Theoretical and Applied Genetics (2018) 131:569–579
https://doi.org/10.1007/s00122-017-3019-9

ORIGINAL ARTICLE

GLABROUS (*CmGL*) encodes a HD-ZIP IV transcription factor playing roles in multicellular trichome initiation in melon

Huayu Zhu[1] · Xiaofen Sun[1] · Qi Zhang[2] · Pengyao Song[1] · Qianmei Hu[1] · Xiaojing Zhang[1] · Xiang Li[1] · Jianbin Hu[1,4] · Junsong Pan[2] · Shouru Sun[1,4] · Yiqun Weng[3] · Luming Yang[1]

Received: 5 September 2017 / Accepted: 10 November 2017 / Published online: 17 November 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

Key message **Map-based cloning identified *CmGL* that encodes a HD-ZIP type IV transcription factor that controls multicellular trichome initiation in melon.**

Abstract Trichomes are small hairs covering the aerial parts of plants that originate from the epidermal cells, which can protect plants against the damage by insects and pathogens. The regulatory pathway of unicellular trichomes has been well studied in the model plant *Arabidopsis*. Little is known about the genetic control and regulation of trichome development in melon (*Cucumis melo* L.) which has multicellular trichomes. In this study, we identified a melon mutant, *cmgl*, which showed completely glabrous on all aerial organs. A bulked segregant analysis was conducted to identify polymorphic markers for linkage analysis in a population with 256 F_2 plants, which allowed to locate the *cmgl* locus in melon chromosome VIII. Next-generation sequencing-aided marker discovery and fine mapping in a large population with 1536 F_2 plants narrowed the candidate gene region to 12 kb that harbored only one candidate gene for *cmgl*, which encoded a class IV homeodomain-associated leucine zipper transcription factor. Four SNPs in the coding region of the *CmGL* gene were identified between the two parental lines; a single base substitution from C to A resulted in a premature termination codon and a truncated protein in the *cmgl*. The SNP was converted into a dCAPS marker, which showed co-segregation in the F_2 population and 564 melon accessions. Result of this study will be helpful for better understanding of genetic control of trichome development in melon and marker-assisted selection in developing new cultivars.

Introduction

Melon (*Cucumis melo* L.,) is one of the most highly diversified species in the Cucurbitaceae family which also includes other important crops such as cucumber, watermelon and squash/pumpkin. It shows a high level of morphological variability in fruit characteristics such as shape, color, texture, taste and composition. Melon is divided in two major subspecies: subsp. *melo* and subsp. *agrestis* (Kirkbride 1993). The genome of melon has been sequenced in 2012 using NGS sequencing technology and a double haploid line DHL 92, and the genome assembly was further improved using high-density SNP markers (Garcia-Mas et al. 2012; Argyris et al. 2015). With the rapid development of genetic and genomic resources, melon has become a model of choice for studying several important biological processes such as plant sex determination (Boualem et al. 2008, 2015; Martin et al. 2009), and phloem transport (Zhang et al. 2010). Furthermore, several important genes controlling fruit quality in melon have been cloned such as *CmKFB* gene regulating

Communicated by Herman J. van Eck.

Huayu Zhu and Xiaofen Sun contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-017-3019-9) contains supplementary material, which is available to authorized users.

✉ Luming Yang
ylmhort@163.com

[1] College of Horticulture, Henan Agricultural University, 63 Nongye Road, Zhengzhou 450002, China

[2] School of Agriculture and Biology, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai 200240, China

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

[4] Henan Key Laboratory of Fruit and Cucurbit Biology, Zhengzhou 450002, China

flavonoid accumulation (Feder et al. 2015) and *CmPH* gene determining fruit acidity (Cohen et al. 2014; Sherman et al. 2013).

Trichomes are hair-like structures which can protect plants against the damage by insects, pathogens, UV irradiation, low temperature and excessive transpiration (Hulskamp 2004; Hulskamp et al. 1999; Schellmann and Hulskamp 2005). In addition, trichomes may help plants attract pollinators and disperse seeds (Serna and Martin 2006). Trichomes exhibit high morphological variation, and can be divided into several forms, which could be unicellular or multicellular, glandular or glandless, and branched or unbranched (Werker 2000). As the representative of unicellular trichomes, the regulatory pathway of trichomes has been widely studied in the model plant *Arabidopsis thaliana* (Hulskamp 2004; Hulskamp et al. 1999; Schellmann and Hulskamp 2005). The initiation of trichome differentiation is regulated by *GL2* (*GLABRA2*), which belongs to the class IV homeodomain-leucine zipper (HD-ZIP IV) family (Ishida et al. 2008; Pesch and Hulskamp 2009; Schiefelbein 2003). The *GL2* was activated by an upstream MYB–bHLH–WD40 trimeric complex which is comprised by three transcription factors, *GL1* (*GLABRA1*), which encodes an R2R3 MYB protein (Kirik et al. 2005), *GL3* (*GLABRA3*) and its homolog *EGL3* (*ENHANCER OF GLABRA3*), which are basic helix–loop–helix proteins (Payne et al. 2000), and *TTG1* (*TRANSPARENT TESTAGLABRA1*), a WD40-repeat protein (Walker et al. 1999). *TRY* (*TRIP-TYCHON*), *CPC* (*CAPRICE*), and ETC (*ENHANCER OF TRY AND CPC*) are negative regulators, which all belong to the R3 single-repeat MYB family. They can compete with the R2R3 MYB proteins to bind the bHLH protein which will form an inactivating complex and thereby inhibit the initiation of trichome differentiation (Kirik et al. 2004; Wester et al. 2009).

The regulatory mechanisms of multicellular trichome formation are still largely unknown. The *MIXTA*, an R2R3-type MYB transcription factor encoding the conserved DNA-binding domain similar to *GL1*, could trigger long-stalked multicellular trichome formation when ectopically expressed in tobacco, while the ectopic expression of *MIXTA* in *Arabidopsis gl1-1* mutant failed to restore the trichome phenotype (Payne et al. 1999). Moreover, over-expression of *GL1* in tobacco had no effect on trichome formation, suggesting that the developments of unicellular and multicellular trichomes are likely controlled by different molecular mechanisms. This is further confirmed by the *Wo* gene which was identified in tomato encoding a HD-ZIP IV transcription factor and could regulate the formation of type I multicellular trichomes. Overexpression of *Wo* gene had no effect on trichome initiation in *Arabidopsis* (Yang et al. 2011). These data demonstrate that the initiation of multicellular trichomes in tobacco and tomato and unicellular trichomes in *Arabidopsis* may be regulated by different pathways.

