Theor Appl Genet (2013) 126:2187–2196
DOI 10.1007/s00122-013-2128-3

ORIGINAL PAPER

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

Yuhong Li · Changlong Wen · Yiqun Weng

Received: 27 January 2013 / Accepted: 8 May 2013 / Published online: 21 May 2013

Abstract In cucumber, *Cucumis sativus* L., the spine and skin colors are two important fruit quality traits for variety improvement. In this study, we investigated the inheritance of spine and mature fruit skin colors in F_2 and F_3 populations derived from a cross between two inbred lines WI7200 (black spine and orange fruit skin colors) and WI7201 (white spine and creamy fruit skin colors). We confirmed that a single, dominant gene, *B*, controlled both black spine color and orange mature fruit color. Initial framework mapping with microsatellite markers located the *B* locus in the distal region of the short arm of cucumber chromosome 4. Fine mapping was conducted with draft genome scaffold-assisted chromosome walking and stepwise increase of mapping population sizes, which allowed for the assignment of the *B* locus to a 50 kb genomic DNA region with two flanking markers that were 0.06 and 0.09 cM, respectively, from the *B* locus in a mapping population of 2,001 F_2 plants. Gene annotation of this 50 kb region identified six genes including one encoding for a R2R3-MYB transcription factor. Sequence alignment of the R2R3-MYB homologs between the two parent inbreds identified a 1 bp deletion in the third intron of this gene in WI 7201. A molecular marker based on this indel was co-segregating with the spine and fruit colors. Quantitative RT-PCR revealed higher level of expression of this R2R3-MYB gene in WI7200 than in WI7201 in both immature and mature fruits. This R2R3-MYB gene seems to be the best candidate gene for the *B* locus conditioning black spine and orange mature fruit colors of cultivated cucumber.

Communicated by I. Paran.

Y. Li and C. Wen contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2128-3) contains supplementary material, which is available to authorized users.

Y. Li
Horticulture College, Northwest A&F University, Yangling 712100, China

Y. Li · C. Wen · Y. Weng (✉)
Horticulture Department, University of Wisconsin, Madison, WI 53706, USA
e-mail: yiqun.weng@ars.usda.gov; weng4@wisc.edu

C. Wen
Beijing Vegetable Research Center and National Engineering Research Center for Vegetables, Beijing Academy of Agricultural and Forestry Sciences, Beijing 100097, China

Y. Weng
Vegetable Crops Research Unit, USDA-ARS, 1575 Linden Drive, Madison, WI 53706, USA

Introduction

Cucumber, *Cucumis sativus* L. ($2n = 2x = 14$), which belongs to the gourd family Cucurbitaceae, is an economically important vegetable crop worldwide. Its immature fruits are consumed fresh or cooked, or processed as pickles. Fruit-related characteristics are considered the most important traits for cucumber varietal improvement. As such, considerable attention has been placed on fruit quality. A number of simply inherited genes have been identified for skin texture, fruit color, spine characteristics, and internal quality (reviewed in Pierce and Wehner 1990). Several genes for fruit skin texture were found tightly linked which include heavy netting (*H*), smooth (*tu*), dull fruit skin (*D*), fruit ribbing (*Fr*), uniform immature fruit color (*u*), numerous spines (*ns*), small spines (*ss*), and

tender fruit (*te*) (Fanourakis and Simon 1987; Vakalounakis 1992; Walters et al. 2001; Yuan et al. 2008; Miao et al. 2011). Interestingly, specific allele combinations of these genes are characteristic of different market classes. For example, the European greenhouse cucumbers often have uniform immature fruit color (*u*), smooth (*tu*) and glossy (*d*) fruits with heavy fruit netting (*H*) and tender skin (*te*). Meanwhile, typical fresh market north China cucumbers have ribbed (*Fr*) and dull (*D*) fruits with non-netting (*h*), and American cucumbers are generally warty (*Tu*), mottled (*U*) with dull (*D*), thick skin (*Te*) but no ribs (*fr*).

The fruit spine is an important trait for cucumber fruit quality. Cucumber varieties differ in spine size (large vs. small), density (numerous vs. very few) and color (black or brown vs. white). Genetic studies indicated that black or brown spines are dominant to white and controlled by a single gene, *B* (Strong 1931; Tkachenko 1935; Fanourakis and Simon 1987; Vakalounakis 1992; Walters et al. 2001). Shanmugasundarum et al. (1971) found different segregation ratios among the F_2 and BC_1 progeny from crosses of cucumber line 9362 with PI 212233 and Pixie, and suggested two dominant genes, *B* and *B2* (Pierce and Wehner 1990) controlling spine color in cultivated cucumber. The spine color in the wild cucumber *C. sativus* var. *hardwickii* seems to be under control by a different genetic mechanism. Cowen and Helsel (1983) crossed white-spined cultivated cucumber lines and black-spined wild cucumber line LJ90430, and found that the black spine color was controlled by two genes (9 black:7 white in F_2), which was also observed in Walters et al. (2001) and our studies (unpublished data). The two genes in *C. sativus* var. *hardwickii* were designated *B3* and *B4*, respectively (Pierce and Wehner 1990). The allelic relationship of *B* with either *B3* or *B4* is unknown. Nevertheless, in a cross between white-spined cultivated cucumber with the wild cucumber, Pitchaimuthu et al. (2012) found brown spine color in the *C. s.* var. *hardwickii* line was controlled by a single dominant gene. Therefore, more work is needed to clarify the inheritance of spine color in the wild cucumber.

