Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-019-03468-0

ORIGINAL ARTICLE

An irregularly striped rind mutant reveals new insight into the function of PG1β in cucumber (*Cucumis sativus* L.)

Mengfei Song[1] · Mengru Zhang[1] · Feng Cheng[1] · Qingzhen Wei[1] · Jing Wang[1] · Marzieh Davoudi[1] · Jinfeng Chen[1] · Qunfeng Lou[1]

Received: 9 March 2019 / Accepted: 24 October 2019

Abstract

Key message **Via bulked segregant analysis sequencing combined with linkage mapping, the *ist* gene responsible for the irregularly striped rind mutation was delimited to a 144-kb region in cucumber. Sequencing and expression analysis identified *Csa1G005490* as the candidate gene.**

Abstract The rind appearance of cucumber is one of the most important commercial quality traits. Usually, an immature cucumber fruit has a uniform rind that varies from green to yellow to white among different cultivated varieties. In the present paper, we isolated a novel fruit appearance cucumber mutant, *ist*, that has an irregularly striped rind pattern. The mutant displayed green irregular stripes on a yellow-green background at the immature fruit stage. Genetic analysis revealed that a single recessive gene, *ist*, is responsible for this mutation. A BSA (bulked segregant analysis) sequencing approach combined with genetic mapping delimited the *ist* locus to an interval with a length of 144 kb, and 21 predicted genes were annotated in the region. Based on mutation site screening and expression analysis, two single-nucleotide polymorphisms within the candidate gene, *Csa1G005490*, were identified as constituting the mutation. *Csa1G005490* encodes a polygalacturonase-1 noncatalytic subunit beta protein (PG1β) known to be involved in fruit softening. The expression of *Csa1G005490* was significantly lower in the *ist* mutant than in the wild type. Transcriptome analysis identified 1796 differentially expressed genes (DEGs) between the *ist* mutant and wild type. Gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses indicated that these DEGs were enriched mostly in photosynthesis and chlorophyll metabolism pathways. Decreased expression patterns of several chlorophyll synthesis genes in the mutant suggest that *ist* plays a key role in chlorophyll biosynthesis. These results will provide new insight into the molecular mechanism underlying rind appearance polymorphisms in cucumber.

Introduction

Rind appearance is an important determinant of the commercial value of cucumber, which has rich genetic diversity. In recent years, numerous genetic studies have focused on rind appearance traits of cucumber, including fruit color and skin

Communicated by Albrecht E. Melchinger.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03468-0) contains supplementary material, which is available to authorized users.

✉ Jinfeng Chen
jfchen@njau.edu.cn

✉ Qunfeng Lou
qflou@njau.edu.cn

Mengfei Song
2017204016@njau.edu.cn

Mengru Zhang
2017104068@njau.edu.cn

Feng Cheng
2018204019@njau.edu.cn

Qingzhen Wei
wqz0619@163.com

Jing Wang
2013104044@njau.edu.cn

Marzieh Davoudi
2018204052@njau.edu.cn

[1] State Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Weigang Street No. 1, Nanjing 210095, China

Published online: 16 November 2019

texture (Xie and Wehner 2001; Weng and Wehner 2018). In terms of its diversity, cucumber fruit color can be dark green, light green, green, yellow-green or white (Lun et al. 2015; Zhou et al. 2015; Liu et al. 2016; Hao et al. 2018). In addition, the skin texture can be dull or glossy, netted or smooth, ribbed or nonribbing, and mottled or uniform (Miao et al. 2011; Wang et al. 2014; Yang et al. 2014a, b). Although several genetic and regulatory mechanisms involved in rind appearance have been identified, information on irregularly striped rinds is still not available in cucumber.

The striped rind pattern among the typical characteristics of cucurbit vegetable fruits, especially in melon, watermelon and pumpkin (Paris 2003; Gama et al. 2015; Lv et al. 2018). The irregularly striped rind of the pumpkin is regulated by a series of genes (*broad striped l-I*^{Bst}, *narrow striped l-I*st, *irregular striped l-I*^{ist}), and these genes have obvious genetic relationships with the dark green fruit-controlling gene *L-I* and the light green fruit-controlling gene *l-I* (*L-I* > *l-I*^{Bst} > *l-I*st > *l-I*^{ist} > *l-I*) (Paris2003). This multigene regulatory pattern of the irregularly striped rind trait also exists in watermelon: The inheritance of fruit rind pattern is controlled by *G* (medium or dark solid green), g^W (wide stripe), g^M (medium stripe), g^N (narrow stripe) and *g* (solid light green or gray). The dominance of this series is as follows: $G > g^W > g^M > g^N > g$. In addition, the *csm* gene has been proposed to control clear stripe margins and is recessive to the gene controlling blurry stripe margins (*Csm*) (Lou and Wehner 2016). Well-defined stripe patterns are controlled by a dominant gene, and two molecular markers have been identified as being linked to the stripe pattern in watermelon (Gama et al. 2015). In melon, there are two different genetic patterns of rind stripes. Melon rind stripes were first thought to be controlled by the recessive gene *st* (Danin-Poleg et al. 2002). However, other studies confirmed that the gene controlling the striped pattern, named *CmSp-1*, is dominant over the gene controlling the nonstriped pattern (Lv et al. 2018). The stripped rind trait in melon is also controlled by a single dominant gene (*st3*), and the location of *st3* was narrowed to a 172.8-kb region on chromosome 4 (Liu et al. 2019). This different genetic pattern may be caused by differences in materials. Striped rind traits are also found in other horticultural crop species (Choudhury 1977; Telias et al. 2011). The red stripe trait of apple peels is mainly caused by the upregulation of anthocyanin-related gene expression levels regulated by the MYB10 transcription factor (Telias et al. 2011). Moreover, similar to the change in peel color, stripe color also changes with fruit developmental stage (Biebel and Mazourek 2013).

The emergence of a striped pattern is caused by different accumulations of pigments in the peel. The diversity of fruit color also results largely from the difference in the type and proportion of pigments (Lightbourn et al. 2008; Wang et al. 2010). There is an obvious regulation and genetic relationship between fruit color and stripe pattern. Cucumber has a wide diversity of fruit colors, and the genes underlying the orange mature fruit color (*B*), white immature fruit color (*w*), light green peel (*lgp*) and yellow-green peel (*ygp*) have been cloned (Li et al. 2013; Zhou et al. 2015; Liu et al. 2016; Hao et al. 2018). Related studies have indicated that the molecular mechanisms underlying peel color in cucumber are involved in the cyclase step of chlorophyll biosynthesis, chloroplast division, cytokinin signaling and the biosynthesis of anthocyanin-containing compounds (Lun et al. 2015; Zhou et al. 2015; Liu et al. 2016; Hao et al. 2018).

Cucumber rinds generally display a nonstriped pattern. However, in contrast to uniform rinds, some cucumber varieties may have netted or mottled rinds (Miao et al. 2011; Yang et al. 2014b). Heavy netting at the fruit maturity stage is controlled by a single dominant gene, *Heavy netting* (*H*) (Miao et al. 2011). This gene is located in a 297.7-kb region of chromosome 5 and is linked to the *dull fruit skin* (*d*), *uniform immature fruit color* (*u*) and *fruit ribbing* (*fr*) genes (Miao et al. 2011; Wang et al. 2014). The color at the flowering end of immature cucumber primarily falls into two types: mottled or uniform fruit color (Yuan et al. 2008). The mottled fruit color phenotype (*U*) is dominant over the uniform fruit color phenotype (*u*), and a single gene is responsible for phenotypic segregation. The *u* gene is closely linked to the *Tu* gene (*Tuberculate fruit*) and the *D* gene (*Dull fruit skin*). A single recessive gene, *u* (*uniform immature fruit color*), has been mapped to a 313.2-kb region on chromosome 5 (Yang et al. 2014b).