Trichomes are commonly found in melon plants covering the leaves, stems, branches, flowers, and tendrils, and the trichomes show multicellular, unbranched, glandless which are different from the unicellular trichomes in *A. thaliana*. There is still little study on the molecular mechanisms of trichome development in melon, but several spontaneous glabrous mutants have been characterized and studied in its close relative cucumber. The first mutant *csgl1* shows no observable trichomes on leaves, stems, tendrils, and floral organs, but has obvious trichomes on the hypocotyl. Map-based cloning has revealed that *CsGL1* encodes a Class I HD-ZIP TF, and the loss-function of *csgl1* is due to a 2649-bp genomic DNA deletion spanning the first and second exons (Chen et al. 2014; Zhao et al. 2015). The *csgl3/tril* mutant exhibits a completely glabrous phenotype on cotyledons, hypocotyl, young leaves, fruits, and fruit stalks. It has been mapped on cucumber chromosome 6 and encodes a class IV HD-ZIP TF which lost its function due to the insertion of a 5-kb long terminal repeat (LTR) retrotransposon in the fourth exon of *CsGL3* (Cui et al. 2016; Pan et al. 2015; Wang et al. 2016). In melon, a completely glabrous mutation was first reported by Foster (1963) which can be recognized readily making it an excellent seedling marker in breeding. Genetic analysis indicated that *glabrous* (*gl*) is controlled by a single recessive gene. However, the *gl* gene is still not mapped and the nucleotide sequence information for this gene remains unknown. Thus, cloning and characterizing the *gl* gene are necessary for understanding the mechanism of multicellular trichome formation in melon.

In this study, a completely glabrous muskmelon line NSL73046 was used to cross with a wild-type line M465 for mapping of the *glabrous* (*cmgl*) gene. Based on the resequencing of the two parental lines, Indel markers were developed in the candidate region which were then used in genotyping a large F_2 mapping population. A candidate gene for *cmgl* was identified encoding a HD-ZIP IV transcription factor, and the loss-of-function in the mutant was due to a single base substitution resulting in a premature termination codon.

Materials and methods

Plant materials and mapping population

NSL73046 (the mutant type) is a muskmelon mutant from USDA-ARS GRIN. It displayed completely glabrous on all aboveground organs including leaves, stems, flowers, and a permanent yellow-green leaf color throughout its lifespan which have been proved to be controlled by two different recessive genes, *glabrous* and *yellow-green plant* (Foster 1963). To confirm the inheritance mode and fine mapping of *cmgl* gene in melon, NSL73046 was used to cross with a

normal muskmelon material M465 (the wild type) to generate a large F_2 mapping population. The χ^2-test for goodness-of-fit was used to test for deviation of the observed data from the theoretically expected segregation for trichome phenotype data in F_2 plants.

To investigate the allelic diversity of the *cmgl* candidate gene in natural populations, the trichome phenotypes of a panel including 564 melon accessions with diverse genetic backgrounds were examined using *CmGL* co-segregating dCAPS marker. Information for these melon materials was provided in Supplementary Table S1. All the melon lines and F_2 population were grown in the greenhouse at Maozhuang Research Station of Henan Agricultural University (Zhengzhou, China). Unexpanded young leaves from these plants were collected into 1.5 mL microcentrifuge tubes, lyophilized in a freeze dryer, and ground into fine powder. Genomic DNA was extracted using the CTAB method (Murray and Thompson 1980).

Scanning electron microscopy

The samples of leaves, stems and hypocotyls from NSL73046 and M465 were fixed in FAA (formaldehyde acetic acid–ethanol which contained 50% (v/v) ethanol, 5% (v/v) acetic acid and 3.7% (v/v) formaldehyde dissolved in water) at 4 °C for 24 h, then dehydrated through gradient ethanol elution [50, 60, 70, 85, 90, 95 and 100% (v/v)] and critical point dried in a Leica EM CPD030 desiccator. The dried specimens were sputter-coated with gold–palladium and observed under JSM-6360LV and JEM-2010HT scanning electron microscopes.

Marker development, genetic mapping strategy and linkage analysis

A BSA strategy was used to screen for polymorphic SSR markers between two DNA pools selected from the F_2 population for quick identification of molecular markers linked with the *cmgl* locus. The DNA pools of mutant and normal plants consisted of ten glabrous plants and 10 non-glabrous plants, respectively. 384 SSR markers evenly distributed in 12 melon chromosomes (Zhu et al. 2016) were selected to screen for polymorphisms. After initial anchor of *cmgl* in chromosome 8, a scaffold-based chromosome walking strategy was taken to identify more closely linked markers. New SSR markers in the target scaffolds were selected for primary mapping. At fine mapping stage, SNPs and Indels were explored based on the sequence difference by comparing the genome resequencing data between NSL73046 and M465. For Indels, only those with ≥ 3 bp differences were utilized for primer design with Primer3 (http://primer3.ut.ee/). For SNP genotyping, dCAPS markers were developed with dCAPS Finder 2.0 (Neff et al. 1998).

The PCR amplification of molecular markers and gel electrophoresis was conducted as described in Zhu et al. (2016). Linkage analysis of the *cmgl* locus with molecular markers was performed with the Kosambi mapping function using JoinMap 3.0 with the threshold LOD score of 4.0.

DNA sequencing, sequence comparison and gene prediction

Genome resequencing of NSL73046 and M465 was carried out following the standard Illumina protocol, and the library was used for paired-end sequencing on the Illumina HiSeq 2500 analyzer. After removing short reads and low quality reads, the clean reads were used for mapping to melon reference genome DHL92 (https://melonomics.net/) (Garcia-Mas et al. 2012) using software BWA (Li and Durbin 2009). SNPs and small indels detected from the alignments were called using Samtools and output was given in pileup format (Li et al. 2009). The quality score of the SNPs was assigned by Samtools to evaluate the reliability of SNP calling based on the Phred-scaled probability that the consensus is identical to the reference. The SNPs and small indels between two parental lines were detected using the GATK software tool package (DeLuca et al. 2012), and the reliable SNPs and small indels were noted and predicted using SnpEff software (Chen et al. 2009).

Expression analysis with qRT-PCR

To examine expression patterns of the candidate gene, qRT-PCR analysis was performed with tissues from NSL73046 and M465 including cotyledon, leaf, flower, fruit and tendril. Total RNA of different fresh tissues was extracted using the Plant RNA kit (Omega, USA) and reverse-transcribed using SuperScript III reverse transcriptase (Invitrogen, USA). All experiments were performed with three biological and three technical replicates. The actin gene (MELO3C008032) of melon was used as an internal control. qRT-PCR was performed using the SYBR Green PCR master mix (Applied Biosystems Inc., USA) in iCycleriQTM 5 Multicolor Real-Time PCR detection system (Bio-Rad, USA). The threshold cycle (C_t) value of actin protein was subtracted from that of *CmGL* to obtain a ΔC_t value. The C_t value of control sample M465 was subtracted from the ΔC_t value to obtain a $\Delta\Delta C_t$ value. The fold changes in expression level relative to M465 were expressed as $2^{-\Delta\Delta C_t}$. Student's *t* test was used for comparison between NSL73046 and M465. A difference was considered to be statistically significant when $P < 0.05$. The primers used in this study were listed in Supplementary Table S2.