Like other cucurbit crops, cucumber exhibits a wide spectrum of fruit skin colors. The immature fruit color can vary from light green, green, dark green, yellow green to white, and the mature fruits (30–45 days after anthesis) can be from white, creamy, light green, green, dark green, yellow, orange to red. Hutchins (1940) was the first to study mature fruit color in cucumber who grouped the colors into four classes (red, orange, yellow, and cream) and suggested two genes, *R* and *C*, controlled these colors. According to Hutchins (1940), dominant alleles at both loci (*R_C_*) produced red fruit, dominance at the *R* locus (*R_cc*) produced orange fruit, dominance at the *C* locus (*rrC_*) produced yellow fruit, while homozygous recessive (*rrcc*) produced cream-colored fruit. Peterson and Pike (1992) studied the inheritance of green mature fruit color, and proposed two major genes (*R* and *Gn*) in which *R_ _ _* for orange, *rrgngn* for green, *rrGnGn* for cream, and *rrGngn* for an intermediate color between cream and green mature fruit colors. Both Hutchins (1940) and Peterson and Pike (1992) noted that the spine color was pleiotropic or tightly linked with the *R* locus. That is, fruits with black spines had red or orange mature fruit color in cultivated cucumbers. This association was also observed in several other studies (Fanourakis and Simon 1987; Vakalounakis 1992). However, in wild cucumber the association of black spine color and orange/red mature fruit color does not seem to be true (Cowen and Helsel 1983).

Compared with field crops, the genetic and genomic resources of cucumber are relatively limited, but the availability of cucumber draft genome sequences (Huang et al. 2009; Woycicki et al. 2011; Yang et al. 2012) makes it possible to develop molecular markers for a number of fruit quality-related genes like *H* (heavy netting), *Tu* (warty fruit), *D* (dull fruit skin), *Fr* (fruit ribbing), and *u* (uniform immature fruit color) (Yuan et al. 2008; Zhang et al. 2010; Miao et al. 2011). No molecular markers are available for spine and fruit colors, which are important for both cucumber varietal development and understanding the domestication of cucumber (Sebastian et al. 2010). In the present study, the inheritance of fruit spine color and mature fruit skin color in cucumber was analyzed using an F_2 population derived from WI7200 (black-spined, orange mature fruit) and WI7201 (white-spined, creamy mature fruit). Molecular markers closely linked to the spine or mature skin color gene were developed. Fine genetic mapping allowed for identification of a candidate gene for the *B* locus.

Materials and methods

Plant materials and mapping populations

Two cucumber inbred lines, WI7200 and WI7201 were used to develop segregating populations for spine color and mature color of fruits. WI7200 and WI7201 were selfed for four generations from two plant introduction lines PI 249561 and PI 308915, respectively. The fruits of WI7200 had coarse, black spines and orange mature fruit color, whereas WI7201 had fruits with large, coarse, white spines, and light green to creamy mature fruit color (Fig. 1).

Three mapping populations with increasing sizes, F2-S (small), F2-M (medium) and F2-L (large) from the WI7200 × WI7201 mating were used for linkage analysis of the black spine *B* locus with molecular markers, which contained 137 $F_{2:3}$ families, 543 F_2 and 2,001 F_2 plants, respectively (successively larger population included plants

Fig. 1 Spine and mature fruit skin colors of parental lines (**a**, **c**), F_1 (**b**) and F_2 plants in the greenhouse (**d**) or field (**e**) from the populations derived from WI72000 and WI7201. *Insets* in **a** and **c** are enlarged black and white fruit spines of 7200A and 7201A, respectively, under a dissecting microscope

from the previous smaller one). These populations were also employed in a previous study for fine mapping of the compact plant growth habit gene *cp* by Li et al. (2011a), but the F2-L population for fine mapping of the black spine locus *B* was larger in this study.

Phenotypic data collection

Two parents, F_1, F_2, and F_3 family plants were grown both in the Walnut Street Greenhouse of the University of Wisconsin-Madison and the field (Hancock Agriculture Research Station, Hancock, WI, USA) from 2009 fall greenhouse to 2012 summer field seasons. For each F_3 family, at least 15 individuals were planted and scored for segregation at the fruit spine color and mature fruit skin color loci. For each recombinant F_2 plant defined in an interval with two flanking markers, at least 30 F_2-derived F_3 plants in each family were scored to infer the genotype at the *B* locus in the F_2 plant. The colors of fruit spines (white or black) were taken in young, immature fruits (within 20 days after pollination). Mature fruit colors were visually scored as creamy or orange at 30–45 days after pollination in both the greenhouse and field trials.

Molecular marker development and mapping strategy

In our earlier study (Li et al. 2011a), we developed a cucumber linkage map using 46 plants from the same WI7200 × WI7201 F_2 population. Initial linkage analysis placed the *B* locus for the black spine in the distal region of

the short arm of chromosome 4 linked to the microsatellite (or SSR, simple sequence repeat) marker SSR14026. A whole genome scaffold-based chromosome walking strategy was taken to identify more closely linked markers with this gene. Draft genome scaffold assemblies from both the Gy14 (Version 1.0) (Yang et al. 2012) and 9930 (Version 2.0) (Li et al. 2011b) cucumber lines were used. In the target region of the scaffolds associated with SSR14026, new SSR markers were selected from a collection of 83,689 SSRs that were developed from the Gy14 draft genome assembly (Cavagnaro et al. 2010; Yang et al. 2012). If no suitable SSRs were available, new SSRs were developed from the target region using the SSR Locator program (available at http://minerva.ufpel.edu.br/~lmaia.faem/) (da Maia et al. 2008). The uniqueness of expected PCR products from these newly developed SSR primer pairs was verified through in silico PCR (Cavagnaro et al. 2010).