Research on cucumber rind appearance formation has important theoretical and practical significance for understanding the mechanism underlying rind appearance polymorphisms and for breeding new cucumber germplasm resources. In this study, an irregularly striped cucumber rind mutant (*st*) was obtained, and the objectives of this study were to determine the genetic basis and identify the candidate gene underlying the irregularly striped rind mutant.

Materials and methods

Plant materials and phenotypic data collection

The irregularly striped rind (*ist*) cucumber material was identified from a spontaneous mutation in the genetic background of inbred line CCMC, which has a uniform dark green rind. An F_2 population from a cross between *ist* and CCMC was used for genetic analysis and BSA sequencing. Another F_2 segregating population from a cross between the *ist* mutant and Hazerd, an inbred line with a uniform green rind, was used for linkage mapping of the *ist* gene. The presence of rind stripes was determined by visual observations.

All materials were grown in a greenhouse at the BaiMa Cucumber Research Station of Nanjing Agricultural University, Nanjing, China.

Pigment content measurement and transmission electron microscopy

Peel samples from fruit of the *ist* mutant and CCMC at 10 days after pollination (dap) were collected to measure the chlorophyll and carotenoid contents according to the methods described by Arnon (1949) and Soto-Zamora et al. (2005). The following calculation formula was used:

$$Ca = 13.95 \times OD665 - 6.88 \times OD649$$

$$Cb = 24.96 \times OD649 - 7.32 \times OD665$$

$$Cc = (1000 * OD470 - 2.05 * Ca - 114.8 * Cb)/245$$

The same samples were then prepared for transmission electron microscopy (TEM). Peel sections were vacuumized and fixed in 2.5% glutaraldehyde at 4 °C for 4 h. The samples were then dehydrated in a graded ethanol series, and critical-point drying was performed using liquid CO_2 in a Bal-Tec CPD 030 critical-point drier. Afterward, the samples were coated with 15 nm gold on aluminum stubs in a Sputter Coater Bal-Tec SCD 005. A Hitachi S-3500N scanning electron microscope was subsequently used for sample observations.

Bulked segregant analysis sequencing (BSA-seq)

BSA-seq (Zhang et al. 2018) was used to map the *ist* gene. Genomic DNA was extracted from young leaves of F_2 individuals generated from a cross between *ist* and CCMC. An s pool containing equal amounts of DNA from 24 *ist* individuals and a WT (wild type) pool containing 20 individuals with uniform peels were generated for whole-genome resequencing with an Illumina HiSeq 2500 (500 bp). Short reads obtained from the two DNA pools were aligned against the 'Chinese Long' cucumber reference genome to obtain a consensus sequence (Huang et al. 2009). Single-nucleotide polymorphism (SNP) calling, aligned data filtering, SNP index calculation and sliding window analysis were performed as described in our previous work (Song et al. 2018).

Genetic mapping of the *ist* candidate gene

Since the *ist* mutant and CCMC have a low level of polymorphism, a new F_2 population ($n = 134$) from a cross between *ist* and Hazerd was applied to map the *ist* locus. Polymorphic markers between *ist* and Hazerd were identified from 450 published SSR markers evenly distributed across seven chromosomes (Cavagnaro et al. 2010). Via polymorphic markers and two DNA sample pools from eight *ist* mutant and eight individuals with uniform peels in the new F_2 population, the linkage between the markers and *ist* phenotype was determined. A large F_2 population ($n = 1578$) from the same cross was used to finely map the candidate gene. Whole-genome resequencing (Illumina HiSeq 4000 platform) was conducted for *ist*, CCMC and Hazerd to develop Indel and SNP markers within the initial interval. Polymorphisms of the Indel markers were detected by polyacrylamide gel electrophoresis. The polymorphism of the SNP markers was detected by Sanger sequencing. All primer sequences used in this study are listed in Table S1.

Sequence analysis, annotation and identification of the *ist* candidate gene

Potential mutation sites were identified using resequencing data from *ist*, CCMC, Hazerd and two BSA pools. Candidate genes in the target region were predicted using the 'Chinese Long' cucumber reference genome (http://cucurbitgenomics.org/). Functional annotations of the candidate genes were acquired from the Cucumber Genome Database and the NCBI (National Center for Biotechnology Information) database (https://www.ncbi.nlm.nih.gov/). Total RNA was extracted from the leaves of the *ist* mutant and CCMC plants using Trizol. cDNA was then synthesized using a Prime Script™ RT Reagent Kit (TaKaRa) following the manufacturer's instructions. The coding sequences of the candidate gene from the *ist* mutant and WT were sequenced. Full-length cDNA sequences and corresponding protein alignments were analyzed using DNAMAN software.

The homologous amino acid sequences of the predicted proteins from various plant species were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/). MEGA 7.0.21 software was then used to construct a neighbor-joining tree based on 1000 bootstrap replications (Saitou and Nei 1987).

Quantitative real-time PCR (qRT-PCR) analysis

Peel samples at different development stages (5, 10, 30 dap) from the *ist* mutant and CCMC plants were harvested to perform quantitative real-time PCR (qRT-PCR). The experiment was performed with a SYBR Premix Ex Taq™ Kit (TaKaRa) in a Bio-Rad iQ1 Real-time PCR system (Bio-Rad) as described by Li et al. (2014), and the values from triplicate reactions were averaged. Relative mRNA expression data were analyzed using the ΔCt method. The threshold cycle (Ct) value of each gene was calculated and then normalized to the Ct value of *Cs-Actin*. The primer sequences used are listed in Table S1.

RNA-seq analysis

Total RNA was isolated from 10-dap peels of the *ist* mutant and CCMC using Trizol, with three repetitions. RNA quality and quantity were measured using an Agilent 2100 Bioanalyzer. RNA-seq analysis was conducted by the Oebiotech Company (Shanghai, China). The libraries were constructed using a TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, CA, USA). The libraries were then sequenced on an Illumina sequencing platform (Illumina HiSeq X Ten), and 125-bp/150-bp paired-end reads were generated. After removing the reads containing poly-N and low-quality reads, the clean reads were mapped to the 'Chinese Long' cucumber reference genome using HISAT2 (Kim et al. 2015). The FPKM value of each gene was calculated using cufflinks (Trapnell et al. 2010; Roberts et al. 2011). Moreover, DEGs were identified using the DESeq R package functions estimateSizeFactors and nbinomTest. A P value < 0.05 and a fold change > 2 or a fold change < 0.5 were used as threshold criteria for significantly different expression (Anders and Huber 2013). GO enrichment and KEGG pathway enrichment analyses of the DEGs were performed using R based on a hypergeometric distribution. The RNA-seq data were deposited in the NCBI public database (SRA accession: PRJNA541993).

Results

Characterization of the *irregularly striped rind* cucumber mutant

Compared with that of the WT (CCMC), the peel of the *ist* mutant showed green irregular stripes on a yellow-green background at the immature fruit stage (approximately 10 dap) (Fig. 1a) and a white-yellow background at the mature fruit stage (approximately 30 dap) (Fig. 1b). The stripes were scattered on the peel of the *ist* mutant, with one end having dense stripes and the other having sparse ones (Fig. 1).