Identification of HD-ZIP IV gene in melon and phylogenetic analysis

To investigate the phylogenetic relationships of melon CmGL protein with other identified members of HD-ZIP IV family, the proteins used in this study followed those described in Pan et al. (2015) and an additional protein GhHOX3 regulating cotton fiber development was also included (Shan et al. 2014). The sequences in *Arabidopsis*, cucumber (*Cucumis sativus*), tomato (*Solanum lycopersicum*), maize (*Zea mays*), rice (*Oryza sativa*) and cotton (*Gossypium* spp.) were downloaded from NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/). The full-length protein sequences of different species were aligned using the CLUSTX program (Chenna et al. 2003) and the alignment report was visualized by Gendoc. The phylogenetic tree was constructed using the maximum likelihood method in MEGA 5 with a bootstrap of 1000 replicates (Tamura et al. 2011).

We further identified the genome-wide HD-ZIP IV gene members in melon genome using the method followed by Fu et al. (2013). All the identified HD-ZIP IV proteins were further confirmed by Pfam (http://pfam.xfam.org) and SMART (http://smart.embl-heidelberg.de/). The 11 HD-ZIP IV genes identified from cucumber 9930 genome Version2 were downloaded from the Cucurbit Genomics Database (http://www.icugi.org/) (Fu et al. 2013). All full-length protein sequences of non-redundant HD-ZIP IV genes in cucumber and melon were aligned using the CLUSTX program (Chenna et al. 2003) and a phylogenetic tree was constructed using the maximum likelihood method in MEGA 5 (Tamura et al. 2011).

Results

Morphological characterization and inheritance of glabrous mutant in melon

The wild-type melon line M465 had lots of trichomes covering the hypocotyl, cotyledons, young leaves, stem, tendrils, flowers, ovaries and fruits, while the mutant line NSL73046 showed a completely glabrous phenotype on all aboveground organs (Fig. 1). Consistent with these results, we further examined the hypocotyl, leaves and stems of the mutant under a scanning electron microscope (Fig. 2) and found no trichomes on these organs. The trichomes of the wild-type plant were all multicellular consisting of three distinct cell types (head cell, stalk cell, and base cell), which was similar to cucumber (Cui et al. 2016; Pan et al. 2015; Wang et al. 2016).

Among the F_2 population of NSL73046 × M465, the segregation of wild-type and glabrous plants was in agreement with a single recessive gene underlying the trichome-free phenotype (Table 1), which was designated *cmgl* hereinafter. This result was also consistent with early observations by Foster (1963) who suggested that the glabrous phenotype was controlled by a single gene.

Fine mapping of *cmgl* gene

The BSA method was employed for framework mapping of the *cmgl* locus. Of 384 SSR markers selected for polymorphism screening, three were polymorphic between the non-glabrous and glabrous pools, all of which were located on melon chromosome VIII. Then 96 additional SSR markers around this region were selected from a SSR primer list developed from the melon draft genome (Zhu et al. 2016) and five of them were polymorphic between the two bulks. All eight polymorphic SSR markers were used for genotyping 256 F_2 plants, and the *cmgl* gene was mapped between markers CmSSR19480 and CmSSR19495 with a genetic distance of 0.93 and 0.01 cM from the target gene, respectively (Fig. 3a). Information about these markers used in the present study is provided in Supplemental Table S2.

To further narrow down the candidate region harboring the *cmgl* locus, the two parental lines NSL73046 and M465 were re-sequenced using Illumina Hi-Seq 2500. After trimming low quality reads, 83,532,402 (25.06 Gb) and 89,589,250 (26.88 Gb) clean reads were obtained for NSL73046 and M465, respectively. Alignment of these clean reads against the reference genome identified 159 high quality SNPs and 78 Indels in the candidate gene region between two parental lines. Six Indels with difference larger than 4 bp were selected for marker development, and four of them, CmSSR19480, CmSSR19495, Indel1 and Indel6, were polymorphic between two parental lines, which were used for genotyping 438 F_2 plants. This allowed to further narrow the *cmgl* gene down to a 50 kb region in scaffold00007 with a genetic distance of 0.3 and 0.1 cM to Indel1 and Indel6, respectively (Fig. 3b). The two flanking markers Indel1 and Indel6 were genotyped in an extended F_2 population containing 1536 plants, and 8 recombinant plants were identified. Two new polymorphic markers, Indel2 and Indel5, were developed inside this 50 kb region and used for genotyping these recombinant plants. The *cmgl* gene was finally mapped in the region of 11.58 kb between Indel2 and Indel5 (Fig. 3c). Searching the melon genome database (https://melonomics.net/) and manual annotation, only one candidate gene, MEL03COO7377, was identified in this region supporting it is the only candidate gene for the *cmgl* locus.

We examined the expression level of *CmGL* with qRT-PCR in the cotyledon, leaf, flower, fruit and tendril of both parental lines. The *CmGL* gene was expressed in all these organs with the highest expression detected in the leaf

