



# Fine mapping of *lcm1*, a gene conferring chlorophyll-deficient golden leaf in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*)

Wei Fu · Xueling Ye · Jie Ren · Qingqing Li ·  
Jiangtao Du · Ailin Hou · Fabo Mei · Hui Feng ·  
Zhiyong Liu

Received: 30 April 2018 / Accepted: 12 February 2019 / Published online: 16 March 2019  
© Springer Nature B.V. 2019

**Abstract** Leaf color mutants are valuable resources for studying regulatory mechanisms of photosynthetic pigment metabolism. In this study, a chlorophyll-deficient golden leaf mutant *lcm1* of Chinese cabbage was identified from its wild-type “FT” by ethyl methanesulfonate (EMS) mutagenesis. The phenotype of the golden leaf mutant *lcm1* was that the leaves remained golden throughout development. Pigment measurements showed that the chlorophyll content of the mutant *lcm1* was less than that of the wild-type “FT”. There were no obvious grana lamellae observed in the chloroplast microstructure of the mutant *lcm1*. Genetic analysis revealed that the mutant *lcm1* phenotype was controlled by a single recessive nuclear gene, *lcm1* (*BrChlH*). The MutMap method and Kompetitive Allele Specific PCR genotyping were used to predict that *Bra006208* encoding a Mg-chelatase H subunit, the candidate gene for the mutant *lcm1*. The 4249-bp *lcm1* gene had five exons and a single nucleotide substitution (G to A) in the third exon that resulted in an amino acid substitution (A to V). This SNP2523351 occurred in the

highly conserved CobN-Mg\_chel domain of the LCM1 (*BrChlH*) protein. This study suggests that *lcm1* gene plays an important role in the function of CHLH and provides a solid foundation for further studies on the development of plant leaf color.

**Keywords** Chinese cabbage · EMS · MutMap · Golden leaf mutant

## Introduction

Leaves are the main organs of photosynthesis and respiration in plants and directly affect crop yield and quality. Leaf color mutations are ubiquitous in plants and are important for the study of chlorophyll (CHL) biosynthesis, chlorophyll regulation and degradation, chloroplast development, and chloroplast-nuclear signal transmission (Colombo et al. 2008; Pogson and Albrecht 2011; Xu et al. 2013). Cloning leaf color mutant genes and studying their functions have important theoretical and practical implications to further understand the regulatory mechanisms of chloroplast development and to cultivate new varieties of higher plants with high photosynthetic efficiency.

The molecular mechanisms of leaf color mutations are complex. Mutant genes may directly or indirectly interfere with the metabolism of CHL, causing leaf color variations in a variety of ways. With regard to existing

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11032-019-0945-z>) contains supplementary material, which is available to authorized users.

W. Fu · X. Ye · J. Ren · Q. Li · J. Du · A. Hou · F. Mei ·  
H. Feng · Z. Liu (✉)  
Department of Horticulture, Shenyang Agricultural University,  
120 Dongling Road, Shenhe District, Shenyang 110866, People’s  
Republic of China  
e-mail: lzyky99@163.com

leaf color mutants, the affected genes can be subdivided into four categories:

1. Genes in the metabolic pathways of photosynthetic pigments. Higher plant photosynthetic pigments include CHL and carotenoids, and the content and proportions of these two pigments are directly related to leaf color. Plant CHL biosynthesis involves 15 enzymatic reactions. Any gene mutation in this pathway can interfere with the formation of CHL, changing the proportion of various pigments in chloroplasts and causing leaf color variation (Beale 2005; Wu et al. 2007; Mochizuki et al. 2010; Braumann et al. 2014; Gao et al. 2016). Mutations in some of the key genes of the CHL catabolic pathway may produce a stay-green phenotype and delay senescence (Kusaba et al. 2007).
2. Genes related to the heme→phytochrome chromophore pathway. CHL biosynthesis and heme biosynthesis are two branches of the tetrapyrrole biosynthetic pathway. If the heme→phytochrome chromophore pathway is disrupted, it leads to changes in the content of the heme, which affects CHL biosynthesis via feedback, causing leaf color variations (Terry and Kendrick, 1999; Chen et al. 2013; Shi et al. 2013).
3. Genes encoding other chloroplast proteins. Plant chloroplasts contain more than 3000 proteins, but chloroplast DNA encodes only ~100 proteins; the remaining proteins are nuclear-encoded. Slight changes in chloroplast proteins may lead to leaf color changes (Kusumi et al. 2011; Dong et al. 2013).
4. Genes that are not directly related to the photosynthetic system. In addition to the light system-related pathways described above, other genetic mutations may cause leaf color variation (Sozer et al. 2010; Lee et al. 2014). Mutation of one DUF3353 family member was found to cause chlorosis of rice leaves, but its function is unknown (Zhou et al. 2012). Other investigators have found that gene mutations involving the absorption and transport of certain metals (e.g., iron and manganese) can also cause chlorotic leaf (Ihnatowicz et al. 2014; Mamidi et al. 2014).