When SSR markers were exhausted in the target region, DNA fragments at selected intervals from WI7200 and WI7201 were sequenced. Single nucleotide polymorphisms (SNPs) and indels (insertion/deletions of DNA sequences) were identified in these sequences as new markers. Procedure for DNA sequencing and SNP or indel discovery followed Li et al. (2011a). For SNP genotyping, SNP-based PCR amplification of multiple specific alleles (PAMSA) (Gaudet et al. 2007), CAPS (cleaved amplified polymorphic sequence) (Neff et al. 1998) or derived CAPS (dCAPS) (Michaels and Amasino 1998) markers were developed. For PAMSA marker, three primers, a forward common primer and two reverse allele-specific primers were designed. Tails of different lengths, 5 bases for one primer and 15 for the other primer were added to each of the two allele-specific primers to obtain a difference of 10 bp between the amplification products (Gaudet et al. 2007). Primers for CAPS and dCAPS markers were designed with Primer Premier 5.0 (http://www.premierbiosoft.com/) and dCAPS Finder 2.0 (Neff et al. 2002), respectively.

All newly developed markers were first mapped with the F2-S population to identify co-segregating markers, which were then applied to F2-M population to find recombinant plants in a region defined by flanking markers. Finally, fine mapping was performed in the F2-L population with closely linked or co-segregating markers.

DNA exaction, PCR amplification of molecular markers and gel electrophoresis was conducted as described in Li et al. (2011a).

DNA sequencing, annotation and gene prediction, and candidate gene identification

To clone candidate gene sequences for the *B* locus from WI7200 and WI7201, sequencing primers were designed using Primer Premier 5.0 in the predicted gene region. Each primer pair was designed such that the expected amplicon size was from 600 to 900 bp, and neighboring primer pairs were designed to amplify PCR products with at least 100 bp overlap with preceding fragment. Gene prediction in target genomic DNA regions was performed with the computer program FGENESH (Salamov and Solovyev 2000) (http://sunl.softberry.com/) and checked manually. Candidate gene sequence alignment between two parental lines was performed using the software ClustalW2 (Larkin et al. 2007).

Quantitative reverse-transcription PCR (qPCR) analysis of candidate gene

The skins and associated spines of immature (3 days after pollination) and mature (30 days after pollination) fruits from both parents were peeled off with a knife and flash frozen in liquid nitrogen. There were two biological replications (2 fruits from 2 plants) for each parent. Total RNA was isolated with RNeasy Plant Mini Kit (Qiagen, Germany) following manufacturer's instructions. The first-strand cDNA synthesis was performed using RevertAid^{TM} First Strand cDNA Synthesis Kit (Fermentas, USA). All experiments were performed with two biological and three technical replicates. Primers of the R2R3-MYB candidate gene were designed with the Oligo 6 program (http://www.oligo.net/). The cucumber ubiquitin extension protein gene was used as the reference (Wan et al. 2010). Quantitative real-time PCR was performed using the SYBR Green PCR master mix (Applied Biosystems Inc., USA) in iCycler iQ^{TM} 5 Multicolor Real-Time PCR detection system (Bio-Rad, USA) following Xia et al. (2009) with three technical replications. Relative quantification was calculated according to Livak and Schmittgen (2001). The threshold cycle (C_t) value of ubiquitin extension protein was subtracted from that of R2R3-MYB to obtain a ΔC_t value. The C_t value of control sample WI7201 was subtracted from the ΔC_t value to obtain a $\Delta\Delta C_t$ value. The fold changes in expression level relative to WI7201 were expressed as $2^{-\Delta\Delta C_t}$.

Data analysis

Chi-square tests were performed for goodness-of-fit on phenotypic data of spine color and mature fruit color to test for deviations from the expected 1:2:1 segregation ratio in $F_{2:3}$ populations (3:1 for segregation in the F_2 populations). Linkage analysis of the *B* (black spine) and orange mature fruit color loci with molecular markers was performed with the Kosambi mapping function using JoinMap 3.0 with the threshold LOD score of 6.0.

Results

Inheritance of fruit spine color and mature fruit color

Spine colors, black in WI7200, and white in WI7201, could be scored easily on young developing fruits after pollination. The spines gradually fell off the fruit when the fruits grew toward maturity, which started approximately 3 weeks after pollination. The spine color and mature fruit skin color on the fruits of F_1 plants was black and orange, respectively, indicating the dominant nature of black to white spines, and orange to white mature fruit color (Fig. 1b).

In 2009 greenhouse season, among 137 F_2 plants from WI7200 × WI7201, 30 had fruits with white spines, and 107 with black spines ($P = 0.40$ for 3 black to 1 white segregation ratio). In 2010 field season, of the 126 available F_2-derived F_3 families, 30 were homozygous white (*bb*), 35 homozygous black (*BB*), and 61 heterozygous black (*Bb*) at the spine color locus ($P = 0.82$ in χ^2 test against 1:2:1 expected segregation ratio). Meanwhile, among the 406 F_2 plants from the same cross growing in the field, 115 had white-spined and 291 had black-spined fruits ($P = 0.12$ in χ^2 test against 3:1 segregation ration). For the 1,458 F_2 plants for fine genetic mapping, segregation at the *B* locus was 344 *bb* (white spine):1,114 *B_* (black spine) ($P = 0.22$ in χ^2 test against 3 black:1 white). These segregations supported earlier studies (Tkachenko 1935; Fanourakis and Simon 1987; Vakalounakis 1992) that black fruit spine color in cultivated cucumber is controlled by a single dominant gene *B*.