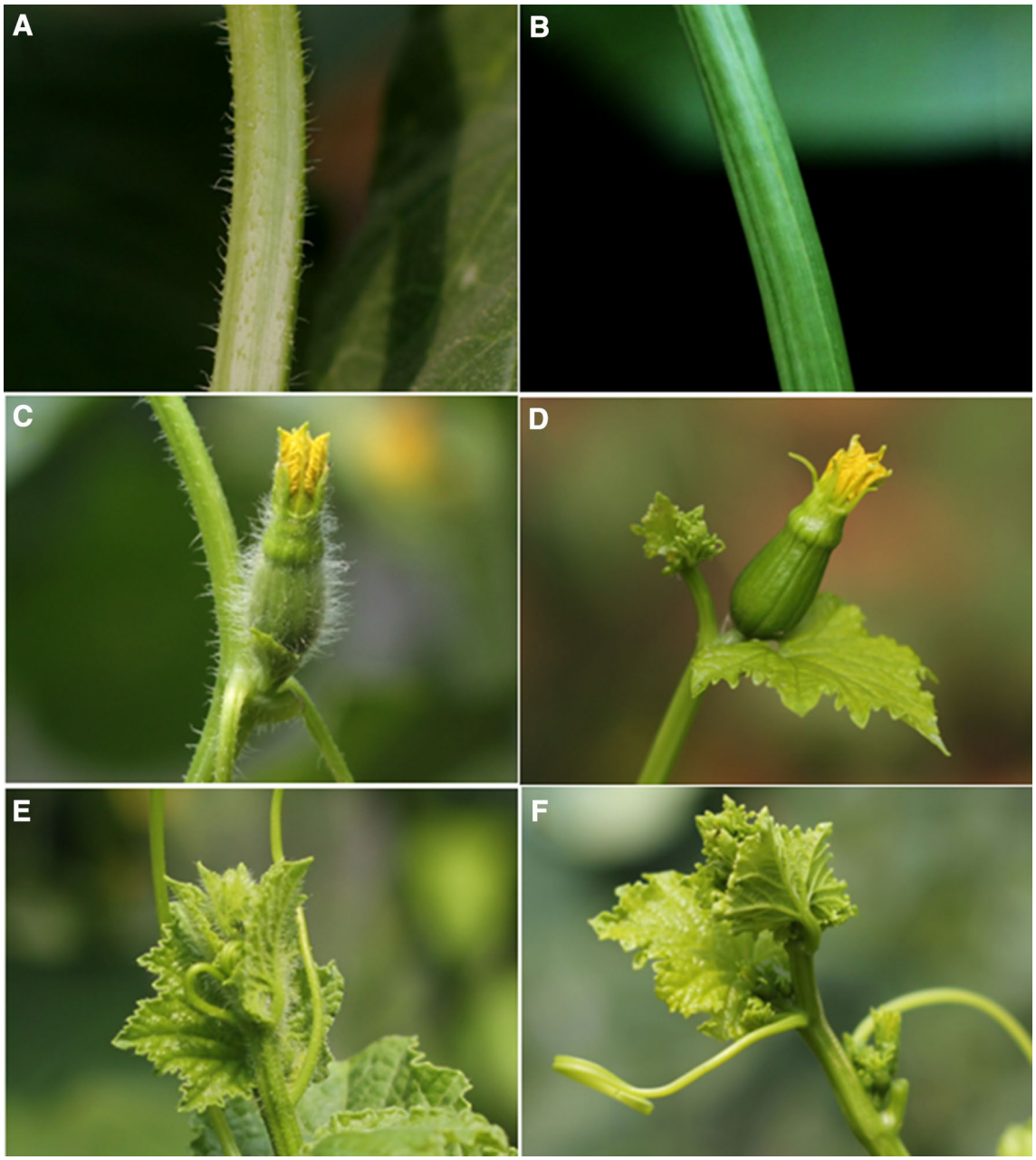

Fig. 1 Trichome phenotypic characterization of stem, female flower and leaf in M465 (**a**, **c**, **e**) and NSL73046 (**b**, **d**, **f**)

(Fig. 4) and the expression level in leaf in the mutant line was significantly lower than that in the wild-type.

Gene annotation, sequence alignment and allelic diversity of *CmGL*

The length of genomic DNA of MEL03COO7377 in the melon reference genome DHL92 was 5367 bp, which was predicted to contain ten exons and the full-length CDS was 2166 bp. The genomic sequence of the candidate gene was compared between two parental lines NSL73046 and M465 by aligning the NGS resequencing reads to the reference genome, and ten SNPs were identified between the wild-type line and mutant line. Among of them, six SNPs were located in the introns, while four SNPs were located in the exons 4, 8 and 10, respectively. The single base substitution in the fourth exon changing from C to A resulted in a premature termination codon in the mutant line, which led to two truncated ORFs predicted in the same genomic region (Fig. 5a). Based on this SNP difference, a dCAPS marker, dCPAS1, was developed which showed co-segregation with glabrous plants in the F_2 population (Fig. 3c). The deduced

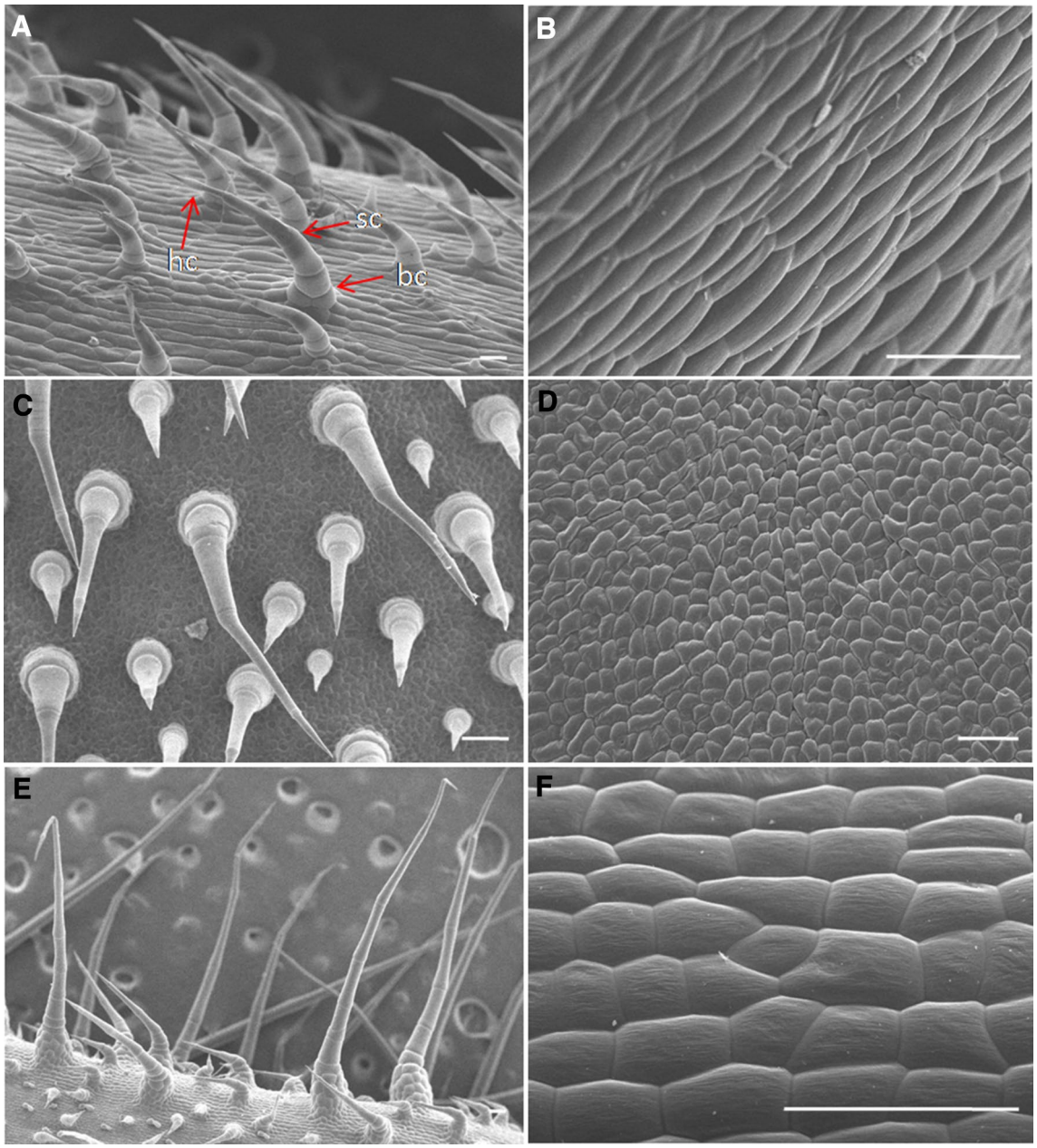

Fig. 2 Scanning electron microscopy images of hypocotyl, young leaf and stem in M465 (**a**, **c**, **e**) and NSL73046 (**b**, **d**, **f**). Arrows indicate the head cell (upper), stalk cell (middle), and base cell (lower) of trichome. Bars = 50 µm