At present, there are several methods for mapping mutant genes. One of the most commonly used is the classical map-based cloning, in which many genes are

isolated. However, traditional methods for the fine mapping of mutant genes are costly and time-consuming. The recently developed MutMap method is a new forward genetic approach based on whole genome sequencing (WGS) for gene mapping and genetic analysis, providing a powerful tool for rapid mining and utilization of genetic variation (Abe et al. 2012). The advantage of this method is that it can be performed quickly and is highly efficient. New approaches extending from this technology are also emerging, such as MutMap+ (Fekih et al. 2013) and MutMap-Gap (Takagi et al. 2013). These methods do not require the establishment of cumbersome mapping populations, the use of genetic crosses or linkage information, thus accelerating the identification process of genomic regions with variation associated with the phenotype of interest. F<sub>2</sub> segregating population was constructed for MutMap, then the mutant, wild-type, and DNA pools of the recessive individuals were analyzed using NGS technology to re-sequence and quickly find the varied SNP (Abe et al. 2012). In recent years, this strategy has been widely used in the cloning of ethyl methanesulfonate (EMS) mutagenesis genes (Takagi et al. 2015; Wachsmann et al. 2017; Zou et al. 2017) and has been successfully applied to map multiple genes in different species (Takagi et al. 2015; Imamura et al. 2018; Jiao et al. 2018; Xu et al. 2018). The MutMap method has been used to isolate some rice genes, including *OsRR22*, a gene responsible for the salinity-tolerant phenotype of *hst1* (Takagi et al. 2015); *WBI*, a gene related to endosperm development (Wang et al. 2018); *Pii*, a gene enhancing rice blast resistance (Takagi et al. 2013); and *OsLAP6/OsPKS1*, a gene controlling pollen exine formation (Zou et al. 2017). Several mutant genes of cucumber produced through EMS mutagenesis have been identified using the MutMap method, including *Csdw*, a regulatory gene associated with dwarf phenotypes (Xu et al. 2018) and *CsMYB36* that can regulate the yellow-green peel coloration in cucumber (Hao et al. 2018). Using the MutMap, researchers showed that the causal gene in *bm40* mutants encodes a GDSL-like lipase in sorghum (Jiao et al. 2018). In quinoa seedlings, *CqCYP76AD1-1* was found to be involved in betalain biosynthesis during the hypocotyl pigmentation process using the MutMap+ method (Imamura et al. 2018). With the MutMap technology involving the known genomic

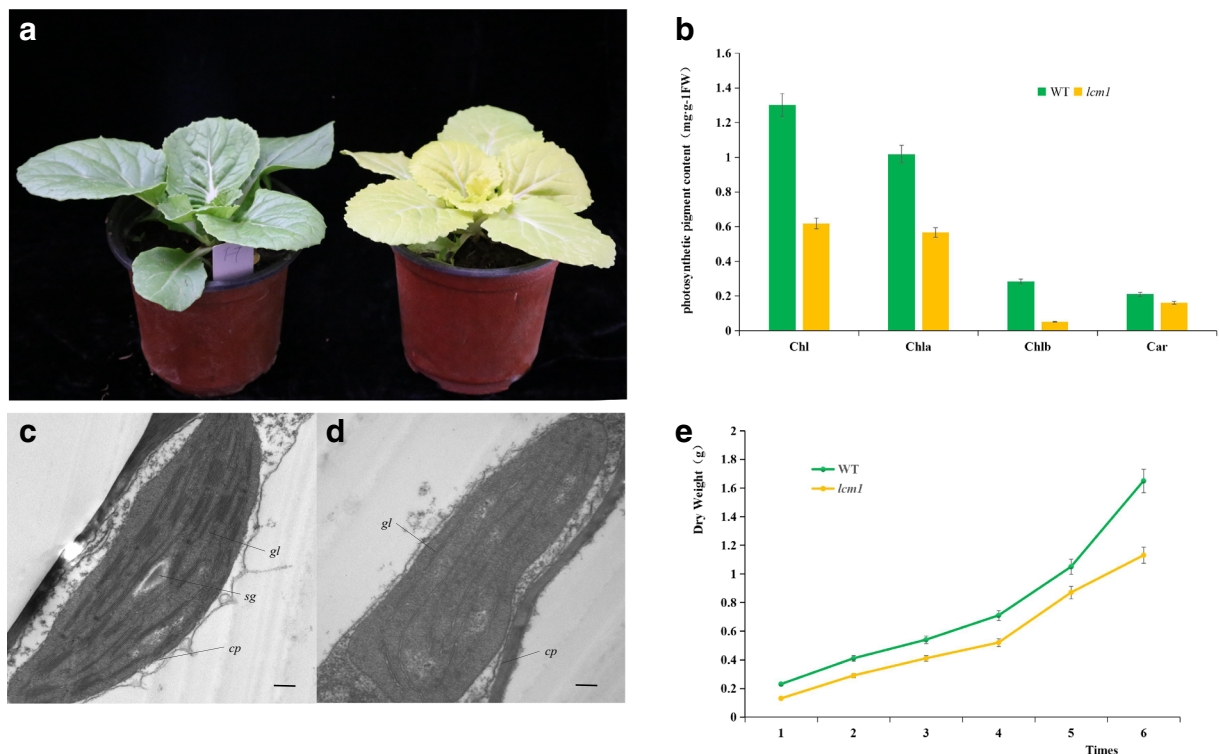
sequences, genomic regions harboring a causal mutation for a given phenotype can be rapidly identified (Abe et al. 2012).

Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) is widely planted in East Asia and is the largest vegetable crop in China. The leaves are not only the photosynthetic organs of cabbage but they are also the harvested product. Leaf color mutation research is important, but currently, there have been few relevant studies of the Chinese cabbage. The release of the Chinese cabbage genomic sequence provided reliable data to support functional genomics studies of *B. rapa* (Wang et al. 2011). In the present study, we obtained a golden leaf mutant in Chinese cabbage by EMS mutagenesis. By using a modified MutMap method and Kompetitive Allele Specific PCR (KASP) results, we determined the SNP2523351 and cloned the corresponding mutant gene. The aims of this research were to provide valuable information to understand the development of leaf color in Chinese cabbage and to provide theoretical support for the breeding of Chinese cabbage possessing high photosynthetic efficiency.

## Materials and methods

### Plant materials

The golden leaf mutant *lcm1* (Fig. 1a) was obtained from the mutagenized population of the Chinese cabbage “FT” treated by EMS mutagenesis, which was a doubled haploid (DH) line and obtained by microspore culture. The specific steps of EMS treatment were as follows: firstly, germinating at 25 °C for 12 h then treating with 0.8% EMS for 12 h, rinsing for 12 h, seeding the treated seeds ( $M_0$ ) in the hole plate, and self-crossing to obtain  $M_1$  seeds. A total of 25 plants were planted of each  $M_1$  strain, and golden leaf plants were identified. The mutant *lcm1* was crossed with wild-type “FT” to construct  $F_1$ ,  $F_2$ , and  $BC_1$  progenies for genetic analysis. Segregation ratio in the  $F_2$  and  $BC_1$  populations was analyzed using the chi-square ( $\chi^2$ ) tests. For MutMap mapping of mutant *lcm1*, the  $F_2$  population was used. All plants were grown in greenhouses at 15–28 °C, at Shenyang Agriculture University in 2017.



**Fig. 1** Plant appearance and physiological characteristics of wild-type “FT” and mutant *lcm1*. **a** Phenotype of parental lines. **b** Photosynthetic pigment content contained corresponding total Chl, Chla, Chlb, and carotenoid (Car) contents in milligrams per gram of fresh weight (FW) in true leaves of both materials. **c-d**

Chloroplast ultrastructure of wild-type “FT” and mutant *lcm1* at the seedling stage (×40,000), *cp* chloroplast, *sg* starch grain, *pg* plastoglobule, *gl* grana lamella; scale bar, 500 nm. **e** Dry weight, each value is the mean of three independent experiments, the *error bars* represent standard error (SE) of the means

### Growth characteristics and photosynthetic pigment determination

At the third true leaf stage and every 3 days thereafter, we removed the aerial part of the wild-type “FT” and mutant *lcm1* to determine their dry weights; each measurement was repeated three times. The fresh samples were dried at 65 °C for 1 h, heated at 105 °C for 24 h, and then weighed to measure the dry weight. At the same time, the third true leaves of wild-type “FT” and mutant *lcm1* were collected; 0.1 g samples were shredded into 50-mL tubes containing 10 mL of extraction solution (80% (v/v) acetone to ethanol) and then kept in the dark for 24 h (Arnon 1949). The amount of chlorophyll a (Chla), chlorophyll b (Chlb), and carotenoid in the solution was measured at wavelengths of 663, 645, and 470 nm, respectively, using a DU® 800 ultraviolet spectrophotometer (Beckman Coulter, USA). Each measurement involved at least three biological replicates.

### Observation of the chloroplast ultrastructure

The third true leaves of wild-type “FT” and mutant *lcm1* were fixed in 4% glutaraldehyde at 4 °C for 3 days, washed with 1% phosphate buffer (four times at 15 min per wash), and then fixed in 1% osmic acid (pH 7.2) for 12 h at 4 °C. The specimens were washed three times with 1% phosphate buffer (15 min each) and then dehydrated for 15 min in a graded series of ethanol concentrations (50, 70, 80, 90, and 95%). The leaves were subsequently dehydrated with absolute ethanol and 100% acetone, embedded in epoxy resin, and sectioned with a LKB2088 ultra-thin microtome (LKBS, Sweden). After double staining with uranyl acetate and lead citrate solutions, the sections were examined under a transmission electron microscope (H-7650, Hitachi, Japan).

### Screening of the candidate genes by MutMap

A modified CTAB method was used to extract the DNA from the wild-type “FT”, mutant *lcm1* and golden leaf plants in the F<sub>2</sub> population (Murray and Thompson 1980). The mutation gene was mapped using modified MutMap method (Abe et al. 2012). The DNA of 15 chlorosis plants in the F<sub>2</sub> population was mixed equally as the DNA pool. Then, the wild-type “FT”, mutant *lcm1* and DNA pool were re-sequenced using the

HiSeq® 2500 Sequencing System from Illumina. Clean reads were obtained by removing low-quality reads from the originals using the NGSQC toolkit. The filtration standard was to remove reads with < 70% Q20 bases, reads containing *N*, and reads with lengths < 70 bp; low-sequence-quality reads at the 3' end were also removed (Dai et al. 2010). The clean reads were aligned to the reference genome (<http://brassicadb.org/brad/>) using Burrows-Wheeler Aligner (BWA) software and the SNPs were detected using SAMtools software (Li and Durbin. 2009; Li et al. 2009a). In order to ensure the reliability of SNP detection, ≥ 4 SNP reads were required with a quality value of ≥ 20. SNPs that can cause mutations were detected in the following operational order: (1) the parental wild-type “FT” SNPs were subtracted from the SNPs in the DNA pool to find those in common with the mutant and (2) from the abovementioned screening results, we found the common SNPs with mutant *lcm1* homozygous SNPs, and the mutant SNPs in the CDS region were screened.