In both parental lines and segregating populations, mature fruit colors could largely be classified as creamy or orange. However, depending on developmental stage or environmental conditions, some variations were observed. The creamy category could be light green under greenhouse conditions, or white with light yellow hues in the field, especially when the fruits were exposed to direct sun for extended time. More variations were observed in the orange category which varied from yellow to orange, brown or red (Fig. 1). However, all fruits with black spines were classified into the orange category, and all white-spined fruits into the creamy category. No recombinants were observed suggesting black spine color and orange mature fruit skin color were controlled by the same gene (see "Discussion"). Following Pierce and Wehner (1990), we assigned the gene symbol *R* for the orange mature fruit color in WI7200, and *rr* for white/creamy fruit color in WI7201.

Chromosome location and molecular markers for the black spine locus *B*

In our previous study, a framework cucumber genetic map was developed using 46 WI7200 × WI7201 F_2 plants which was segregating for plant growth habit (compact vining type by gene *cp*) and spine color (black spine color controlled by gene *B*) loci (Li et al. 2011a). The map for chromosome 4 with 35 mapped loci is shown in Fig. 2a. Information of these 35 SSRs is provided in Table S1 (online materials). Initial linkage analysis indicated that the *B* locus was in the distal region of the short arm of chromosome 4 co-segregating with SSR07209 (Fig. 2a), which is physically located in the Gy14 scaffold00919 (Cavagnaro et al. 2010). We tested 144 additional SSR markers located in this scaffold and identified 15 polymorphic ones. Among the 15, four (UW084260, UW084262, UW084366, and UW084367, Table S2 of online materials) showed segregation distortion (paternal inheritance) in the F2-S mapping population; 11 were mapped with SSR07209 remaining the closest marker to the *B* locus on the resulting map (Fig. 2b).

To identify flanking markers for the *B* locus, new SSR markers were developed based on the cucumber genetic and physical maps across this region (Ren et al. 2009; Weng et al. 2010; Yang et al. 2012). Forty-eight primer pairs were designed from five scaffolds of the Gy14 draft genome assembly, which included scaffold00968, scaffold01936, scaffold04158, scaffold02511, and scaffold02901. Six polymorphic markers were identified and placed on the genetic map with 137 $F_{2:3}$ families. Among them, two (SSR84781 and SSR84783) were 2.8 and 3.3 cM away from the *B* locus, respectively; four (UW024693, UW049135, UW058682, and UW083447) were at the distal end of the long arm of chromosome 4 (Fig. 2b). Now the closest flanking markers were UW084781 and SSR07209 that were 2.8 and 0.4 cM away from the *B* locus, respectively. The names, associated scaffolds and scaffold locations of all markers on the map are listed in Table S2. The order of these mapped loci was consistent with their physical positions in the scaffolds.

Fine mapping of black fruit spine locus *B*

Because SSR07209 and UW084781 flanking the *B* locus belong to two different scaffolds in either the Gy14 (scaffold00919 and scaffold02901) or 9930 (scaffold000009_1 and scaffold000192) draft genome assembly (Table S2), the physical distance between the two markers remains unknown. However, since SSR07209 was only 0.4 cM away from the *B* locus, it was still possible that *B* is in the distal ~564 kb region of Gy14 scaffold00919. We decided to explore STS markers across this region. To increase the chance of finding SNPs or indels between the two parents, we hypothesized that the regions with high level of polymorphisms between the 9930 and Gy14 draft genomes could potentially also be more polymorphic between WI7200 and WI7201. Thus, we aligned genomic DNA