Table 1 Segregation analysis of the trichome trait in the F_2 population

Population	Total	Glabrous type	Wild type	Expected ratio	χ^2	*P*
F_1	19	0	19	–	–	–
F_2	256	65	191	3:1	0.01	0.92
F_3	1536	380	1156	3:1	0.56	0.45

amino acid sequence of *CmGL* contained 721 amino acids and was predicted to encode a HD-ZIP IV transcription factor containing a conserved HOX domain (from 48 to 111 aa) and the START domain (from 237 to 459 aa) (Fig. 5b), while they were separated into two truncated ORFs in the mutant line. BLASTP search in NCBI showed that it shared a high

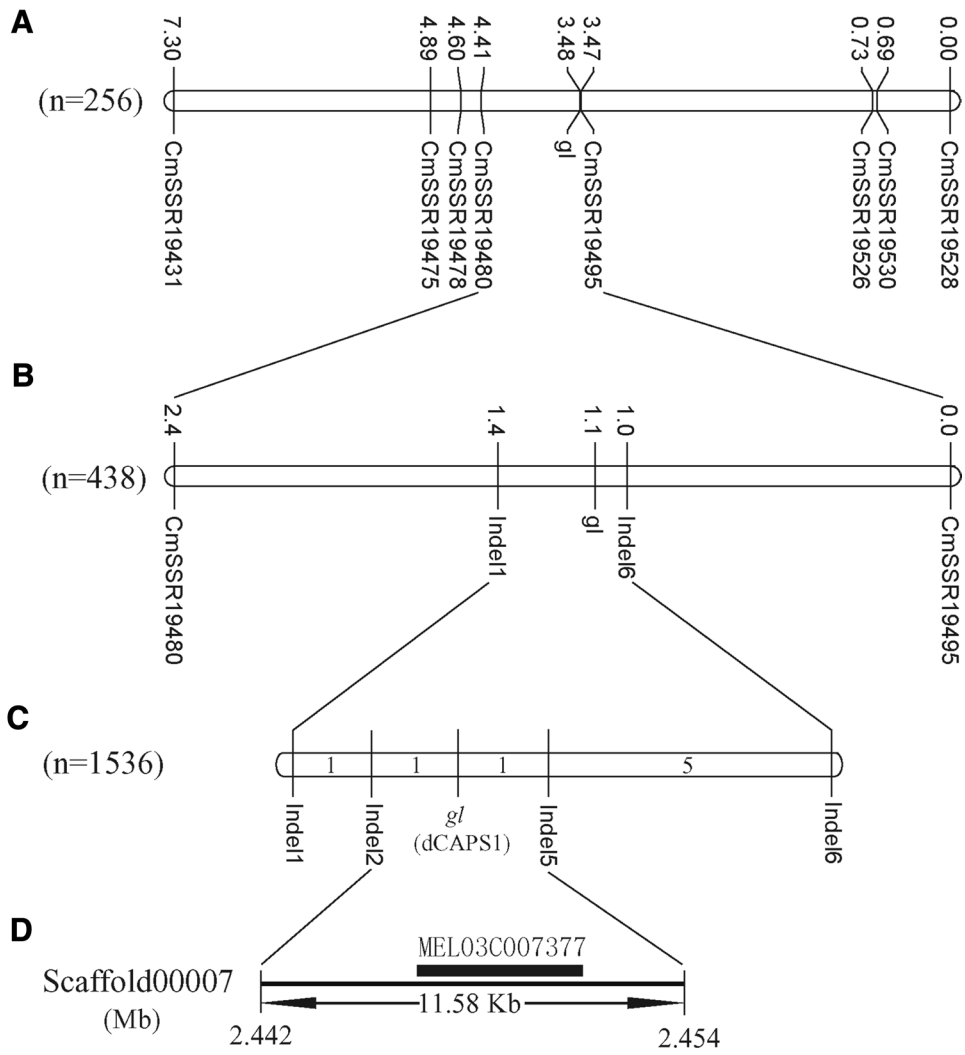

Fig. 3 Fine genetic mapping of *CmGL* gene in melon. **a** Preliminary mapping of *cmgl* gene using SSR markers in 256 F_2 plants. **b** Fine mapping of *cmgl* gene using flanking markers in 438 F_2 plants. **c** The *cmgl* gene was finally mapped between Indel2 and Indel5 marker in 1536 F_2 plants. **d** Only one gene was predicated in the candidate region

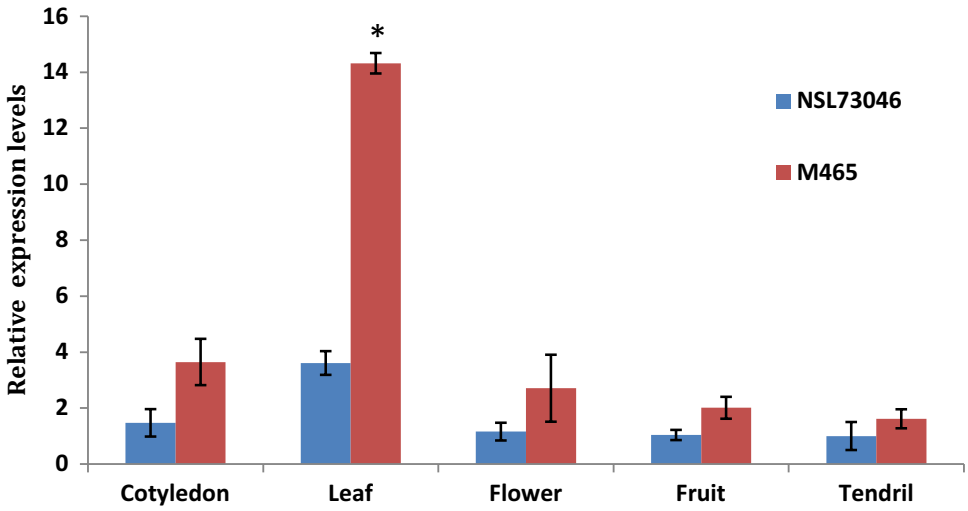

Fig. 4 The expression analysis of *CmGL* gene in the wild-type and mutant line using qRT-PCR in cotyledon, leaf, flower, fruit and tendril. Asterisk represent statistically significant difference ($P < 0.05$) between NSL73046 and M465

identity to cucumber HD-ZIP IV protein, CsGL3, with 97.6% identity in CDS level and 99.6% identity in amino acid level. *CsGL3* has been shown to regulate trichome initiation in cucumber (Cui et al. 2016; Pan et al. 2015; Wang et al. 2016). A phylogenetic tree was also constructed using CmGL and other 31 HD-ZIP IV proteins identified