Compared with that of the WT, the peel of the *ist* had a lower chlorophyll content, including Chl a and Chl b, and there was a low level of carotenoid contents in both the *ist* and WT (Fig. 2a). Furthermore, chloroplast development in wild type and the *ist* mutant was characterized by analyzing the ultrastructure of plastids of 10-dap peels using TEM. There was an evident difference in the chloroplast numbers in the peel cells between the wild type and *ist* mutant. There were fewer chloroplasts in the peel cells of the *ist* mutant than in the WT (Fig. 2b, d). Granum stacks were less dense and appeared scattered in the *ist* mutant compared to the wild type (Fig. 2c, e). Therefore, the observed irregularly striped rind phenotype is due to impaired chloroplast development in the *ist* mutant.

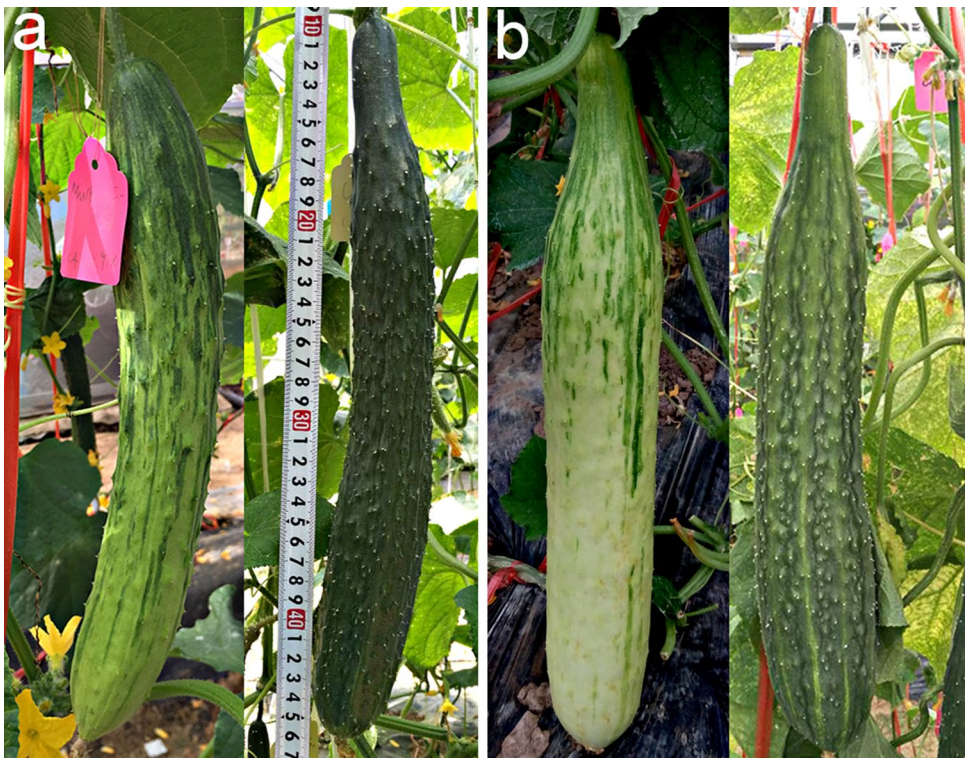

Fig. 1 Phenotypic analysis of the *ist* mutant. Field-grown irregularly striped rind mutant and WT plants at the immature (**a**) and mature stages (**b**)

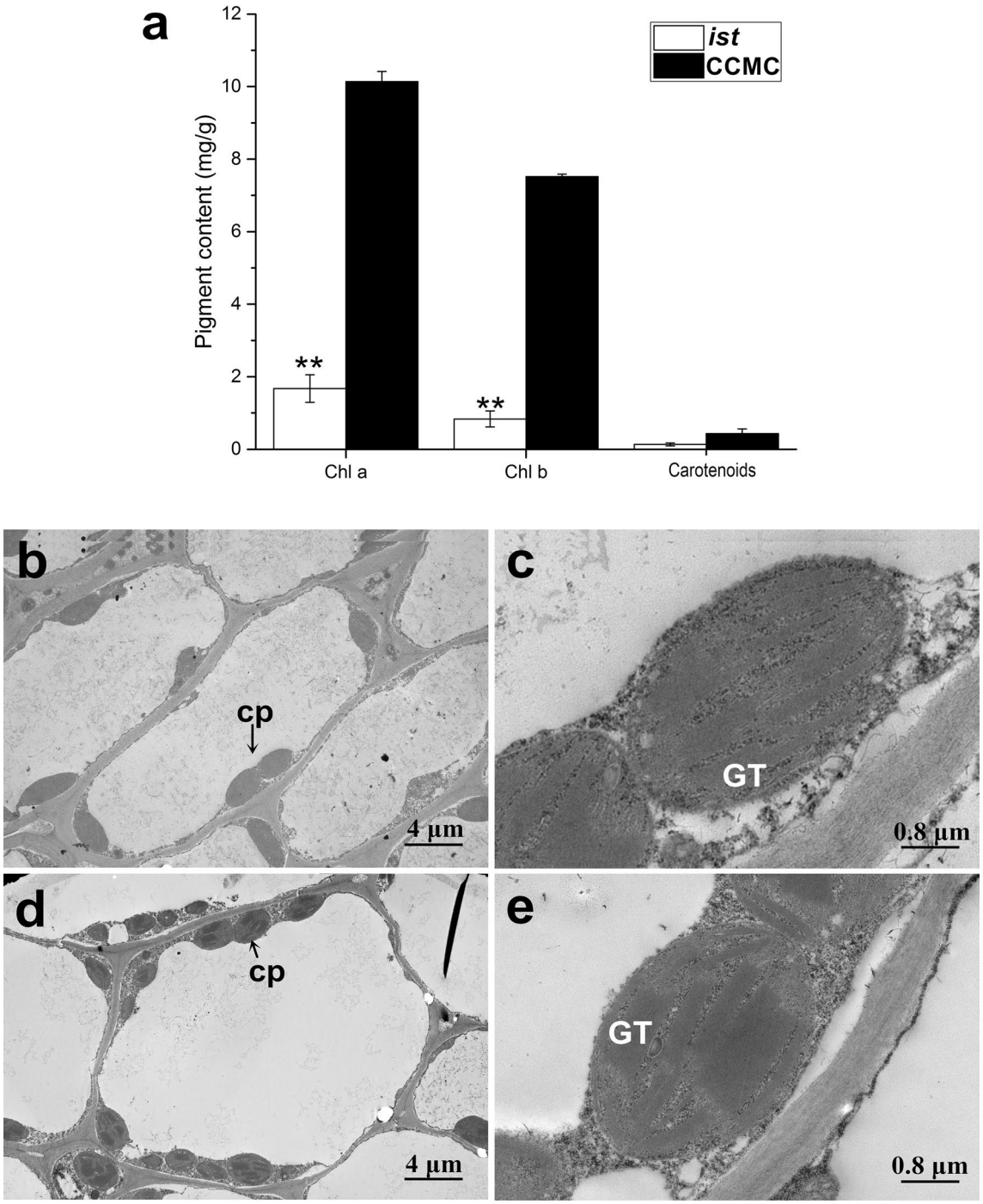

Fig. 2 Pigment contents and chloroplast structure of *ist* and WT. **a** Chlorophyll and carotenoid contents in 10-dap peels of *ist* and WT. The values are the means ± SDs. $^{**}P < 0.01$ (Student's *t* test). Electron microscopy of chloroplast ultrastructures in *ist* (**b**, **c**) and WT (**d**, **e**). CP, chloroplast; GT, granum thylakoid stacks

The genetic characteristics of the *ist* trait were investigated based on populations from crosses of the *ist* mutant with CCMC and Hazerd. All F_1 plants displayed green peels without any stripe phenotype. In the segregating groups, the F_2 plants from the cross between CCMC and *ist* exhibited a segregation pattern that included 29 plants with irregularly striped rinds and 119 plants with uniform rinds, while the F_2 population from the cross between Hazerd and *ist* exhibited a segregation pattern that included 34 plants with irregularly striped rinds and 128 plants with uniform rinds. Both F_2 populations presented a segregation ratio of 3:1, suggesting that the irregularly striped rind trait is controlled by a

single recessive gene. In addition, the number, distribution and width of the stripes in the *ist* individuals of the F_2 population are different (Fig. S1).