### Analysis of candidate gene

SnEff software was used to annotate the identified SNPs (Cingolani et al. 2012). The obtained candidate SNP was used to download the sequence from the Chinese cabbage database (<http://brassicadb.org/brad/index.php>). We designed specific primers with Primer 5.0 software. The candidate SNPs were PCR-amplified, cloned, and sequenced to verify that there was a difference between the wild-type “FT” and mutant *lcm1*.

### Verification of the candidate gene

Two strategies (Sanger sequencing and KASP technology) were used to verify whether the candidate SNP co-segregated with the mutant phenotype. We used 432 recessive plants in the F<sub>2</sub> population that derived from wild-type “FT” and mutant *lcm1*. In the first strategy, all F<sub>2</sub> individuals were identified with KASP technology on the LGC genomics SNPlane™ Genotyping Platform, judging whether the individuals of the mapping group were consistent with the mutant. In the second strategy, we performed a preliminary experiment, using primers to amplify a region containing the SNP. The DNA of wild-type “FT” and mutant *lcm1* was then mixed at a ratio of 1:10, and PCR amplification was performed using the mixed DNA as a template. Then, an equal

quantity of DNA from ten of the chlorosis plants was mixed to build pools, generating 43 DNA pools in total. Primers were designed flanking the candidate SNP. The wild-type “FT”, mutant *lcm1* and 43 mixed DNA pool were used individually as templates for PCR amplification. After the purification, the amplified products were sequenced directly on the 3730XL sequencing platform.

#### Phylogenetic analysis of LCM1 homologs in plants

To investigate the phylogenetic relationships between the LCM1 (BrCHLH) protein identified in the present study and its orthologs in plants, six homologs of different plant species from six families were downloaded from the NCBI database. The species and their GenBank accession numbers are as follows: AAL79577 (*Arabidopsis thaliana*), XP\_015631165 (*Oryza sativa*), AQK47214.1 (*Zea mays*), XP\_013727263 (*Brassica napus*), NP\_001237903 (*Glycine max*), XP\_015582646.1 (*Ricinus communis*), NP\_001312642 (*Nicotiana tabacum*), and XP\_008460769 (*Cucumis melo*). The CHLH protein sequences were aligned and the resulting alignment was visualized using DNAMAN. A phylogenetic tree was created according to the neighbor-joining (NJ) method based on 1000 bootstrap replications using the MEGA5.0 program. LCM1 protein was then examined by the SMART tool (<http://smart.embl-heidelberg.de/>) to confirm the presence of the putative domain. In addition, the online software SWISS-MODEL (<https://swissmodel.expasy.org/>) was used to construct the three-dimensional structures of the proteins for wild-type “FT” and mutant *lcm1*.

#### RNA isolation and transcriptome analysis

Total RNA was isolated from the third leaf of wild-type “FT” and mutant *lcm1* using TRIzol reagent (Invitrogen, USA) according to the manufacturer’s instructions. The quality and integrity were monitored using 1% agarose gel and an Agilent 2100 Bioanalyzer (Agilent Technologies, USA), respectively. The cDNA libraries were constructed by extracting equal amounts of the total RNA from the wild-type “FT” and mutant *lcm1* plants. The library products were subjected to sequencing analysis on the Illumina HiSeq™ 2000 platform (Novogene, China). Clean reads were obtained by filtering reads

containing the adapter, poly-*N*, and low-quality reads from the raw data. The clean reads were aligned to the Brassica reference database (BRAD; <http://brassicadb.org/brad/>).

To identify differentially expressed genes (DEGs) between the two samples, the gene expression levels were calculated using the reads per kilobase per million mapped reads (RPKM) method (Mortazavi et al. 2008). DEGs were identified with the combination of a false discovery rate (FDR)  $\leq 0.001$  and the absolute value of  $\log_2(\text{ratio}) \geq 1$ . The FDR was used to adjust the *P* value in multiple tests (Benjamini and Hochberg, 1995). The Gene Ontology (GO) (Ashburner et al. 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2008) databases were used to identify putative biological functions and biochemical pathways for all DEGs. Compared with the genome background, the significantly enriched GO terms and KEGG pathways in DEGs were identified using a *P* value  $\leq 0.05$  and *Q* value  $\leq 0.05$ , respectively (Abdi, 2007).

SNPs were detected using SOAPsnp software (Li et al. 2009b), which is a member of the SOAP (Short Oligonucleotide Analysis Package). The SOAPsnp program is a re-sequencing utility that can assemble consensus sequences for the genome/transcriptome of a newly sequenced individual based on the alignment of the sequencing reads on a known reference genome. The SNPs can then be identified using the consensus sequences identified through the comparisons with the reference database.

#### Gene expression patterns analysis

Gene expression patterns were determined by quantitative real-time PCR (qRT-PCR) in order to identify the partial DEGs and candidate genes in wild-type “FT” and mutant *lcm1* leaves. The qRT-PCR reactions were conducted with 25  $\mu\text{L}$  2X UltraSYBR Mixture (CW BIO, Beijing, China), 1  $\mu\text{L}$  of each specific primer, 2  $\mu\text{L}$  diluted cDNA, and 21  $\mu\text{L}$  water, for a 50  $\mu\text{L}$  total volume. Gene-specific primers were designed using Premier 5.0, and sequences were listed in Table S1. The *actin* gene was used as an internal control. All experiments were performed on three biological replicates. The relative expression level of each gene was calculated using the  $2^{-\Delta\Delta\text{CT}}$  method. qRT-PCR was performed and analyzed on QuantStudio™ 6 Flex Real-Time PCR System (Applied Biosystems, USA).