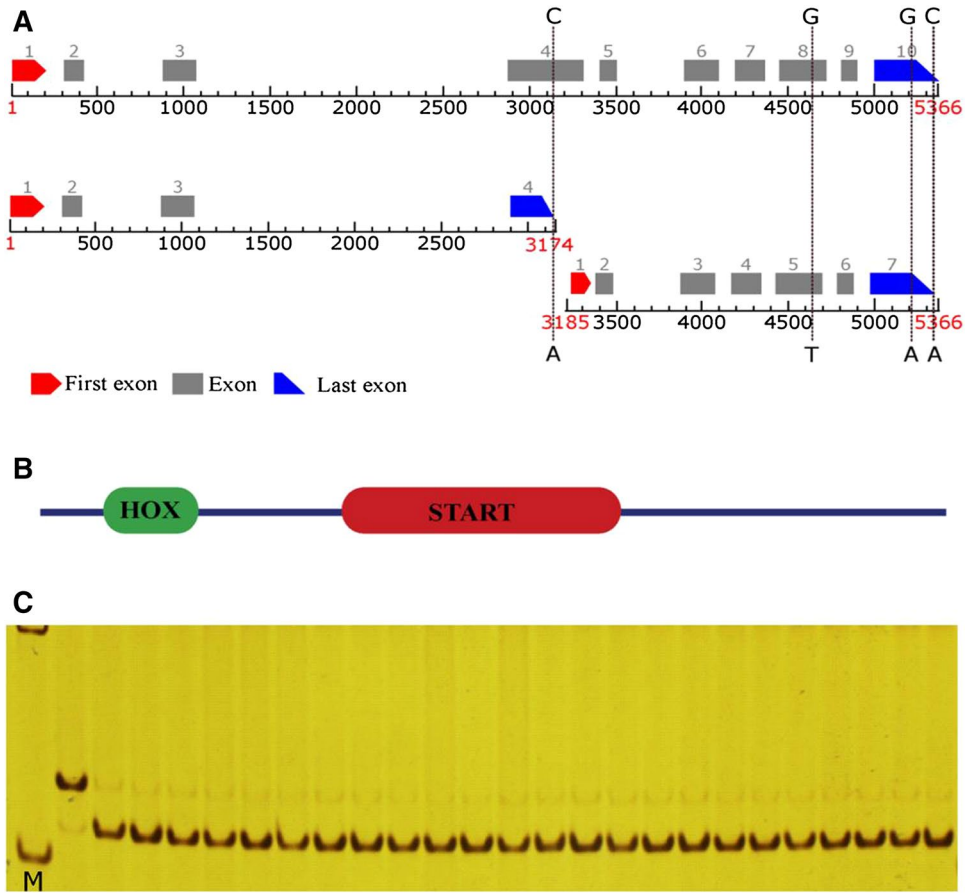

Fig. 5 Gene prediction, sequence comparison and allele diversity of *CmGL* gene in melon. **a** The gene prediction and sequence alignment of *CmGL* gene between the wild-type (upper) and mutant line (lower), and four SNPs were detected in the exons. A single base substitution in the fourth exon from C to A led to a premature termination codon in the mutant. **b** The *CmGL* gene was annotated to have two conserved domains in the deduced amino acid sequences. **c** Allelic diversity of *CmGL* gene in different melon germplasms using *CmGL* co-segregating marker dCAPS1. The first lane represents 100 bp DNA ladder, the second lane represents the mutant line and the remaining ones represent the natural melon germplasms

from cucumber, *Arabidopsis*, tomato, maize, rice, and cotton. The results revealed that CmGL showed a closer relationship to cucumber CsGL3, cotton GhHD1 and tomato WO (Supplementary Figure S1). These proteins are involved in trichome initiation and cell fate determination, suggesting CmGL in melon may have similar functions in regulating multicellular trichome development. Furthermore, 12 non-redundant genes encoding HD-ZIP IV proteins were identified in melon genome including the candidate gene *CmGL*, and they were used together with 11 HD-ZIP IV proteins identified in cucumber genomes (Fu et al. 2013) for multiple sequence alignments. A phylogenetic tree was constructed to unveil the relationship of HD-ZIP IV genes in cucumber and melon, and CmGL showed the closest relationship to CsGL3 suggesting that they were orthologous genes (Supplementary Figure S2).

To examine allelic diversity of the *CmGL* gene in melon natural populations, trichomes of 564 melon accessions (Supplementary Table S1) were observed in 2015 and 2016 field trials. All of them had trichomes as found in M465 wild-type parental line except for the mutant line NSL73046. The *cmgl* co-segregating marker dCAPS1 was used for genotyping all 564 melon accessions (Fig. 5c). The genotypic data at this marker locus were completely consistent with the phenotyping analysis, suggesting that the SNP within the *cmgl* gene was a causal mutation in NSL73046 and further confirmed it was the candidate gene.

Discussion

Plant trichomes play an important role in protecting plants from damage caused by herbivores and pathogens. Different studies have revealed that there is a close relationship between trichome and resistance against aphids and viruses in many species, such as potato (Gregory et al. 1986) and tomato (Kennedy and Sorenson 1985). In melon, leaf trichome density has been also shown to be associated with the rejection of *A. gossypii* Glover to settle (Sarria et al. 2010). A major QTL, *tric11* mapped on melon chromosome XI was identified for controlling the density of glandular trichomes, which explained between 23.8 and 58.7% of its phenotypic variation (Palomares-Rius et al. 2016). In this study, we identified a completely glabrous mutant in the melon accession NSL73046 and we show a region of ~ 12 kb harboring the *cmgl* locus was associated with this mutation (Fig. 3). Only one gene encoding a HD-ZIP IV transcription factor in this candidate region was identified and a dCAPS marker

was developed based on the SNP in the fourth exon showing co-segregation with glabrous phenotype in the F_2 plants and natural populations, further suggesting that it was the candidate gene for the *cmgl* locus (Fig. 5). This was the first report in melon to clone a gene for glabrous mutations, which was different from the *tricl1* major-effect QTL for trichome density reported earlier (Palomares-Rius et al. 2016). The glabrous phenotype will be useful in the identification of seed purity for melon hybrid in the seeding stage for its ease to differentiate the leaf and hypocotyl in the glabrous plants with normal plants.

Many transcription factors have been identified in regulating trichomes development in the model species *Arabidopsis* (Hulskamp 2004; Hulskamp et al. 1999; Schellmann and Hulskamp 2005; Ishida et al. 2008; Pesch and Hulskamp 2009; Schiefelbein 2003), and the molecular mechanisms involved in the initiation and development of trichomes have been extensively investigated (Kirik et al. 2004, 2005; Payne et al. 2000; Walker et al. 1999; Wester et al. 2009). Several studies in tobacco and tomato have revealed that there may be different regulation networks controlling the development of multicellular and unicellular trichomes (Payne et al. 1999; Yang et al. 2011). However, the molecular mechanism of trichomes development in melon is still largely unknown. Further functional analysis on *CmGL* gene will be helpful for better understanding the regulation of multicellular trichomes development in melon.