Bulked segregant analysis sequencing and linkage mapping of the *ist* locus

Based on the phenotype of the F_2 population from the cross between the *ist* mutant and CCMC, 24 individuals with irregularly striped rinds and 20 individuals with uniform peels were selected and bulked as the s pool and WT pool, respectively. For BSA-seq, a total of 13.7 GB and 14.7 GB of clean data were obtained from the s pool and WT pool, with genome coverages of 67.7- and 72.9-fold, respectively (Table S2). In total, 184,246 and 190,573 SNPs were identified on seven chromosomes from the s pool and WT pool, respectively. To identify the causal region, the Δ (SNP index) value was calculated, and the results showed that the Δ (SNP indices) values were approximately 0.5 for most parts of the genome. A single 2.1-Mb genomic region harboring a high SNP index with a cluster of SNPs was identified on chromosome 1 as the *ist* locus candidate region (Fig. S2).

To validate the candidate region, a new F_2 segregating population ($n = 134$) from a cross of *ist* with Hazerd was used to map the *ist* locus using the BSA method. First, two DNA pools, an s pool (eight individuals with irregularly striped rinds) and a u pool (eight individuals with uniform peels), were constructed. Fifty-two polymorphic SSR markers (11.6%) were identified between *ist* and Hazerd among 450 markers evenly distributed across seven chromosomes. After polymorphism testing among these 52 markers between the two DNA pools, two SSR markers (UW085383 and UW044593) were determined to be linked to the *ist* phenotype. The physical location of the two markers was determined in the 'Chinese Long' reference genome by BLAST software (168,193/705,692 on chromosome 1). This result was consistent with the BSA-seq interval (Fig. 3a).

To precisely narrow the potential region, the resequencing data of *ist* and Hazerd were analyzed for the potential 2.1-Mb region to develop additional polymorphic markers. Thirteen out of 32 Indel markers were polymorphic between both parents and were applied to genotype the 134 F_2 individuals. An initial linkage map was constructed using Join Map 4.0, and the *ist* locus was localized within the 0.8-cM region flanked by SFIndel5 and five cosegregating markers (Fig. 3b). A large F_2 population comprising 1578 plants was adopted to further map the target gene. Three new Indel and two SNP polymorphic markers were developed. SFIndel5 and SFIndel14 were used to screen the 1578 F_2 individuals, and 51 recombinants were obtained. Consequently, the *ist* gene was delimited within a 144-kb region between markers SFSNP-1 (918,885 bp) and SFSNP-18 (1,057,836 bp) on chromosome 1, cosegregating with Sindel 61 (Fig. 3c).

Candidate mutation sites and gene identification

According to the 'Chinese Long' cucumber genome database, 21 predicted genes were annotated in the above-mentioned 144-kb region (Fig. 3d, Table 1). To identify potential gene and mutation sites, the resequencing data of *ist*, CCMC, Hazerd and the two BSA pools were analyzed. SNPs, Indels and SVs were detected in *ist*, and the s pool (homozygous) and WT pool (heterozygous) instead of CCMC and Hazerd were selected. In total, 168 SNPs and one Indel were selected in the potential region corresponding to the criteria. Among them, 156 mutation sites were located in the intergenic region or intron region. For the remaining 13 SNPs, seven caused a synonymous variant, while the others (*SNP1G975302*/*SNP1G983456*/*SNP1G983522*/*SNP1G998065*/*SNP1G998383*/*SNP1G1058630*) caused a nonsynonymous mutation between *ist* and the two parents (CCMC and Hazerd), indicating that these SNPs were the causative mutation sites for the *ist* mutant.

According to the Cucumber Genome Database, *SNP1G975302* is located in gene *Csa1G004970*, both *SNP1G983456* and *SNP1G983522* are located in *Csa1G004980*, *SNP998065* and *SNP998383* are located in *Csa1G005490*, and *SNP1G1058630* is located in gene *Csa1G005600*. qRT-PCR was performed on 10-dap peels of *ist* and WT plants to determine the candidate gene. The results showed that the *Csa1G005490* gene was more highly expressed in the WT than in the *ist* mutant, while the other three genes (*Csa1G004970, Csa1G004980* and *Csa1G005600*) exhibited similar expression levels in the *ist* mutant and WT. The gene annotation analysis revealed that *Csa1G004970* has functions related to the CASP-like protein, *Csa1G004980* is related to the EXECUTER protein, *Csa1G005490* is related to the polygalacturonase-1 non-catalytic subunit beta protein, and *Csa1G005600* is related to the aminoacylase-1 protein (Table 1). PG1β can cause tomato fruit to soften gradually and can cause a change from the green astringent stage to the red ripening stage in tomato (Zheng et al. 1992). All of this evidence suggests that *Csa1G005490* is a good candidate gene for the *ist* locus (Fig. 4a).

The full-length genomic DNA of *Csa1G005490* in the 'Chinese Long' reference genome is 1938 bp and contains only one exon. The coding sequence of the candidate gene was aligned between the *ist* mutant and CCMC by DNAMAN software, and two SNPs corresponding to *SNP998065* and *SNP998383* were identified (Fig. 4b). The two base substitutions resulted in nonsynonymous mutations, which led to an amino acid substitution (Fig. S3). Functional annotation of the candidate genes indicated that *Csa1G005490*