## Results

### Identification, phenotypic observation, and genetic analysis of *lcm1* mutant

In this study, the golden leaf mutant *lcm1* was created by EMS mutagenesis of the DH line “FT” seeds of Chinese cabbage. A total of six chlorosis mutants were obtained, respectively named *lcm1*, *lcm2*, *lcm3*, *lcm4*, *lcm5*, and *lcm6*. After planting and identifying the mutants in spring and autumn, we found that mutant *lcm1* displayed a permanent golden yellow color of all above-ground organs throughout its growth (Fig. 1a), and its mutation traits were not affected by external environmental factors. The stems, buds, and pods of the mutants all appeared golden yellow to varying degrees.

The mutant *lcm1* was crossed with the wild-type “FT”, and the leaf color of the F<sub>1</sub> population plants was green. Our results revealed that 432 plants had golden yellow leaf, and 1383 plants displayed the wild-type phenotype. These findings were consistent with the 1:3 ( $\chi^2 < \chi^2_{0.05} = 3.84$ ) segregation ratio, indicating that the golden yellow leaf trait was controlled by a single recessive nuclear gene. The segregation data of the leaf color in all generations are shown in Table 1.

### Analysis of physiological properties of wild-type “FT” and mutant *lcm1*

As the plants developed, the total CHL, Chla, and Chlb in the mutant *lcm1* leaves were significantly lower than those in the wild-type “FT” leaves, which were 47.54%, 55.70%, and 18.31% of the “FT” values, respectively. However, the carotenoid content in the mutant *lcm1* leaves was slightly lower than that in the “FT” leaves (Fig. 1b). Compared with the wild-type “FT”, the

growth at the seedling stage showed that the dry weight of the *lcm1* mutant was significantly reduced (Fig. 1e). These results indicated that the reduced Chl contents and growth in *lcm1* mutant were related to the chlorophyll-deficient leaf phenotype in the mutant.

Consequently, we compared the chloroplast ultrastructure of wild-type “FT” and mutant *lcm1* to investigate how chloroplast development is affected. The chloroplast ultrastructure of the wild-type “FT” and mutant *lcm1* mutants was noticeably different. In the wild-type “FT”, the thylakoids were tightly packed and had a distinct grana structure, but the mutant *lcm1* thylakoids had singular prolamellar bodies with no obvious basal grana structure (Fig. 1c, d).

### Prediction of candidate genes and co-segregation analysis

A modified MutMap method was applied to isolate the *lcm1* gene (Abe et al. 2012). Through genome re-sequencing of the wild-type “FT”, mutant *lcm1*, and the DNA pool, 128,203,724; 153,339,460; and 162,779,565 high-quality reads were obtained, respectively, with average sequencing depths higher than 30×. Compared with the SNPs of the reference genome Chiffu (<http://brassicadb.org/brad/index.php>), 2,088,227; 2,115,201; and 2,120,392 SNPs were obtained, respectively. The SNPs from the DNA pool were deducted from all of the “FT”, leaving the SNPs associated with the mutant. According to the SNP ratio  $\geq 0.9$  and covered reads  $\geq 4$ , nine SNPs were obtained and found to be distributed on the five chromosomes of the Chinese cabbage. Nine SNPs were located on gene coding sequences; among them, eight SNPs were non-synonymous mutations, and one caused a splice region variant. We designed primers on the basis of the mutation information of these nine SNPs (Table S2), using

**Table 1** Genetic analysis of the leaf chlorosis mutant *lcm1* in Chinese cabbage

Generation	Green-colored plants	Yellow-colored plants	Total	Segregation ratio	$\chi^2$ test
P <sub>1</sub> (“FT”)	96	0	96		
P <sub>2</sub> ( <i>lcm1</i> )	0	73	73		
F <sub>1</sub> (P <sub>1</sub> × P <sub>2</sub> )	158	0	158		
F <sub>1</sub> (P <sub>2</sub> × P <sub>1</sub> )	172	0	172		
BC <sub>1</sub> (F <sub>1</sub> × “FT”)	454	0	454		
BC <sub>1</sub> (F <sub>1</sub> × <i>lcm1</i> )	221	215	436	1.03: 1	0.06 (1:1)
F <sub>2</sub>	1383	432	1815	3.20: 1	1.39 (3:1)