Melon and cucumber are two important crops in the genus *Cucumis*, (family Cucurbitaceae) that were diverged around 10 million years ago (Sebastian et al. 2010). The *CmGL* gene identified in this study shared a very high similarity to the cucumber *CsGL3* gene (Pan et al. 2015) with 97.6% nucleotide sequence identity and 99.6% amino acid sequence identity. The cucumber *csgl3* (Pan et al. 2015) and melon *cmgl* (this study) mutants both showed the completely glabrous phenotype suggesting they may perform similar functions. The loss-of-function mutation in melon *cmgl* was due mainly to a single base substitution from C to A which led to a premature termination codon and a truncated protein without the START domain, while that in the cucumber *csgl3* mutant was because of insertion of a 5005 bp LTR retrotransposon (LTR-RT) (Pan et al. 2015). We checked the physical locations of these two homologous genes and found that they were located on a highly conserved syntenic block between cucumber and melon genomes (Garcia-Mas et al. 2012; Li et al. 2011; Yang et al. 2014). It is interesting to note that the LTR-RT has a much higher fraction in the assembled melon draft genome than that in cucumber; the activity of LTR-RT is also much higher in the melon lineage after the divergence (Garcia-Mas et al. 2012). Using *CmGL* co-segregating dCAPS marker, we examined the allelic diversity at the *CmGL* locus in a diverse melon panel including the wild melon and two subspecies: subsp. *melo* and subsp. *agrestis*, there was no difference among these natural germplasms, suggesting the SNP within the *cmgl* gene may be the result of a natural mutation rather than due to domestication or human selection.

The plant-specific HD-Zip transcription factor gene family is characterized by a highly conserved 60–61 amino acid homeodomain (HD) (Ariel et al. 2007; Mayda et al. 1999). These proteins exhibit the singular combination of a HD domain with a leucine zipper acting as a dimerization motif. They can be classified into four subfamilies, HD-ZIp I–IV, according to a set of distinctive features that include DNA-binding specificities, gene structures, additional common motifs and physiological functions (Mukherjee et al. 2009). Several HD-ZIP transcription factors have been identified in regulating multicellular trichome development. For example, the *Wo* gene is essential for trichome formation in tomato, which encodes a HD-ZIP IV transcription factor, and mainly regulates the formation of type I trichomes (Yang et al. 2011). Another HD-ZIP IV transcription factor OCL4, regulating the differentiation of the anther cell wall, has been also identified in inhibiting trichome development in both maize and *Arabidopsis* (Vernoud et al. 2009). In cucumber, *CsGL1* and *CsGL3* encode a HD-ZIP I and HD-ZIP IV transcription factor, respectively, and they regulate the trichomes development in different manners. The *csgl1* mutant shows no visible trichomes on leaves, stems, tendrils, and floral organs except for the hypocotyl, and the role of *CsGL1* in trichome development in cucumber has not been reported in *Arabidopsis* (Chen et al. 2014; Zhao et al. 2015). The *csgl3* mutant shows a completely glabrous phenotype on the aerial parts of plants (Cui et al. 2016; Pan et al. 2015; Zhao et al. 2015). The *CmGL* gene identified in our study controlling the completely glabrous phenotype in melon was also encoded by a HD-ZIP IV transcription factor, and a phylogenetic tree constructed using all HD-ZIP IV proteins identified in cucumber and melon revealed that it was an orthologous gene of CsGL3 (Supplementary Figure S2). Thus, it seems that most HD-ZIP genes regulating trichome development are members of the HD-ZIP IV subfamily, which have been reported to generally regulate the processes related to epidermal cell differentiation, trichome formation, and anthocyanin synthesis (Nakamura et al. 2006). The START domain is often identified in the HD-ZIP IV subfamily proteins, but their functions in plants are still unknown; their conservation between plants and animals suggest that they may be involved in lipid/sterol signal transduction (Schrick et al. 2004). The loss-of-function in the melon *cmgl* mutant resulted from a truncated protein with a lost START domain, and the domain was also missing in the *tril* mutant in cucumber (Wang et al. 2016). In addition, the START domain was

also lost in *cmgl* locus in the melon glabrous mutant. The role of the START domain in trichome formation in higher plants remains to be elucidated in future studies.

Author contribution statement XS, PS, QH, XZ and XL performed phenotyping in F_2 plants and fine mapping. HZ, JH and SS contributed to data processing and phylogenetic analysis. QZ and JP contributed to SEM microscopic analysis. LY, HZ and YW wrote the manuscript. All authors reviewed and approved this manuscript.

Acknowledgements This work was supported by grants from the National Natural Science Foundation of China (No. 31601776), the Scientific Innovation Fund of Henan Agricultural University (KJCX2016A15), and the International Scientific and Technological Cooperation Project of Henan Province.

Compliance with ethical standards

Ethical standard The experiments comply with the current laws of the country in which we were performed.

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

Accession codes The genomic resequence datasets of two melon materials M465 and NSL73046 have been deposited into the NCBI sequence read archive (SRA) SRP119828 under accessions SRS2585930 and SRS2585931.

References

Argyris JM, Ruiz-Herrera A, Madriz-Masis P, Sanseverino W, Morata J, Pujol M, Ramos-Onsins SE, Garcia-Mas J (2015) Use of targeted SNP selection for an improved anchoring of the melon (*Cucumis melo* L.) scaffold genome assembly. BMC Genom 16:4

Ariel FD, Manavella PA, Dezar CA, Chan RL (2007) The true story of the HD-Zip family. Trends Plant Sci 12:419–426

Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari MA, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science 321:836–838

Boualem A, Troadec C, Camps C, Lemhemdi A, Morin H, Sari MA, Fraenkel-Zagouri R, Kovalski I, Dogimont C, Perl-Treves R, Bendahmane A (2015) A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. Science 350:688–691

Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, McGrath SD, Wendl MC, Zhang Q, Locke DP, Shi X, Fulton RS, Ley TJ, Wilson RK, Ding L, Mardis ER (2009) BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods 6:677–681

Chen CH, Liu ML, Jiang L, Liu XF, Zhao JY, Yan SS, Yang S, Ren HZ, Liu RY, Zhang XL (2014) Transcriptome profiling reveals roles of meristem regulators and polarity genes during fruit trichome development in cucumber (*Cucumis sativus* L.). J Exp Bot 65:4943–4958

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31:3497–3500

Cohen S, Itkin M, Yeselson Y, Tzuri G, Portnoy V, Harel-Baja R (2014) The *PH* gene determines fruit acidity and contributes to the evolution of sweet melons. Nat Commun 5:4026

Cui JY, Miao H, Ding LH, Wehner TC, Liu PN, Wang Y, Zhang SP, Gu XF (2016) A new glabrous gene (*csgl3*) identified in trichome development in cucumber (*Cucumis sativus* L.). PLoS One 11(2):e0148422

DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, Reich M, Winckler W, Getz G (2012) RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 28:1530–1532

Feder A, Burger J, Gao S, Lewinsohn E, Katzir N, Schaffer AA, Meir A, Davidovich-Rikanati R, Portnoy V, Gal-On A, Fei Z, Kashi Y, Tadmor Y (2015) A Kelch domain-containing F-box coding gene negatively regulates flavonoid accumulation in muskmelon. Plant Physiol 169:1714–1726

Foster RE (1963) Glabrous, a new seeding marker in muskmelon. J Hered 54:113–114

Fu R, Liu W, Li Q, Li J, Wang L, Ren Z (2013) Comprehensive analysis of the homeodomain-leucine zipper IV transcription factor family in *Cucumis sativus*. Genome 56:395–405

Garcia-Mas J, BenjakA Sanseverino W, Bourgeois M et al (2012) The genome of melon (*Cucumis melo* L.). Proc Natl Acad Sci USA 109:11872–11877