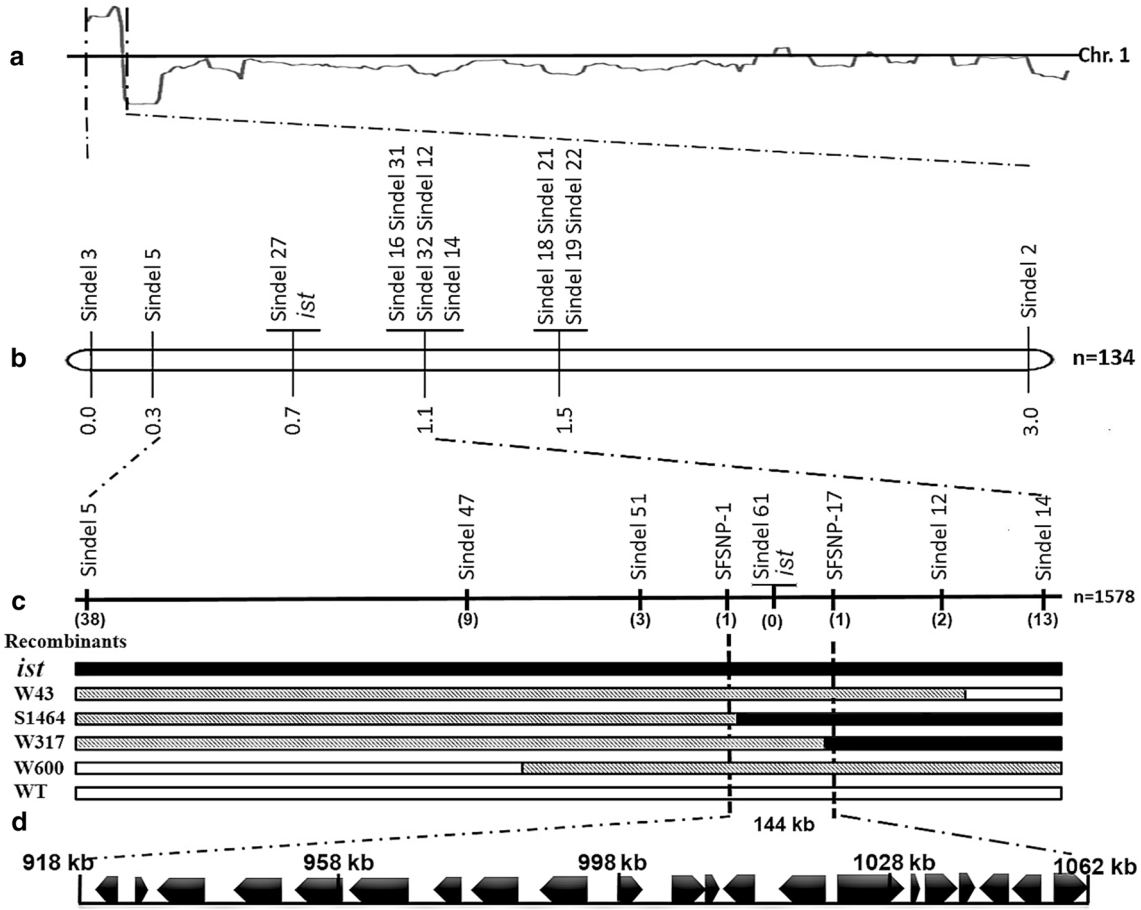

Fig. 3 Linkage mapping of the *ist* gene. **a** BSA-seq mapped the *ist* locus to one end of chromosome 1 (the SNP index is represented by the black curve). **b** A genetic map that delimited the *ist* locus to a 0.8 cM region. **c** Fifty-one recombinants and eight polymorphic markers were applied to narrow the *ist* locus to a 144-kb region. The numbers in parentheses indicate the number of recombinant plants of each marker. The black box indicates the mutant genotype, the white box indicates the wild-type genotype, and the striped box indicates the heterozygous genotype. **d** Physical position of the mapping region. The black arrows indicate genes in the interval

encodes a polygalacturonase-1 noncatalytic subunit beta protein (PG1β) containing a BURP superfamily domain. Alignment of 14 homologous protein sequences from other plant species revealed a high degree of conservation of the BURP superfamily domain, varying from 92.91% (melon homolog) to 60.12% (Arabidopsis homolog). A phylogenetic tree was subsequently constructed using the neighbor-joining (NJ) algorithm (Fig. 4c). The cucumber and melon PG1β proteins, which showed the highest percentage of identity in the alignment, clustered together.

The expression level of the *Csa1G005490* gene in the peels was measured at different stages of fruit development (5 dap, 10 dap and 30 dap) in the *ist* mutant and WT. According to qRT-PCR results, the expression level of the gene in *ist* was significantly lower than that in WT. The expression of *Csa1G005490* was higher during immature fruit stages and gradually decreased as the fruit matured in both the *ist* mutant and WT (Fig. 4d).

Comparative transcriptome profiling of CCMC and the *ist* mutant

To explore the possible mechanism of rind appearance formation in the *ist* mutant from the perspective of gene expression, RNA-seq analysis was performed with the 10-dap peels of the WT and *ist* mutant. In total, there were 1796 differentially expressed genes (DEGs) between the *ist* mutant and wild-type samples, among which 347 and 1449 of the DEGs were upregulated and downregulated, respectively. The GO annotations and functional classifications indicated that these DEGs were mainly enriched in photosynthesis and light harvesting in photosystem I in the biological process category, plastid large ribosomal subunit in the cellular component and pigment binding in the molecular function category (Fig. S4). The DEGs were also analyzed by KEGG enrichment and were mainly enriched in metabolic pathways such as those involving

Table 1 Predicted genes in the 144-kb region

No.	Gene ID[a]	log2FoldChange[b]	Gene annotation
1	*Csa1G004910*	0.610251	Prohibitin
2	*Csa1G004920*	−0.57811	Protein of unknown function DUF1645
3	*Csa1G004930*	−0.5004	1-Deoxy-D-xylulose-5-phosphate reductoisomerase
4	*Csa1G004940*	0.197032	Peroxisomal membrane protein PEX14
5	*Csa1G004950*	−0.12215	Exosome complex exonuclease RRP40
6	*Csa1G004960*	0.048379	Kinesin light chain
7	*Csa1G004970*	−0.10064	CASP-like protein
8	*Csa1G004980*	−0.2376	Protein EXECUTER
9	*Csa1G004990*	0.215843	RING finger protein
10	*Csa1G005490*	−2.53757	Polygalacturonase-1 noncatalytic subunit beta
11	*Csa1G005500*	0.559848	Protein kinase
12	*Csa1G005510*	0.61018	Protein kinase
13	*Csa1G005520*	1.345753	Aldehyde dehydrogenase
14	*Csa1G005530*	0.239728	Cyclic nucleotide-gated ion channel
15	*Csa1G005540*	−0.14907	Spermidine synthase 1
16	*Csa1G005550*	0.311433	Unknown protein
17	*Csa1G005560*	−0.20143	Trehalose-6-phosphate synthase
18	*Csa1G005570*	0.559292	NHL repeat-containing protein-like
19	*Csa1G005580*	0.16492	RNA-dependent RNA polymerase 2
20	*Csa1G005590*	−0.29672	Ubiquitin-conjugating enzyme h
21	*Csa1G005600*	0.223112	Aminoacylase-1

[a]ID numbers were given by cmb (http://cucurbitgenomics.org/search/genome/2)

[b]The log2(fold change) is the log2 conversion of the expression base of the mean ratio of the *ist* group to the WT group taken from RNA-seq

photosynthesis, photosynthesis–antenna proteins, porphyrin and chlorophyll metabolism and alpha-linolenic acid metabolism (Fig. S5). Moreover, 72 transcription factors were identified among the DEGs, including bHLH, ERF, BZIP and MYB transcription factors.

Using RNA-seq data, we further analyzed the expression patterns of several key genes involved in pigment synthesis, such as chlorophyll, carotenoid and anthocyanin synthesis (Fig. S6). Compared to those in the WT, chlorophyllide a oxygenase (CAO), magnesium chelatase subunit (CHLH), NADPH protochlorophyllide oxidoreductase (POR) and several key synthesis genes in the chlorophyll biosynthesis pathway in the *ist* mutant were significantly downregulated. However, there were no significant differences in the expression levels of the carotenoid synthesis pathway and anthocyanin synthesis genes between the WT and *ist* mutant.

In addition, the expression of 12 cucumber genes was confirmed by qRT-PCR to verify the RNA-seq results. The results showed that the qPCR results were essentially consistent with the expression trend revealed by RNA-seq, which showed the reliability of the RNA-seq data (Fig. S7).

Discussion

As one of the most important species of cucurbits, *Cucumis sativus* L. is widely cultivated worldwide. Rind appearance is the most significant commercial trait of cucumber. The genes underlying fruit color (*B, w, lgp, ygp*) and fruit skin texture (*H*, *u*, *D*, *Fr*) have been mapped or cloned (Weng and Wehner 2018). However, there is no information about irregularly striped rinds in cucumber. In the present study, a novel cucumber rind appearance mutant, irregularly striped rind (*ist*), was isolated. The mutant peel exhibited irregular green stripes on a yellow-green background at the immature stage (Fig. 1). Compared with that in the WT peel, the chlorophyll content in the *ist* mutant peel was lower, and the number and structure of the chloroplasts also changed in the peel cells of *ist* compared with WT.