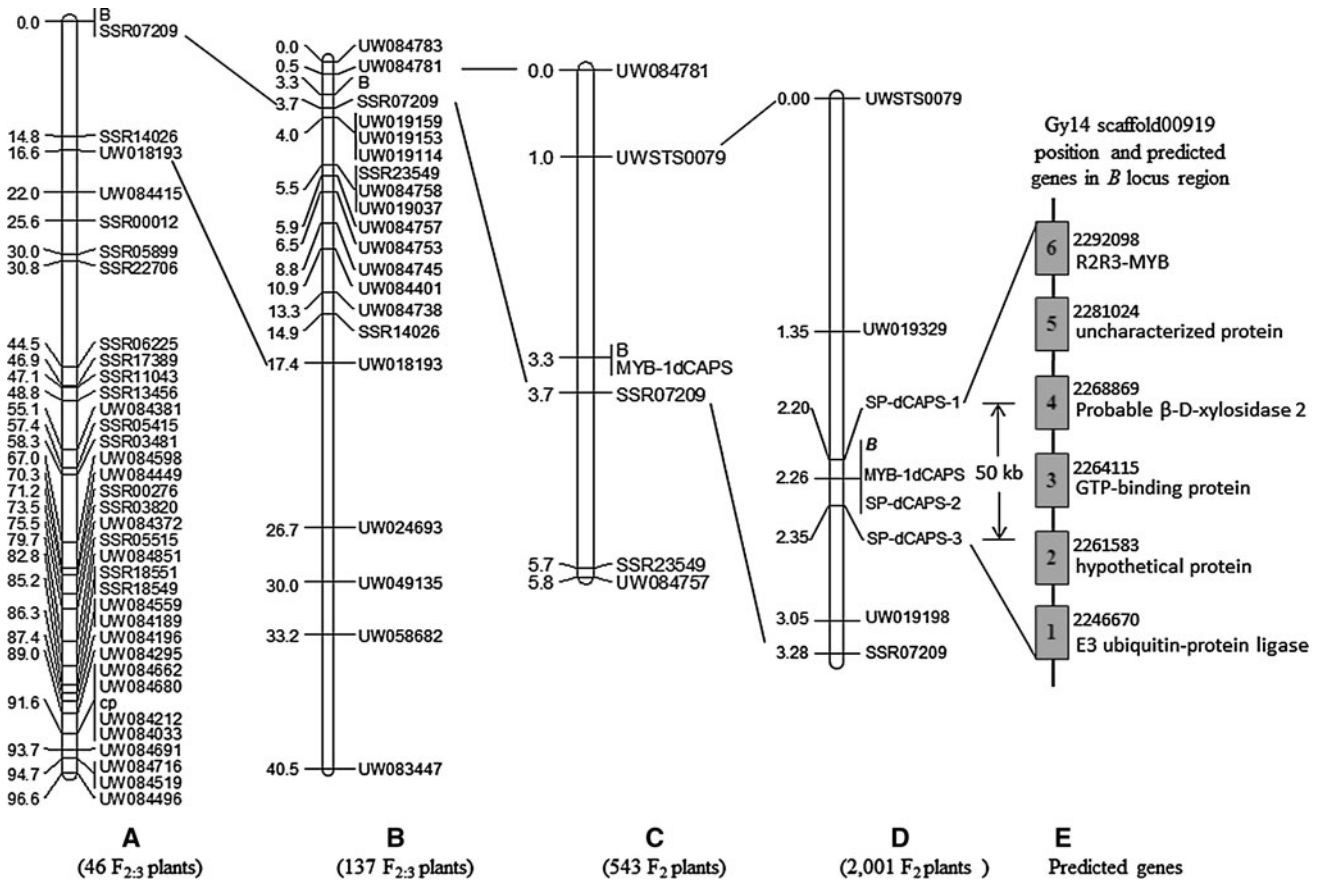

Fig. 2 Fine genetic-physical maps of black spine *B* locus. **a** A framework linkage map based on 46 F_2 plants from WI7200 × WI7201 and 35 SSR markers. The spine color locus (*B*) is placed in the distal region of the short arm of chromosome 4. **b** Chromosome location of *B* locus and molecular markers based on 137 $F_{2:3}$ families. *Vertical bars* delimit co-segregating marker loci. **c** A high-resolution linkage map based on 543 $F_{2:3}$ families. **d** The fine genetic map in the spine color gene (*B*) region based on 2,001 F_2 plants. **e** Six genes annotated in a 50 kb regions defined by two flanking markers of the *B* locus. The sixth R2R3-MYB is the most possible candidate gene

sequences in the ~564 kb region between 9930 scaffold000009_1 and Gy14 scaffold00919, and identified regions with a minimum of six SNPs per 1,000 bp. We also annotated this 564 kb region. Gene predictions identified a R2R3-MYB transcription factor, which is known to be important in controlling fruit coloring in many fruit crops (see "Discussion"). STS primers were designed in 25 selected DNA fragments including this R2R3-MYB gene region. PCR products from WI7200 and WI7201 were sequenced; 34 SNPs and two indels in four DNA fragments were identified between the two parents. Four CAPS, dCAPS, PAMSA, or indel markers (one from each fragment) were designed and run in the segregating population (F2-M). The SNP marker MYB-1dCAPS developed within the R2R3-MYB gene region co-segregated with the black spine color locus (*B*) in 543 F_2 plants (F2-M mapping populations) (Fig. 2c). In addition, two markers, UWSTS0079 and SSR07209, flanked the *B* locus in a 440 kb genomic DNA region (Table S2).

A new set of 1,458 F_2 plants were screened with SSR07209 and UWSTS0079, from which 93 recombinants across this interval were identified. The marker MYB-1dCAPS was used to genotype the 93 recombinant plants. While MYB-1dCAPS was still co-segregating with the *B* locus in this fine mapping population, there were 64 and 29 recombinants in the two intervals defined by the marker pairs UWSTS0079–MYB-1dCAPS and MYB-1dCAPS–SSR07209, respectively.

To further narrow down the region carrying the *B* locus, additional SSR markers from the 440 kb region between UWSTS0079 and SSR07209 were identified. Among 54 SSRs tested, 2 (UW019329 and UW019198) were mapped in this region, which was 0.91 and 0.79 cM, respectively, from the *B* locus, thus assigning *B* to a 237 kb interval (Fig. 2d; Table S2). Next, we designed STS primers to amplify and sequence 20 DNA fragments in this 237 kb region, which allowed for identification of SNPs in four fragments between WI7200 and WI7201. Four SNP-

derived CAPS or dCAPS markers were developed, and three (SP-dCAPS-1, SP-dCAPS-2 and SP-dCAPS-3) were successfully mapped (the marker UWSTS00067 showed paternal inheritance). Linkage analysis with 2,001 F_2 plants indicated that both MYB-1dCAPS and SP-dCAPS-2 co-segregated with the *B* locus, whereas SP-dCAPS-1 and SP-dCAPS-3 flanked the *B* locus at 0.06 and 0.09 cM away, respectively, which is approximately 50 kb in the Gy14 scaffold00919 (Fig. 2d, e; Table S2).

