit shapes in the F_2 population derived from the crossing of hermaphroditic H34 (*m*) and andromonoecious H38 (*m-1*) were diverse (Supplementary Fig. 1). Even the fruits from a same plant showed different shapes, which was not found among the two parental lines and their F_1 progeny. This may imply that the fruit shape is influenced by the transient balance of endogenous hormone or ethylene. The influence function, which is variable in the fruits from different nodes, is much more obvious in allogeneic *m/m-1* background. Since both *m* and *m-1* alleles showed the same loss-of-function of ACS, except for sex expression, the action of the *m* locus on fruit shape might not depend on the enzymatic function of the *CsACS2* gene in cucumber, and the relationships of sex determination, fruit shape, and CsACS activities need to study in future.

The TPMS mapping strategy

In this study, we employed a TMPS to quickly locate the *m-1* locus in cucumber chromosome 1. The primary motivation for TPMS is to reduce number of primers used for polymorphism screening, and thus save the cost for primer synthesis. In a mapping project, many markers are needed to screen the parental lines to identify polymorphic ones for linkage analysis. Initially, we identified a small number of markers to cover all the genome: the number was controlled by the scale of the genetic map and the average interval between markers in the given species (in the present study this was 1 marker/10 cM). All the markers selected would generate a framework marker system, and the system's utility was unlimited in that same species for further mapping research. Therefore, over the long term, the markers cost next to nothing to produce. After initial mapping, which locates the candidate gene into a genomic interval, all the SNPs and putative polymorphic InDel loci can be calculated between the two reference genomes, which permit the development of new markers. Consequently, for a TPMS platform, after establishing a set of markers for the framework mapping, the all markers for a desired phenotype are restricted to a limited number, which reduces the cost. This is very efficient and cost effective as compared with classical genome-wide polymorphism screening using random markers. The two prerequisites for the TPMS strategy are the availability of a linkage map and development of two segregating populations from three parental lines for the target gene. This should be feasible in most cases. Many traditional genetic tools are available for most of the well-characterized species, and usually, a second population is also needed for accurate genetic analysis in some studies.

We used the cucumber plant to validate the TPMS's efficiency. Cucumber is a plant with a narrow genetic background (Kennard et al. 1994; Horejsi and Staub 1999), which made it very difficult to develop polymorphic markers (Li et al. 2009, 2013; Amano et al. 2013). Even in this study, when we initially mapped the *m-1* locus, all of the 81 markers used revealed no polymorphisms between parental lines 9930 and H38. Their similar ecotype (north China type) might partially explain this phenomenon. Nevertheless, since the segregation data of all the makers had been identified from the genetic map between Gy14 and 9930, it meant that these markers should exhibit polymorphisms between Gy14 and H38, and could use to bulk analysis directly. After initial mapping, a narrow down genome region was found, and all the putative polymorphic markers could be identified with sequence alignment and in silico PCR from the first two parental reference genome sequence. Therefore, all the markers designed and synthesized were sequence targeted and gene location aimed, which also made the work faster, cheaper, and less labor-intensive.

Although we are seeing increasing use of the next generation sequencing technologies and SNP-based genotyping in gene mapping and cloning projects, we believe TPMS has its value for many labs that cannot afford the cost of large scale sequencing or re-sequencing and the associated burden of bioinformatic analysis. On the other hand, as we have shown in the present study, TPMS can take advantage of the genomics resources available in a crop species to expedite the process to gene cloning.

Author contribution statement Z Li and Z. Gong conceived the research and designed the experiments. H. Niu and D. Li developed the plant populations. J. Tan, Q. Tao and Z. Zhang performed most of the experiment including mapping, *E. coli* analysis, SDS-PAGE and Western bolt analysis. J. Tan, H. Niu and Z. Li participated in genotyping and phenotyping. Y. Weng and Z. Li oversaw the entire work and wrote the manuscript.

Acknowledgments The authors wish to thank Professors Huanwen Meng, Xiaofeng Wang, and Dr. Yanfeng Zhang (Northwest A&F University) for their assistance with cucumber materials and immunoblot analysis in this study. This work was supported by the National Natural Science Foundation of China (No. 31471879 and No. 31272185), the Innovation of Agricultural Science and Technology in Shaanxi Province (No. 2015NY081), and the Basic Scientific Research Fund from Northwest A&F University (No. QN2012010).