In this study, two materials with uniform rinds were hybridized with the *ist* mutant to produce two F_1 and two F_2 populations. All F_1 materials showed uniform rinds, and the presence or absence of stripes in both F_2 populations was in accordance with the 1:3 Mendelian inheritance law. This result indicated that the irregularly striped rind trait

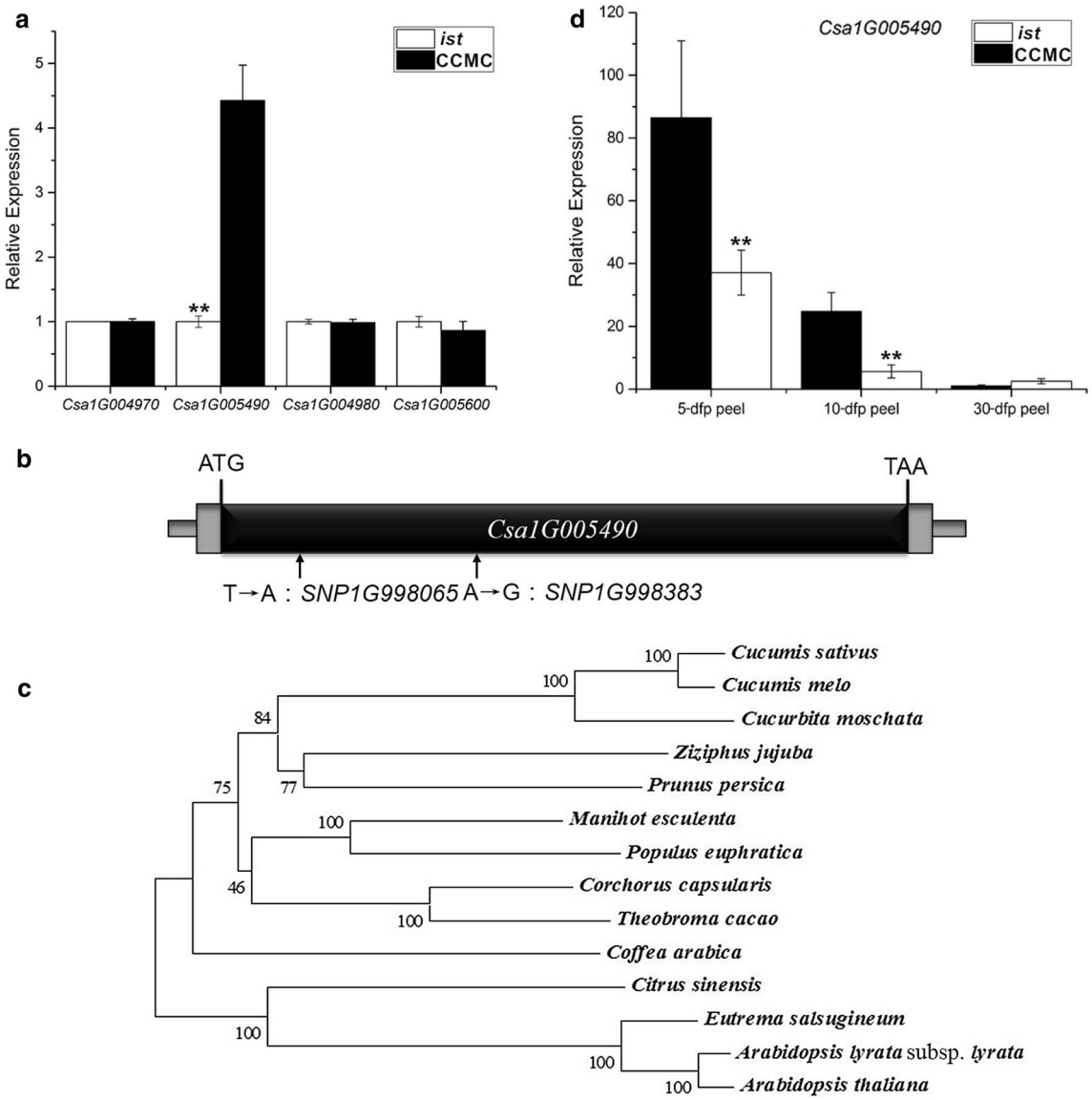

Fig. 4 Identification of the *ist* gene and its encoding protein. **a** Expression evaluation of mutant genes in the candidate region by qRT-PCR. The values are the means ± SDs of three biological and three technical replicates. Asterisks indicate $P < 0.01$ (Student's *t* test). **b** Diagram of the gene structure of *Csa1M005490*. Two single-nucleotide mutations occur in the exon of *Csa1M005490* in the *ist* mutant. **c** Phylogenetic tree of PG1β proteins from different plant species, constructed by MEGA 7. The numbers indicate the percentage of replicate trees in which the associated taxa clustered together according to the bootstrap test (1000 replicates). **d** Relative expression of *Csa1M005490* at different developmental stages of WT and *ist* peels

of cucumber was regulated by a single recessive gene. The inheritance pattern underlying striped rinds has been identified in several cucurbit crop species, such as pumpkin, watermelon and melon. The presence or absence of striped rinds is mainly regulated by a single gene, but the color, shape, density and distribution of the stripes are controlled by a series of genes (Paris 2003; Gama et al. 2015; Lv et al. 2018; Liu et al. 2019). Similarly, different colors, shapes and distributions of stripes were found in different *ist* individuals of the F_2 population (Fig. S1), which may be regulated by other unknown genes or pathways.

In this paper, BSA-seq combined with linkage mapping was used to identify the candidate gene at the *ist* locus. Both methods mapped the *ist* locus to the same location on chromosome 1 (Fig. 3, Fig. S2), which verified the accuracy of the results. The candidate region range from linkage mapping spanned 144 kb, and only six mutation sites that might cause gene changes were obtained by screening the resequencing data of the parents and mixed pools. Two of these sites were located in the gene *Cs1G005490*, which caused differences in expression and changes in its encoded amino acid sequence in the *ist* mutant compared to the wild type

(Fig. 4, Fig. S3). Because other mutant genes in the region did not exhibit a change in their expression level (Fig. 4a), the *Cs1G005490* gene was considered the candidate gene for the irregularly striped rind mutant. However, sometimes not all polymorphisms will be detected due to low coverage of some regions. The expression pattern can show only the difference in the transcription level of this gene between the WT and *ist* and cannot represent the functional difference. Therefore, transformation experiments with candidate genes will be needed to verify this result.