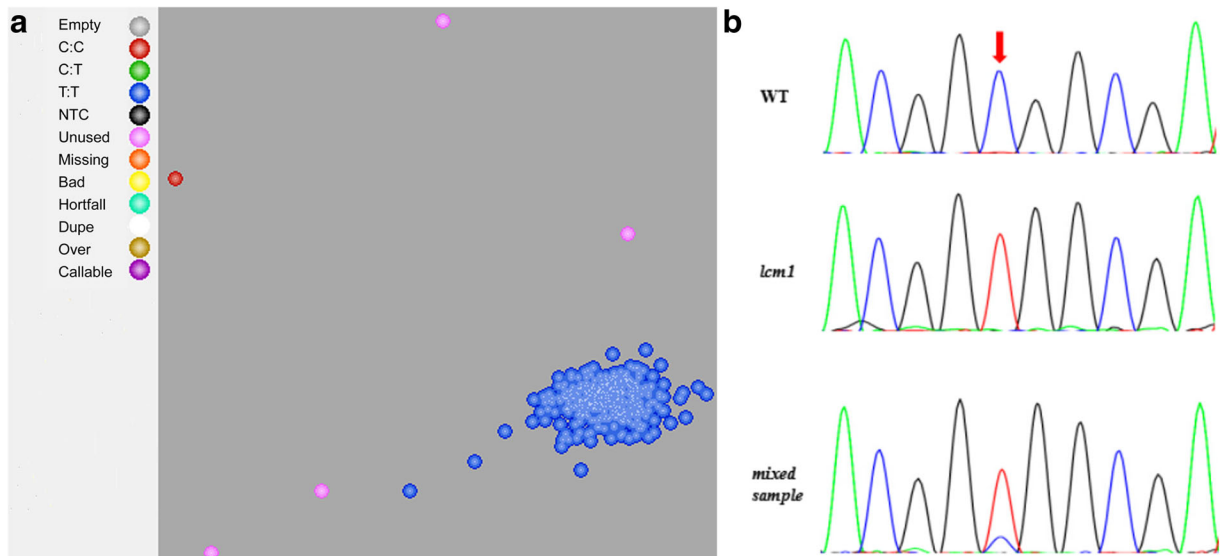
**Table 2** Three candidate SNPs information

Chromosome	SNP position	Reference	Wild type	Mutant	Mutation type	Gene ID	Function
A03	1662998	G	G	A	Missense	Bra006027	Mitochondrial HSP70
A03	2523351	G	G	A	Missense	Bra006208	Magnesium chelatase subunit chlH
A03	3516344	G	G	A	Splice region variant	Bra006446	Transcription regulator NOT2/NOT3/NOT5 family protein

wild-type “FT” and mutant *lcm1* genomic DNA as templates and directly sequenced using PCR production. The sequencing results showed that six SNPs had no difference between wild-type “FT” and mutant *lcm1*, while the other three SNPs might be the candidate SNPs responsible for mutant *lcm1* in Chinese cabbage (Table 2). Of the three SNPs (SNP1662998, SNP2523351, and SNP3516344), only SNP2523351 was located in the Chinese cabbage gene *Bra006208*, which was involved in CHL biosynthesis and encoded a Mg-chelatase H subunit. Therefore, we speculate that *Bra006208* may be the candidate gene for the mutant *lcm1*, and the SNP can be related to the golden leaf phenotype.

To further verify the causal SNP of mutant *lcm1*, we used 432 recessive plants in F<sub>2</sub> population to determine whether the SNP was co-segregated with the mutant phenotype. Two strategies were used to verify that the

SNP2523351 resulted in the golden leaf phenotype. In the first strategy, we used KASP technology to genotype individual plants in F<sub>2</sub> segregation population and revealed that SNP2523351 was T:T genotype in the 432 chlorosis plants (Fig. 2a), which was consistent with the mutants. However, 34 and 48 recombinations were found at the two other SNPs (SNP1662998 and SNP3516344), respectively. In order to verify the accuracy of KASP, we sequenced the PCR products with the recombinant plants’ DNA as templates and found that the results were consistent with that of KASP. In the second verification strategy, we used direct sequencing of PCR products to detect the genotype in the F<sub>2</sub> population. The preliminary experiment results clearly showed that the sequencing peak at this site included both C and T bases (Fig. 2b). Therefore, the DNA from each of the 43 pools was amplified by PCR with the abovementioned primers. The amplified products were



**Fig. 2** Verification of the candidate gene by F<sub>2</sub> population. **a** Detection of 432 genotypes of the chlorosis plants. KASP technology was used to genotype individual plants and found that candidate SNP was T:T genotype in the 432 chlorosis plants. **b**

DNA sequencing peak map of wild-type “FT” and mutant *lcm1* mixed pool, sample DNA of wild-type “FT” and mutant *lcm1* was mixed at a ratio of 1:10. The red arrow indicates the mutation site

sequenced directly. All 43 sequences had nucleobases at the SNP that were identical to the mutant (i.e., both T (thymine)), which means that the SNP2523351 co-segregated with the mutant phenotype. Overall, according to the MutMap method and PCR sequencing results, *Bra006208 (lcm1, BrChlH)*, encoded Mg-chelatase H subunit that harbored the SNP2523351, was identified as the causative mutation for *lcm1*, which had a single nucleotide substitution (G to A) in its third exon (2,523,351 on chromosome A03). *lcm1* gene contains four introns and five exons, with an open reading frame (ORF) length of 4149 bp and encodes a mature protein of 1382 amino acids. The Pfam domain CobN-Mg\_chel (M4CPR9) was from the 244th to 1363rd amino acid (Fig. 3a) by SMART tool, and the SNP2523351 was located at the 1001 amino acid (A to V). The SNP2523351 was determined by cloning and sequencing the wild-type “FT” and mutant *lcm1*.