Genes predicted in the black spine/orange mature fruit color region

Using the FGENESH program, in the 237 kb genomic region delimited by UW019329 and UW019198, 39 genes were predicted, and six were in the 50 kb region defined by SP-dCAPS-1 and SP-dCAPS-3 (Fig. 2d, e). The predicted functions and associated information of all 39 genes are presented in Table S3 (online materials). Among the six genes in the 50 kb region, the gene encoding a R2R3-MYB transcription factor seemed to be a good candidate for the *B* gene because of co-segregation. We cloned the R2R3-MYB gene homologs from both WI7200 and WI7201, which were 2,612 and 2,611 bp, respectively. Gene annotation suggested three exons and four introns in this gene. Alignment of the R2R3-MYB candidate gene sequences revealed only one base pair difference between the two parental lines: a 1-bp (T) insertion in the third intron in WI7200 (Fig. S1 of online materials).

Expression level of MYB candidate gene in immature and mature fruits

Primer sequences for both reference and MYB candidate genes used in qPCR analysis are provided in Table S4 (online materials). The expression levels of the R2R3-MYB candidate gene in the skins of immature and matures fruits of both parents were measured (Fig. 3), and expression of this candidate gene was 1.7 and 2.5 times as high in WI 7200 (black spine) as compared to WI 7201 (white spine) (significantly higher, $P < 0.05$ in *t* tests).

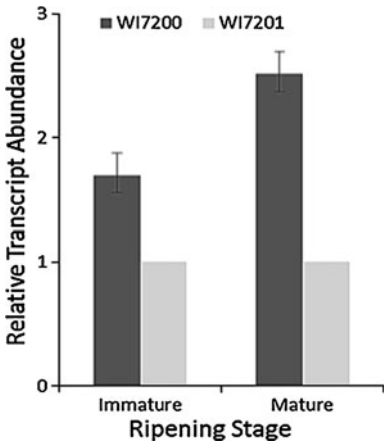

Fig. 3 Expression of R2R3-MYB candidate gene in immature and mature fruits of WI7200 and WI7201 from qPCR analysis. Data are the means of three replicates (±SD). Relative transcript abundance used WI7201 as the reference

Discussion

Inheritance of spine and mature fruit skin color in cultivated cucumber

Results from this study support a dominant gene, *B* controlling black spine in cultivated cucumber, which is consistent with previous studies (Strong 1931; Tkachenko 1935; Vakalounakis 1992; Walters et al. 2001). Among the thousands of plants observed in our segregating populations, all black-spined fruits had yellow, orange or red mature fruit colors, and all white-spined fruits had light green or creamy mature fruit colors. This association is consistent with earlier reports (e.g., Tkachenko 1935; Fanourakis and Simon 1987; Vakalounakis 1992). However, our results do not support the proposal by Hutchins (1940) that two genes, *R* and *C*, control the orange or red mature fruit color in cucumber. The discrepancy in segregation ratios of mature fruit colors between the present study and Hutchins (1940) may be due to different lines employed. Hutchins (1940) scored four mature fruit skin colors as 9 red:3 orange:3 yellow:1 cream in an F_2 population. In our study, all fruits were classified into two categories: orange which may include red, brown, orange, yellow, and light yellow, and creamy which may include green, light green, cream, and white (Fig. 1). Noticeably, the fruits on the same plant may exhibit different colors within the same category depending on the development stage, or node positions of the fruit. The variations in color development may also be due to environmental factors, such as light and temperature during fruit development, which have been observed in many other fruit crops (e.g., Piero et al. 2005; Ubi et al. 2006; Feng et al. 2010; Azuma et al. 2012).

Cucumber spines are multicellular, non-glandular trichomes developed from epidermal cells of the fruit skin (Fig. 1a, c, insets), which are similar in shape and structure to leaf trichomes (Guan 2008). This is also evidenced by the fact that no trichomes (spines) could be seen on either the foliage or the fruit surface in the *glabrous* mutant of cucumber (Yang et al. 2011). Considering the organogenesis of spine and the association of black spine color with orange mature fruit color, it is reasonable to conclude that the mature fruit skin color and spine color in cultivated cucumber may be controlled by the *B* gene.

In cucumber, four genes have been identified that are related with fruit spines: *B* for black spine, *ns* for numerous

spines, *ss* for small spines, and *gl* (*glabrous*) for trichomes formation (Pierce and Wehner 1990). Gene *gl-2* controlling presence or absence of trichomes on the foliage or fruit was mapped in chromosome 2 (Yang et al. 2011). The two genes *ss* and *ns* were located in a cluster of genes (*u*, *D*, *Fr*, *te*, *Tu*) in chromosome 5 controlling fruit epidermal features (Fanourakis and Simon 1987; Vakalounakis 1992; Walters et al. 2001; Yuan et al. 2008; Zhang et al. 2010; Miao et al. 2011). In this study, we placed the black spine locus *B* on the short arm of chromosome 4 (Fig. 1). Therefore, genes controlling spine characteristics in cucumber are clearly located in different chromosomes.

Genetics mapping of the *B* locus

By stepwise increase of the mapping population sizes and taking advantage of the Gy14 and 9930 draft genome assemblies, we were able to identify a candidate gene for the *B*/*R* locus. Nevertheless, map-based cloning in the cucumber genome remains non-trivial. In the 564 kb region of Gy14 scaffold00919 between SSR07209 and the distal end (Fig. 1), among 160 SSR and 45 STS (~30 kb DNA sequences) markers screened, only two SSR markers (UW019198 and UW019329) and five SNPs-derived markers were successfully mapped. While cultivated cucumber has in general low genetic diversity (Weng et al. 2010; Li et al. 2011a), one important reason for the low success rate in the present study was the ~440 kb region between SSR07209 and UWSTS0079 in which a number of markers showed paternal inheritance in the mapping population (Table S2). Segregation distortion is often encountered in genetic mapping studies in cucumber (e.g., Yuan et al. 2008; Weng et al. 2010; Miao et al. 2011), but complete paternal inheritance of markers as found in the present study is not common. In *Cucumis* species, the mitochondrial genome exhibits paternal inheritance (Havey et al. 1998). There are intracellular transfers of DNA sequences between the nuclear and mitochondrial genomes (Alverson et al. 2011). However, when the 440 kb sequences were BLASTed against the cucumber mitochondrial genome assembly (Alverson et al. 2011), no significant hits were found. The paternal inheritance of markers in this region may be genotype specific, because molecular markers developed from this region (scaffold00919) segregated normally in other populations we have worked with (for example, Yang et al. 2012). The reason for this phenomenon may need further investigation.