Functional and structural annotations showed that *Cs1G005490* encodes a polygalacturonase-1 noncatalytic subunit beta protein containing a BURP domain superfamily domain that is highly conserved among homologous proteins from other species (Fig. 4c). The BURP family, whose members are types of plant-specific proteins, is mainly involved in the development of specific tissues and cells and function in response to stress (Hattori et al. 1998). Members of the BURP protein family, which comprises four representative members, BNM2, USPs, RD22 and PG1β, that present similar primary structures but exhibit different expression patterns, play different roles in plant growth and development (Hattori et al. 1998; Zheng et al. 1992; Wang et al. 2003; Harshavardhan et al. 2014). PG1β is a thermostable, acidic glycosylated subunit protein of the PG1 protein, with a molecular weight of 38 kD. This beta subunit of the PG protein can interact with the structural components of the cell wall and the PG2 catalytic subunit to stabilize or modulate the activity of the polygalacturonase complex (Watson et al. 1994). PG has been reported to be involved in fruit ripening and softening in tomato, banana and apple (Smith et al. 1990; Hadfield and Bennett 1998; Asif and Nath 2005; Longhi et al. 2013). The PG1β protein participates in fruit development and maturity and is mostly found in developing fruit. In tomatoes, PG1β is expressed mainly in the pericarp cells of fruit from 20 days after flowering to the mature stage; this protein can gradually soften tomato fruit and make the fruit change from the green astringent stage to the red ripening stage (Zheng et al. 1992). In this paper, the expression level of PG1β in WT peel was significantly higher than that in the *ist* mutant peel (Fig. 4d). Subsequently, the high expression level in both WT and *ist* decreased as the fruit developed. The different expression patterns suggested that PG1β may play different roles between cucumber and tomato. However, it has been shown that PG activity is lacking in the softening process of melon (McCollum et al. 1989). In addition, when *PG* was overexpressed in apples, the color of the leaves of the transgenic plants changed, suggesting that the *PG* gene may be involved in the regulation of pigment synthesis (Atkinson et al. 2002). In this study, we demonstrated that two single amino acid substitutions in the PG1β protein resulted in a change in cucumber rind appearance (Fig. 4b). Given the function of the PG1β protein, we speculate that mutations in this protein may cause dysplasia of fruit epidermal cells, resulting in changes in rind appearance.

In this paper, more than 1700 differentially expressed genes between the *ist* mutant and WT were identified by RNA-seq, and most of them were downregulated. These genes were enriched in photosynthesis, pigment synthesis, chloroplast structure and other pathways related to plant fruit color. The expression changes in these genes have an important regulatory effect on the synthesis and accumulation of pigments in the *ist* mutant. Moreover, the GO annotation and KEGG enrichment analyses of the differentially expressed genes revealed a significant enrichment of a series of items related to photosynthesis (Fig. S4, Fig. S5). Photosynthesis can directly affect the synthesis of pigments and chloroplast formation. Downregulated expression of photosynthesis-related genes is closely related to thylakoid developmental defects.

The degree of fruit coloration depends mainly on the type and proportion of different pigments (Lightbourn et al. 2008). The mutation of any related gene may result in the production of new fruit colors. Mutation of the lycopene synthesis gene causes different tomato fruit colors (Ronen et al. 1999). The difference in the carotenoid synthesis gene makes pumpkin peels show different red colors (Itle and Kabelka 2009). Mutations in the anthocyanin regulatory factor (a MYB transcription factor) cause different fruit colors in horticultural crop species such as tomato, grape and cucumber (Azuma et al. 2008; Dal Cin et al. 2011; Hao et al. 2018). In the present study, the RNA-seq data indicated that several key genes in the chlorophyll synthesis pathway were downregulated in the *ist* mutant. The downregulated expression of the CAO, CHLH and POR genes can cause a decrease in the chlorophyll synthesis rate, affecting the normal accumulation of chlorophyll (Stenbaek and Jensen 2010). Changes in the expression of chlorophyll synthesis genes may also be the direct cause of the rind appearance changes in the *ist* mutant.

It is worth noting that genes encoding several transcription factors were identified among the DEGs, including MYBs, bHLHs and WRKYs. MYB transcription factors have been shown to affect peel color by participating in the synthesis of anthocyanins in some plant species (Azuma et al. 2008; Dal Cin et al. 2011). Mutations of two types of MYB transcription factors, R2R3-MYB and MYB36, cause the formation of white peels and yellow-green peels in cucumber (Liu et al. 2016; Jiao et al. 2017; Hao et al. 2018). By interacting with MYBs, bHLH transcription factors promote anthocyanin accumulation (Xie et al. 2012). We speculate that these transcription factors caused a change in rind appearance through several pathways (such as the anthocyanin pathway).

In summary, we propose that mutations in the *ist* gene lead to a change in cucumber fruit color and that the *ist* gene may be involved in pigment synthesis in the peel. These results will be useful for elucidating the mechanisms of pigment accumulation and regulation in cucumber fruit peels. The complex relationships between *ist* and other chlorophyll synthesis genes deserve further investigation.

Acknowledgments This research was partially supported by the National Natural Science Foundation of China (31772318), the Fund for Independent Innovation of Agricultural Science and Technology of Jiangsu Province [CX(17)3016] and the National Supporting Programs (2016YFD0100204-25) from the Ministry of Science and Technology of China.

Author contribution statement QL and JC conceived the research and designed the experiments. MS performed the research and analyzed the data. MS and QL wrote the manuscript. MZ, FC and JW were involved in the phenotypic selection, DNA extraction and mapping work. QW participated in the genome resequencing analysis, and MD participated in modifying the paper. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest.

References

Anders S, Huber W (2013) Differential expression of RNA-Seq data at the gene level—The DESeq package. Available at http://www.bioconductor.org/packages/devel/bioc/vignettes/DESeq/inst/doc/DESeq.pdf. Accessed 17 Oct 2013

Arnon DI (1949) Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. Plant Physiol 24:1–15

Asif MH, Nath P (2005) Expression of multiple forms of polygalacturonase gene during ripening in banana fruit. Plant Physiol Biochem 43:177–184

Atkinson RG, Schröder R, Hallett IC, Cohen D, MacRae EA (2002) Overexpression of polygalacturonase in transgenic apple trees leads to a range of novel phenotypes involving changes in cell adhesion. Plant Physiol 129(1):122–133

Azuma A, Kobayashi S, Mitani N, Shiraishi M, Yamada M, Ueno T, Kono A, Yakushiji H, Koshita Y (2008) Genomic and genetic analysis of Myb-related genes that regulate anthocyanin biosynthesis in grape berry skin. Theor Appl Genet 117:1009–1019

Biebel N, Mazourek M (2013) Inheritance of rind color and reverse striping in a *Cucurbita pepo* (subsp. Texana) cross. Cucurbit Genet Coop Rep 35–36:16–17

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

Choudhury HC (1977) Genetical studies in West African egg plants. Indian J Genet Plant Breed 37:26–33

Dal Cin V, Tieman DM, Tohge T, McQuinn R, de Vos RC, Osorio S, Schmelz EA, Taylor MG, Smits-Kroon MT, Schuurink RC, Haring MA, Giovannoni J, Fernie AR, Klee HJ (2011) Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit. Plant Cell 23:2738–2753

Danin-Poleg Y, Tadmor Y, Tzuri G, Reis N, Hirschbery J, Katzir N (2002) Construction of a genetic map of melon with molecular markers for horticultural traits, and localization of genes associated with ZYMV resistance. Euphytica 125(3):373–384

Gama RN, Santos CA, Dias RC, Alves JC, Nogueira TO (2015) Microsatellite markers linked to the locus of the watermelon fruit stripe pattern. Genet Mol Res 14:269–276

Hadfield KA, Bennett AB (1998) Polygalacturonases: many genes in search of a function. Plant Physiol 117:337–343

Hao N, Du YL, Li HY, Wang C, Wang C, Gong SY, Zhou SM, Wu T (2018) *CsMYB36* is involved in the formation of yellow green peel in cucumber (*Cucumis sativus* L.). Theor Appl Genet 131:1659–1669

Harshavardhan VT, Son LV, Seiler C, Junker A, Weigelt-Fischer K, Klukas C, Altmann T, Sreenivasulu N, Baumlein H, Kuhlmann M (2014) AtRD22 and AtUSPL1, members of the plant-specific BURP domain family involved in *Arabidopsis thaliana* drought tolerance. PLoS One 9(10):e110065