Compliance with ethical standards

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

References

Abdelmohsin ME and Pitrat M (2008) Pleiotropic effect of sex expression on fruit shape in melon. Cucurbitaceae 2008, Proceedings of IXth EUCARPIA Meeting on Genetics and Breeding of Cucurbitaceae (Pitrat M ed), INRA, Avignon (France), 21–24 May 2008

Amano M, Mochizuki A, Kawagoe Y, Iwahori K, Niwa K, Svoboda J, Maeda T, Imura Y (2013) High-resolution mapping of *zym*, a recessive gene for *Zucchini yellow mosaic virus* resistance in cucumber. Theor Appl Genet 126:2983–2993

Boualem A, Troadec C, Kovalski I, Sari M-A, Perl-Treves R, Bendahmane A (2009) A conserved ethylene biosynthesis enzyme leads to andromonoecy in two *Cucumis* species. PLoS One 4:e6144

Boualem A, Fleurier S, Troadec C, Audigier P, Kumar APK, Chatterjee M, Alsadon AA, Sadder MT, Wahb-Allah MA, Al-Doss AA, Bendahmane A (2014) Development of a *Cucumis sativus* TILLinG platform for forward and reverse genetics. PLoS One 9:e97963

Bulens I, van de Poel B, Hertog ML, de Proft MP, Geeraerd AH, Nicolaï BM (2011) Protocol: an updated integrated methodology for analysis of metabolites and enzyme activities of ethylene biosynthesis. Plant Methods 7:17

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

Chae HS, Faure F, Kieber JJ (2003) The *eto1*, *eto2*, and *eto3* mutations and cytokinin treatment increase ethylene biosynthesis in *Arabidopsis* by increasing the stability of ACS protein. Plant Cell 15:545–559

Cramer CS, Wehner TC (2000) Path analysis of the correlation between fruit number and plant traits of cucumber populations. HortScience 35:708–711

Galun E (1961) Study of the inheritance of sex expression in the cucumber, the interactions of major genes with modifying genetic and non-genetic factors. Genetica 32:134–136

Horejsi T, Staub JE (1999) Genetic variation in cucumber (*Cucumis sativus* L.) as assessed by random amplified polymorphic DNA. Genet Res Crop Evol 46:337–350

Huai Q, Xia Y, Chen Y, Callahan B, Li N, Ke H (2001) Crystal structures of 1-aminocyclopropane-1-carboxylate (ACC) synthase in complex with aminoethoxyvinylglycine and pyridoxal-5′-phosphate provide new insight into catalytic mechanisms. J Biol Chem 276:38210–38216

Huang SW, Li RQ, Zhang ZH, Li L, Gu XF, Fan W et al (2009) The genome of the cucumber, *Cucumis sativus* L. Nat Genet 41:1275–1281

Kamachi S, Sekimoto H, Kondo N, Sakai S (1997) Cloning of a cDNA for a 1-aminocyclopropane-1-carboxylate synthase that is expressed during development of female flowers at the apices of *Cucumis sativus* L. Plant Cell Physiol 38:1197–1206

Kamachi S, Mizusawa H, Matsuura S, Sakai S (2000) Expression of two 1-aminocyclopropane-1-carboxylate synthase genes, *CS-ACS1* and *CS-ACS2*, correlated with sex phenotypes in cucumber plants (*Cucumis sativus* L.). Plant Biothnol 17:69–74

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

Knopf RR, Trebitsh T (2006) The female-specific *CS-ACS1G* gene of cucumber. A case of gene duplication and recombination between the non-sex-specific 1-aminocyclopropane-1-carboxylate synthase gene and a branched-chain amino acid transaminase gene. Plant Cell Physiol 47:1217–1228

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Kubicki B (1962) Inheritance of some characters in muskmelons (*Cucumis melo*). Genet Pol 3:265–274

Kubicki B (1969) Investigations on sex determination in cucumber (*Cucumis sativus* L.). VII. Andromonoecism and hermaphroditism. Genet Pol 10:101–120

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA (1987) MAPMARKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural population. Genomics 1:174–181

Li N, Mattoo AK (1994) Deletion of the carboxyl-terminal region of 1-aminocyclopropane-1-carboxylic acid synthase, a key protein in the biosynthesis of ethylene, results in catalytically hyperactive, monomeric enzyme. J Biol Chem 269:6908–6917