Phylogenetic analysis among the LCM1 and its homologs

To better understand the relation between LCM1 protein and its closest homologs, an expanded alignment and a phylogenetic analysis were performed using *Brassica rapa* and eight other species. The protein is highly homologous to Mg-chelatase subunit H from other plants, such as *Arabidopsis thaliana* (94% homology), *Oryza sativa* (83%), *Zea mays* (82%), *Brassica napus* (99%), *Glycine max* (87%), *Ricinus communis* (88%), *Nicotiana tabacum* (86%), and *Cucumis melo* (86%)

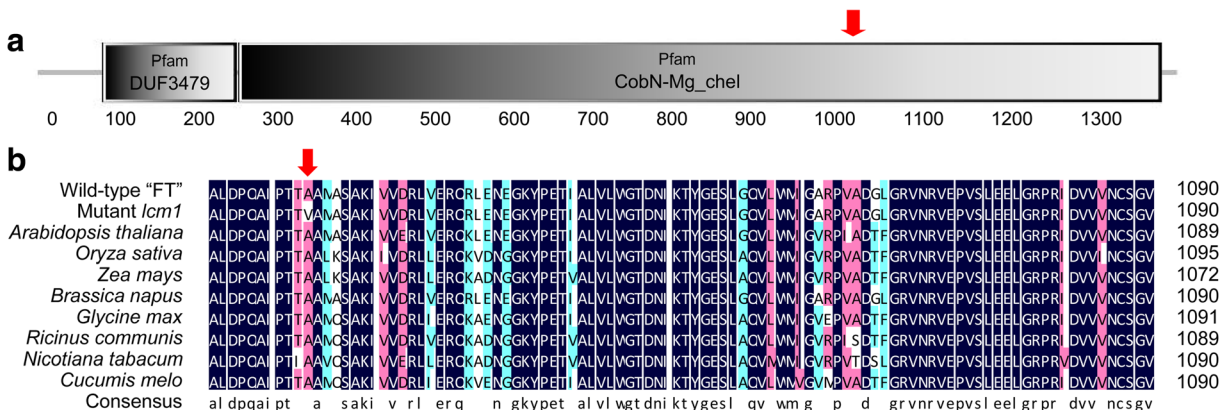
(Fig. 3b). We identified LCM1 protein that possessed the conserved CobN-Mg\_chel domain and the amino acid where the non-synonymous mutation was located in the conserved domain (Fig. 3a). To gain insights into the phylogenetic analysis among LCM1 protein and other homologs, the resulting phylogenetic tree showed that LCM1 protein was grouped together with the homologs from other Cruciferae species, suggesting that they may have a similar function (Fig. S1). In addition, we found the three-dimensional structures of the proteins between “FT” and mutant *lcm1* were different at the mutant site (Fig. 4).

Expression analysis of *lcm1* gene

In our study, we used qRT-PCR to detect the expression level of *lcm1* gene in Chinese cabbage leaves and found that there were no significant differences in the expression between the wild-type and mutant.

Transcriptional analysis of wild-type and mutant leaves

The transcriptome is consistent with the previous qRT-PCR results. In order to analyze the changes in the transcriptome caused by the presence of the mutant gene, we used the next-generation sequencing (NGS) technology to re-sequence the wild-type “FT” and mutant *lcm1* leaf transcriptomes, using three replicates per sample. Then, 50 M reads were detected each in wild-type “FT” and mutant *lcm1*, and approximately 50% of these reads could be mapped to the genome. There were

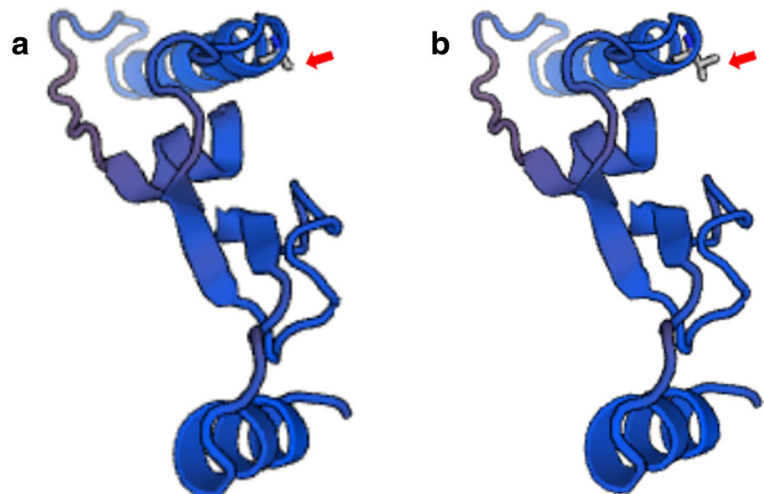


**Fig. 3** Analysis of Mg-chelatase H subunit protein sequence. **a** CobN-Mg\_chel domain in LCM1 protein from 244th to 1363rd amino acid, and the mutant amino acid was located in the conserved domain; the red arrow indicates the location of the mutated amino acid. **b** Alignment of wild-type “FT” and mutant *lcm1* Mg-

chelatase H subunit protein sequences with 8 species; the protein is highly homologous to Mg-chelatase subunit H from other plants. Single amino acid substitutions resulting in leaf color mutation in different species are pink block-typed



**Fig. 4** Three-dimensional structure of wild-type “FT” and mutant *lcm1* ChlH protein sequences. Since the other sequences are identical, we constructed three-dimensional structures of 90 amino acid sequences including the SNPs, respectively. The red arrow represents the mutant amino acid. **a** A in wild-type “FT”. **b** V in mutant *lcm1*



27,827, 27,753 and 27,852 expressed genes detected in the mutant *lcm1* leaves, and 27,929; 27,967; and 28,016 expressed genes detected in the wild-type “FT”. The screening standard for DEGs was  $FDR \leq 0.001$ , and the difference in multiples was more than two times. Finally, 606 DEGs (wild-type “FT” vs mutant *lcm1*) were identified in the two samples, among which 308 genes were upregulated and 298 were downregulated. These genes were enriched for those involving biosynthesis of secondary metabolites, carbon fixation in photosynthetic organisms, phenylpropanoid biosynthesis, metabolic pathways, and microbial metabolism in diverse environments.