R2R3-MYB as the candidate gene for black spine or orange mature skin color in cucumber

Anthocyanins are water-soluble flavonoids that are responsible for the red, blue, and purple colors in plants (Tanaka et al. 2008). In almost all plant species studied so far, structural genes of the flavonoid biosynthetic pathway are largely regulated at the level of transcription through a complex of MYB transcription factors, basic helix-loop-helix (bHLH) TFs and WD repeat proteins (reviewed by Hichri et al. 2011). Most MYB transcription factors regulating the flavonoid biosynthetic pathway belong to the R2R3-MYB type which comprises one of the largest families of transcription factors in plants (Feller et al. 2011). In the cucumber genome, there are approximately 55 members in the R2R3-MYB family (Li et al. 2012). Most genes cloned so far that are responsible for fruit color changes belong to the R2R3-MYB family (Hichri et al. 2011). For example in grapevine (*Vitis vinifera*), a deletion in the coding region of the *VvmybA1* gene was responsible for the change of fruit color from black to white (Yhaushiji et al. 2006); moreover, a retrotransposon insertion in the promoter region of the same gene changed skin color from red to white (Kobayashi et al. 2004). In apple (*Malus domestica*), a 23-bp repeat motif in the promoter of the *MdMYB10* gene is found only in red-fleshed apples (Espley et al. 2009); meanwhile, methylation levels of the *MYB10* gene promoter region correlates with the peel phenotypes (Telias et al. 2011).

In the present study, fine genetic mapping identified a R2R3-MYB transcription factor that seems to be the best candidate gene for both the black spine color and orange mature fruit color. The expression of black spine color occurs very early in fruit development, while skin color develops progressively from green or light green to light yellow, yellow, orange or red. The development of mature fruit color in WI7200 is likely due to biosynthesis of anthocyanin (Weng et al., unpublished data). How the same R2R3-MYB gene regulates the biosynthesis of anthocyanin for color expression in the spine and fruit skin requires further investigation. No changes except a single nucleotide insertion in the third intron of WI7200 were found at the DNA sequence level within the candidate gene between the two parental lines (Fig. S1 of online materials). However, preliminary qPCR analysis indeed revealed significant differences between the two parental lines in expression of this candidate gene at both immature and mature fruit stages. The transcription level of this R2R3-MYB candidate gene in WI7200 (black spine, orange fruit color) was significantly higher than that in WI7201 (white spine, creamy/light green skin color) (Fig. 2). Nucleotide polymorphism or the methylation level in the promoter region of this candidate gene needs to be examined. Since there are six genes in the 50 kb region harboring the *B* locus (Fig. 2e), only functional validation by transformation will provide the definite evidence to prove the identity of this candidate gene.

Acknowledgments The authors thank Linda Crubaugh for technical help, and Michael J. Havey for critical reading of the manuscript. This

research was supported by a Grant from USDA-SCRI (Project #2011-51181-30661) to YW. YL's work was partly supported by the National Natural Science Foundation of China (Project #31171955). CW's research was sponsored by the China Scholarship Council.

References

Alverson AJ, Rice DW, Dickinson S, Barry K, Palmer JD (2011) Origins and recombination of the bacterial-sized multi-chromosomal mitochondrial genome of cucumber (*Cucumis sativus*). Plant Cell 23:2499–2513

Azuma A, Yakushiji H, Koshita Y, Kobayashi S (2012) Flavonoid biosynthesis-related genes in grape skin are differentially regulated by temperature and light conditions. Planta 236:1067–1080

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

Cowen NM, Helsel DB (1983) Inheritance of two genes for spine color and linkages in a cucumber cross. J Hered 74:308–310

da Maia LC, Palmieri DA, de Souza VQ, Kopp MM, de Carvalho FIF, Oliveira AC (2008) SSR Locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. Int J Plant Genomics: 412696

Espley RV, Brendolise C, Chagné D, Kutty-Amma S, Green S, Volz R, Putterill J, Schouten HJ, Gardiner SE, Hellens RP, Allan AC (2009) Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. Plant Cell 21:168–183

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

Feller A, Machemer K, Braun EL, Grotewold E (2011) Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J 64:94–116

Feng SQ, Wang YL, Yang S, Xu YT, Chen XS (2010) Anthocyanin biosynthesis in pears is regulated by a R2R3-MYB transcription factor PyMYB10. Planta 232:245–255

Gaudet M, Fara AG, Sabatti M, Kuztninsky E, Mugnozza GS (2007) Single-reaction for SNP genotyping on agarose gel by allele-specific PCR in black poplar (*Populus nigra* L.). Plant Mol Biol Report 25:1–9

Guan Y (2008) Mapping and cloning of related gene for fruit spines formation in cucumber. Ph.D. thesis of Shanghai Jiaotong University (in Chinese with English abstract)

Havey MJ, McCreight JD, Rhodes B, Taurick G (1998) Differential transmission of the *Cucumis* organellar genomes. Theor Appl Genet 97:122–128