Li Z, Pan JS, Guan Y, Tao QY, He HL, Si LT, Cai R (2008) Development and fine mapping of three co-dominant SCAR markers linked to the *M/m* gene in the cucumber plant (*Cucumis sativus* L.). Theor Appl Genet 117:1253–1260

Li Z, Huang SW, Liu SQ, Pan JS, Zhang ZH, Tao QY, Shi QX, Jia ZQ, Zhang WW, Chen HM, Si LT, Zhu LH, Cai R (2009) Molecular isolation of the *M* gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants. Genetics 182:1381–1385

Li Z, Wang S, Tao QY, Pan JS, Si LT, Gong ZH, Cai R (2012) A putative positive feedback regulation mechanism in *CsACS2* expression suggests a modified model for sex determination in cucumber (*Cucumis sativus* L.). J Exp Bot 63:4475–4484

Li YH, Wen CL, Weng YQ (2013) Fine mapping of the pleiotropic locus *B* for black spine and orange mature fruit color in cucumber identifies a 50 kb region containing a R2R3-MYB transcription factor. Theor Appl Genet 126:2187–2196

Lizada MCC, Yang SF (1979) A simple and sensitive assay for 1-aminocyclopropane-1-carboxylic acid. Anal Biochem 100:140–145

Loy JB (2006) Fruit size in melon in monoecious and andromonoecious isolines. Cucurbit Genetics Cooperative Report no 28–29, pp 12–13

Mibus H, Tatlioglu T (2004) Molecular characterization and isolation of the *F/f* gene for femaleness in cucumber (*Cucumis sativus* L.). Theor Appl Genet 109:1669–1676

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

Ren Y, Zhang ZH, Liu JH, Staub JE, Han YH, Cheng ZC, Li XF, Lu JY, Miao H, Kang HX, Xie BY, Gu XF, Wang XW, Du YC, Jin WW, Huang SW (2009) An integrated genetic and cytogenetic map of the cucumber genome. PLoS One 4:e5795

Robinson RW (1978) Association of fruit shape and sex expression in the cucumber. Cucurbit Genetics Cooperative Report 1: 10 (Article 7)

Rosa JT (1928) The inheritance of flower types in *Cucumis* and *Citrullus*. Hilgardia 3:233–250

Rottmann WH, Peter GF, Oeller PW, Keller JA, Shen NF, Nagy BP, Taylor LP, Campbell AD, Theologis A (1991) 1-aminocyclopropane-1-carboxylate synthase in tomato is encoded by a multigene family whose transcription is induced during fruit and floral senescence. J Mol Biol 222:937–961

Sakata Y, Fukino N, Ohara T, Sugiyama M, Yoshioka Y (2013) Effect of monoecious trait on fruit shape and total soluble solid contents of melon. Hort Res (Japan) 12:15–22

Serquen FC, Bacher J, Staub JE (1997) Genetic analysis of yield components in cucumber at low density. J Am Soc Hort Sci 122:522–528

Tarun AS, Lee JS, Theologis A (1998) Random mutagenesis of 1-aminocyclopropane-1-carboxylate synthase: a key enzyme in ethylene biosynthesis. Proc Natl Acad Sci USA 95:9796–9801

Trebitsh T, Staub JE, O'Neill SD (1997) Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the *Female* (*F*) locus that enhances female sex expression in cucumber. Plant Physiol 113:987–995

Vogel JP, Woeste KW, Theologis A, Kieber JJ (1998) Recessive and dominant mutations in the ethylene biosynthetic gene *ACS5* of *Arabidopsis* confer cytokinin insensitivity and ethylene overproduction, respectively. Proc Natl Acad Sci USA 95:4766–4771

Wang GL, Paterson AH (1994) Assessment of DNA pooling strategies for mapping of QTLs. Theor Appl Genet 88:355–361

Wang XF, Goshe MB, Soderblom EJ, Phinney BS, Kuchar JA, Li J, Asami T, Yoshida S, Huber SC, Clouse SD (2005) Identification and functional analysis of in vivo phosphorylation sites of the *Arabidopsis* BRASSINOSTEROID-INSENSITIVE1 receptor kinase. Plant Cell 17:1685–1703

White MF, Vasquez J, Yang SF, Kirsch JF (1994) Expression of apple 1-aminocyclopropane-1-carboxylate synthase in *Escherichia coli*: kinetic characterization of wild-type and active-site mutant forms. Proc Natl Aced Sci USA 91:12428–12432

Yang LM, Koo D-H, 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