Hattori J, Boutilier KA, Campagne MMV, Miki BL (1998) A conserved BURP domain defines a novel group of plant proteins with unusual primary structures. Mol Gen Genet 259:424–428

Huang SW, 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

Itle RA, Kabelka EA (2009) Correlation between L* a* b* color space values and carotenoid content in pumpkins and squash (*Cucurbita* spp.). HortScience 44(3):633–637

Jiao JQ, Liu HQ, Liu J, Cui MM, Xu J, Meng HW, Li YH, Chen SX, Cheng ZH (2017) Identification and functional characterization of APRR2 controlling green immature fruit color in cucumber (*Cucumis sativus* L.). Plant Growth Regul 83:233–243

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357

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

Li J, Wu Z, Cui L, Zhang TL, Guo QW, Xu J, Jia L, Lou QF, Huang SW, Li ZG, Chen JF (2014) Transcriptome comparison of global distinctive features between pollination and parthenocarpic fruit set reveals transcriptional phytohormone cross-talk in cucumber (*Cucumis sativus* L.). Plant Cell Physiol 55:1325–1342

Lightbourn GJ, Griesbach RJ, Novotny JA, Clevidence BA, Rao DD, Stommel JR (2008) Effects of anthocyanin and carotenoid combinations on foliage and immature fruit color of *Capsicum annuum* L. J Hered 99:105–111

Liu H, Jiao JQ, Liang XJ, Liu J, Meng HW, Chen SX, Li YH, Cheng ZH (2016) Map-based cloning, identification and characterization of the *w* gene controlling white immature fruit color in cucumber (*Cucumis sativus* L.). Theor Appl Genet 129:1247–1256

Liu L, Sun TT, Liu XY, Guo Y, Huang X, Gao P, Wang XZ (2019) Genetic analysis and mapping of a striped rind gene (*st3*) in melon (*Cucumis melo L.*). Euphytica 215(2):20

Longhi S, Hamblin MT, Trainotti L, Peace CP, Velasco R, Costa F (2013) A candidate gene based approach validates Md-PG1 as the main responsible for a QTL impacting fruit texture in apple (*Malus* × *domestica Borkh*). BMC Plant Biol 13:37

Lou LL, Wehner TC (2016) Qualitative inheritance of external fruit traits in watermelon. HortScience 51:487–496

Lun YY, Wang X, Zhang CZ, Yang L, Gao DL, Chen HM, Huang SW (2015) A *CsYcf54* variant conferring light green coloration in cucumber. Euphytica 208:509–517

Lv JC, Fu QS, Lai Y, Zhou MD, Wang HS (2018) Inheritance and gene mapping of spotted to non-spotted trait gene *CmSp*-1 in melon (*Cucumis melo* L. var. *chinensis Pangalo*). Mol Breed 38:105

McCollum TG, Huber DJ, Cantliffe DJ (1989) Modification of polyuronides and hemicelluloses during muskmelon fruit softening. Physiol Plant 76(3):303–308

Miao H, Zhang SP, Wang XW, Zhang ZH, Li M, Mu SQ, Cheng ZC, Zhang RW, Huang SW, Xie BY, Fang ZY, Weng YQ, 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 182(2):167–176

Paris HS (2003) Genetic control of irregular striping, a new phenotype in *Cucurbita pepo*. Euphytica 129:119–126

Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol 12(3):R22

Ronen G, Cohen M, Zamir D, Hirschberg J (1999) Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant *Delta*. Plant J 17:341–351

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Smith CJ, Watson CF, Morris PC, Bird CR, Seymour GB, Gray JE, Arnold C, Tucker GA, Schuch W, Harding S et al (1990) Inheritance and effect on ripening of antisense polygalacturonase genes in transgenic tomatoes. Plant Mol Biol 14:369–379

Song MF, Wei QZ, Wang J, Fu WY, Qin XD, Lu XM, Cheng F, Yang K, Zhang L, Yu XQ, Li J, Chen JF, Lou QF (2018) Fine mapping of *CsVYL*, conferring virescent leaf through the regulation of chloroplast development in cucumber. Front Plant Sci 9:432

Soto-Zamora G, Yahia EM, Brecht JK, Gardea A (2005) Effects of postharvest hot air treatments on the quality and antioxidant levels in tomato fruit. Lwt-Food Sci Technol 38:657–663

Stenbaek A, Jensen PE (2010) Redox regulation of chlorophyll biosynthesis. Phytochemistry 71:853–859

Telias A, Kui LW, 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

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28:511

Wang AM, Xia Q, Xie WS, Datla R, Selvaraj G (2003) The classical Ubisch bodies carry a sporophytically produced structural protein (RAFTIN) that is essential for pollen development. Proc Natl Acad Sci USA 100:14487–14492

Wang HL, Wang W, Zhang P, Pan QH, Zhan JC, Huang WD (2010) Gene transcript accumulation, tissue and subcellular localization of anthocyanidin synthase (ANS) in developing grape berries. Plant Sci 179:103–113

Wang M, Gu XF, Miao H, Liu SL, Wang Y, Wehner TC, Zhang SP (2014) Molecular mapping and candidate gene analysis for heavy netting gene (*H*) of mature fruit of cucumber (*Cucumis sativus* L.). Sci Agric Sin 47:1550–1557

Watson CF, Zheng LS, Dellapenna D (1994) Reduction of tomato polygalacturonase beta-subunit expression affects pectin solubilization and degradation during fruit ripening. Plant Cell 6:1623–1634

Weng YQ, Wehner TC (2018) Cucumber gene catalog 2017. Cooperative 2018:17

Xie J, Wehner TC (2001) Gene list 2001 for cucumber. Cucurbit Genet Coop Rep 24:110–136

Xie XB, Li S, Zhang RF, Zhao J, Chen YC, Zhao Q, Yao YX, You CX, Zhang XS, Hao YJ (2012) The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. Plant Cell Environ 35:1884–1897

Yang XQ, Zhang WW, Li Y, He HL, Bie BB, Ren GL, Zhao JL, Wang YL, Nie JT, Pan JS, Cai R (2014a) High-resolution mapping of the dull fruit skin gene *D* in cucumber (*Cucumis sativus* L.). Mol Breed 33(1):15–22

Yang XQ, Li Y, Zhang WW, He HL, Pan JS, Cai R (2014b) Fine mapping of the uniform immature fruit color gene *u* in cucumber (*Cucumis sativus* L.). Euphytica 196:341–348

Yuan XJ, Pan JS, Cai R, Guan Y, Liu LZ, Zhang WW, Li Z, He HL, Zhang C, Si LT, Zhu LH (2008) Genetic mapping and QTL analysis of fruit and flower related traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Euphytica 164(2):473–491

Zhang KJ, Wang X, Zhu WW, Qin XD, Xu J, Cheng CY, Lou QF, Li J, Chen JF (2018) Complete resistance to powdery mildew and partial resistance to downy mildew in a *Cucumis hystrix* introgression line of cucumber were controlled by a co-localized locus. Theor Appl Genet 131(10):2229–2243

Zheng LS, Heupel RC, Dellapenna D (1992) The beta-subunit of tomato fruit polygalacturonase isoenzyme-1-isolation, characterization, and identification of unique structural features. Plant Cell 4:1147–1156

Zhou Q, Wang SH, Hu BW, Chen HM, Zhang ZH, Huang SW (2015) An accumulation and replication of chloroplasts 5 gene mutation confers light green peel in cucumber. J Integr Plant Biol 57:936–942