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V (2011) Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. J Exp Bot 62:2465–2483

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

Hutchins AE (1940) Inheritance in the cucumber. J Agric Res 60:117–128

Kobayashi S, Goto-Yamamoto N, Hirochika H (2004) Retrotransposon-induced mutations in grape skin color. Science 304:982

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Li YH, Yang LM, Pathak M, Li DW, He XM, Weng Y (2011a) Fine genetic mapping of *cp*: a recessive gene for compact (dwarf) plant architecture in cucumber, *Cucumis sativus* L. Theor Appl Genet 123:973–983

Li Z, Zhang ZH, Yan PC, Huang SW, Fei ZJ, Lin K (2011b) RNA-Seq improves annotation of protein-coding genes in the cucumber genome. BMC Genomics 12:540

Li Q, Zhang CJ, Li J, Wang LN, Ren ZH (2012) Genome-wide identification and characterization of R2R3MYB family in *Cucumis sativus*. PLoS ONE 7:e47576

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402–408

Miao H, Zhang SP, Wang XW, Zhang ZH, Li M, Mu SQ, Cheng ZC, Zhang RW, Huang SW, Xie BY, Fang ZY, Zhang ZX, Weng Y, Gu XF (2011) A linkage map of cultivated cucumber (*Cucumis sativus* L.) with 248 microsatellite marker loci and seven genes for horticulturally important traits. Euphytica 172:167–176

Michaels SD, Amasino RM (1998) A robust method for the detecting single-nucleotide changes as polymorphic markers by PCR. Plant J 14:381–385

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

Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends Genet 18:613–615

Peterson GC, Pike LM (1992) Inheritance of green mature seed-stage fruit color in *Cucumis sativus* L. J Am Soc Hortic Sci 117:643–645

Pierce LK, Wehner TC (1990) Review of genes and linkage groups in cucumber. HortScience 25:605–615

Piero ARL, Puglisi I, Rapisarda P, Petrone G (2005) Anthocyanins accumulation and related gene expression in red orange fruit induced by low temperature storage. J Agric Food Chem 53:9083–9088

Pitchaimuthu M, Dutta OP, Swamy KRM, Souravi K (2012) Mode of inheritance of bitterness and spine color in cucumber fruits (*Cucumis sativus* L.). In: Sari N, Solmaz I, Aras V (eds) Cucurbitaceae 2012, Proceedings of the 10th EUCARPIA meeting on genetics and breeding of Cucurbitaceae, Antalya, Turkey, 15–18 October 2012, pp 70–73

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

Salamov AA, Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. Genome Res 10:516–522

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

Shanmugasundarum SP, Williams PH, Peterson CE (1971) Inheritance of fruit spine color in cucumber. HortScience 6:213–214

Strong WJ (1931) Breeding experiments with the cucumber (*Cucumis sativus* L.). Sci Agric 11:333–346

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant J 54:733–749

Telias A, Lin-Wang K, Stevenson DE, Cooney JM, Hellens RP, Allan AC, Hoover EE, Bradeen JM (2011) Apple skin patterning is associated with differential expression of MYB10. BMC Plant Biol 11:93

Tkachenko NN (1935) Preliminary results of a genetic investigation of the cucumber, *Cucumis sativus* L. Bull Appl Plant Breed (Ser 2) 9:311–356

Ubi BE, Honda C, Bessho H, Kondo S, Wada M, Kobayashi S, Moriguchi T (2006) Expression analysis of anthocyanin

biosynthetic genes in apple skin: effect of UV-B and temperature. Plant Sci 170:571–578

Vakalounakis DJ (1992) Heart leaf, a recessive leaf shape markers in cucumber: linkage with disease resistance and other traits. J Hered 83:217–220

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

Wan H, Zhao Z, Qian C, Sui Y, Malik A, Chen J (2010) Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. Anal Biochem 399:257–261

Weng Y, Johnson S, Staub JE, Huang S (2010) An extended microsatellite genetic map of cucumber, *Cucumis sativus* L. HortScience 45:880–886

Woycicki R, Witkowicz J, Gawronski P et al (2011) The genome sequence of the North-European cucumber (*Cucumis sativus* L.) unravels evolutionary adaptation mechanisms in plants. PLoS ONE 6:e22728

Xia XJ, Wang YJ, Zhou YH, Tao Y, Mao WH, Shi K, Asami T, Chen Z, Yu JQ (2009) Reactive oxygen species are involved in brassinosteroid-induced stress tolerance in cucumber. Plant Physiol 150:801–814

Yang SJ, Miao H, Zhang SP, Cheng ZC, Zhou J, Dong SY, Wehner TC, Gu XF (2011) Genetic analysis and mapping of *gl-2* gene in cucumber (*Cucumis sativus* L.). Acta Hortic Sin 38:1685–1692

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

Yhaushiji H, Kobayashi S, Goto-Yamamoto N, Jeong ST et al (2006) A skin color mutation of grapevine from black-skinned Pinot Noir to white skinned Pinot Blanc, is cause by deletion of the functional *VvmybA1* allele. Biosci Biotechnol Biochem 70:1506–1508

Yuan XJ, Pan JS, Cai R, Guan Y, Liu LZ et al (2008) Genetic mapping and QTL analysis of fruit and flower related traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Euphytica 164:473–491

Zhang WW, He HL, Guan Y, Du H, Yuan LH, Li Z, Yao DQ, Pan JS, Cai R (2010) Identification and mapping of molecular markers linked to the tuberculate fruit gene in the cucumber (*Cucumis sativus* L.). Theor Appl Genet 91:53–61