Gregory P, Tingey WM, Ave DA, Bouthyette PY (1986) Potato glandular trichomes: a physicochemical defense mechanism against insects. ACS Symp Ser Am Chem Soc 296:160–167

Hulskamp M (2004) Plant trichomes: a model for cell differentiation. Nat Rev Mol Cell Biol 5:471–480

Hulskamp M, Schnittger A, Folkers U (1999) Pattern formation and cell differentiation: trichomes in *Arabidopsis* as a genetic model system. Int Rev Cytol 186:147–178

Ishida T, Kurata T, Okada K, Wada T (2008) A genetic regulatory network in the development of trichomes and root hairs. Annu Rev Plant Biol 59:365–386

Kennedy GG, Sorenson CF (1985) Role of glandular trichomes in the resistance of *Lycopersicon hirsutum f. glabratum* to Colorado potato beetle (Coleoptera: Chrysomelidae). J Econ Entomol 78:547–551

Kirik V, Simon M, Wester K, Schiefelbein J, Hulskamp M (2004) *ENHANCER* of *TRY* and *CPC* 2 (*ETC2*) reveals redundancy in the region-specific control of trichome development of *Arabidopsis*. Plant Mol Biol 55:389–398

Kirik V, Lee MM, Wester K, Herrmann U, Zheng Z, Oppenheimer D, Schiefelbein J, Hulskamp M (2005) Functional diversification of MYB23 and *GL1* genes in trichome morphogenesis and initiation. Development 132:1477–1485

Kirkbride JH (1993) Biosystematic monograph of the genus *Cucumis* (Cucurbitaceae): botanical identification of cucumbers and melons. Parkway Publishers, Boone

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAM tools. Bioinformatics 25:2078–2079

Li D, Cuevas H, Yang L, Li Y, Garcia-Mas J, Zalapa J, Staub J, Luan F, Reddy U, He X, Gong Z, Weng Y (2011) Syntenic relationships between cucumber (*Cucumis sativus* L.) and melon (*C. melo* L.) chromosomes as revealed by comparative genetic mapping. BMC Genom 12:396

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A (2009) A transposon-induced

epigenetic change leads to sex determination in melon. Nature 461:1135–1237

Mayda E, Tornero P, Conejero V, Vera P (1999) A tomato homeobox gene (HD-zip) is involved in limiting the spread of programmed cell death. Plant J 20:591–600

Mukherjee K, Brocchieri L, Burglin TR (2009) A comprehensive classification and evolutionary analysis of plant homeobox genes. Mol Biol Evol 26:2775–2794

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

Nakamura M, Katsumata H, Abe M, Yabe N, Komeda Y, Yamamoto KT, Takahashi T (2006) Characterization of the class IV homeodomain-leucine zipper gene family in *Arabidopsis*. Plant Physiol 141:1363–1375

Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. Plant J 14:387–392

Palomares-Rius FJ, Yuste-Lisbona FJ, Viruel MA, Lopez-Sese AI, Gomez-Guillamon ML (2016) Inheritance and QTL mapping of glandular trichomes type I density in *Cucumis melo* L. Mol Breed 36:132

Pan YP, Bo KL, Cheng ZH, Weng YQ (2015) The loss-of-function *GLABROUS* 3 mutation in cucumber is due to LTR-retrotransposon insertion in a class IV HD-ZIP transcription factor gene *CsGL3* that is epistatic over *CsGL1*. BMC Plant Biol 15:302

Payne T, Clement J, Arnold D, Lloyd A (1999) Heterologous *myb* genes distinct from *GL1* enhance trichome production when overexpressed in *Nicotiana tabacum*. Development 126:671–682

Payne CT, Zhang F, Lloyd AM (2000) *GL3* encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1. Genetics 156:1349–1362

Pesch M, Hulskamp M (2009) One, two, three…models for trichome patterning in *Arabidopsis*? Curr Opin Plant Biol 12:587–592

Sarria E, Palomares-Rius FJ, Lopez-Sese AI, Heredia A, Gomez-Guillamon ML (2010) Role of leaf glandular trichomes of melon plants in deterrence of *Aphis gossypii* Glover. Plant Biol 12:503–511

Schellmann S, Hulskamp M (2005) Epidermal differentiation: trichomes in *Arabidopsis* as a model system. Int J Dev Biol 49:579–584

Schiefelbein J (2003) Cell-fate specification in the epidermis: a common patterning mechanism in the root and shoot. Curr Opin Plant Biol 6:74–78

Schrick K, Nguyen D, Karlowski WM, Mayer KF (2004) START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. Genome Biol 5:R41

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

Serna L, Martin C (2006) Trichomes: different regulatory networks lead to convergent structures. Trends Plant Sci 11:274–280

Shan CM, Shangguan XX, Zhao B, Zhang XF, Chao LM, Yang CQ (2014) Control of cotton fibre elongation by a homeodomain transcription factor GhHOX3. Nat Commun 5:5519

Sherman A, Eshed R, Harel-Beja R, Tzuri G, Portnoy V, Cohen S, Rubinstein M, Schaffer AA, Burger J, Katzir N, Ophir R (2013) Combining bulk segregation analysis and microarrays for mapping of the pH trait in melon. Theor Appl Genet 126:349–358

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Vernoud V, Laigle G, Rozier F, MeeleyRB Perez P, Rogowsky PM (2009) The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. Plant J 59:883–894

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, EschJJ Marks MD, Gray JC (1999) The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. Plant Cell 11:1337–1350

Wang YL, Nie JT, Chen HM, Guo CL, Pan J, He HL, Pan JS, Cai R (2016) Identification and mapping of *Tril*, a homeodomain-leucine zipper gene involved in multicellular trichome initiation in *Cucumis sativus*. Theor Appl Genet 129:305–316

Werker E (2000) Trichome diversity and development. Adv Bot Res 31:1–35

Wester K, Digiuni S, Geier F, Timmer J, Fleck C, Hulskamp M (2009) Functional diversity of R3 single-repeat genes in trichome development. Development 136:1487–1496

Yang CX, Li HX, Zhang JH, Luo ZD, Gong PJ, Zhang CJ, Li JH, Wang TT, Zhang YY, Lu YE, Ye ZB (2011) A regulatory gene induces trichome formation and embryo lethality in tomato. Proc Natl Acad Sci USA 108:11836–11841

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

Zhang BC, Tolstikov V, Turnbull C, Hicks LM, Fiehn O (2010) Divergent metabolome and proteome suggest functional independence of dual phloem transport systems in cucurbits. Proc Natl Acad Sci USA 107:13532–13537

Zhao JL, Pan JS, Guan Y, Zhang WW, Bie BB, Wang YL, He HL, Lian HL, Cai R (2015) Micro-trichome as a class I homeodomain-leucine zipper gene regulates multicellular trichome development in *Cucumis sativus*. J Integr Plant Biol 57:925–935

Zhu H, Guo L, Song P, Luan F, Hu J, Sun X, Yang L (2016) Development of genome-wide SSR markers in melon with their cross-species transferability analysis and utilization in genetic diversity study. Mol Breed 36:153