## Discussion

In this study, the golden leaf mutant *lcm1* was used for genetic identification and fine mapping. The mutant has stable genetic characteristics and exhibits a golden leaf color throughout the growth period (Fig. 1a). The growth characteristics and total Chl content of mutant *lcm1* were significantly lower than those of “FT” (Fig. 1b, e) and the chloroplast ultrastructure was significantly different from “FT” (Fig. 1c). However, the reduction in the chlorophyll content did not result in an equivalent reduction in growth for the mutant, indicating that the growth of the mutants was not seriously affected. The plant growth rate of mutant *lcm1* was slower than that of wild-type “FT” for the entire growing period. These observations were consistent with other chlorophyll deficiency mutants (Gao et al. 2016).

Therefore, we think that the golden leaf of the chlorosis mutant *lcm1* is due to a decrease in chlorophyll content.

NGS technology detects SNPs with high throughput and high accuracy, laying the foundations for the cloning of EMS mutagenesis genes (Deng et al. 2017). Abe et al. (2012) proposed the MutMap cloning strategy to rapidly clone mutated genes. EMS mutagenesis combining NGS-based mapping is an effective way to identify important genes. We sequenced the DNA pool of recessive plants in an  $F_2$  population using the MutMap method based on NGS and then compared the results with the whole genome sequence of the wild-type combination. Directly sequencing of PCR products showed that only three SNPs (SNP1662998, SNP2523351, and SNP3516344) located on chromosome A03 had real sites. Furthermore, the genotyping of the  $F_2$  population was detected by KASP technology. We found that there were no recombinants at the *lcm1* gene detection SNP2523351 among the genotyped 432 chlorosis plants (Fig. 2a), but the other two SNPs (SNP1662998 and SNP3516344) had recombinations. Besides, we identified LCM1 protein that possessed the conserved CobN-Mg\_chel domain and the mutant amino acid was located in the conserved domain (Fig. 3a). In addition, the causal SNP can be identified by transcriptome sequencing to further confirm the SNP; the splicing quality of most plant genome sequences is difficult to achieve so the reference genomes of model plants (such as *Arabidopsis*) are important. Some of the SNPs of the mutations may not be on the spliced chromosomes but exist on scaffolds that are not spliced to the chromosome. The method used in this study did not impose excessive requirements on the splicing quality of the

reference genome, which expanded the scope of use with the method. In this study, three SNPs (SNP1662998, SNP2523351, and SNP3516344) met the screening requirements, and, with the help of 432 recessive plants in F<sub>2</sub> population, one causal SNP2523351 was finally identified, demonstrating that this method is effective and rapid. This provides a beneficial tool for the cloning of plant EMS mutagenesis genes. These findings also provide a solid basis for *lcm1* gene cloning, identification, and functional analysis.

The biosynthesis of CHL in higher plants is a very complex enzymatic process. Thus far in the model plant *Arabidopsis thaliana*, studies have shown that the plant CHL biosynthesis pathway is a 15-step biochemical reaction that involves the participation of 15 enzymes encoded by 27 genes (Beale 2005; Mochizuki et al. 2010). Certainly, mutant genes in different places in the pathway can produce different phenotypes. For example, inhibition of CHL synthesis in the temperature-sensitive albino mutant “white pigeon” in kale occurs in the synthesis of NADPH-protochlorophyllide (Pchl) from Mg-protoporphyrin IX (Zhou et al. 2013); the production of the yellow-green leaf mutant, *yg11*, in rice is due to the inhibition of CHL synthesis during the formation of CHL from chlorophyllide (Wu et al. 2007); and the rice mutant, *etiolated leaf and lethal (ell)*, caused by a mutation in the magnesium chelatase gene, *OsCHLI*, displays yellow-colored leaves in young seedlings, which occurs from a disruption in protoporphyrin IX to Mg-protoporphyrin IX conversion (Zhang et al. 2015). Mg-chelatase is a key enzyme in CHL synthesis and is responsible for catalyzing the chelation of Mg<sup>2+</sup> into protoporphyrin IX to form Mg-protoporphyrin IX (Tanaka and Tanaka, 2007; Stenbaek and Jensen 2010; Tanaka et al. 2011). Mg-chelatase is an ATP-dependent heterologous polymerase composed of three subunits, CHLI, CHLD, and CHLH (Jensen et al. 1999; Fodje et al. 2001); mutations in any of these three subunits can cause yellow traits. The CHLH subunit is thought to be the catalytic subunit of Mg-chelatase that can bind the substrate, protoporphyrin IX (Reid and Hunter 2002). In a cucumber, a single point mutation in the CHLI subunit results in the production of golden leaf color (Gao et al. 2016), which was concluded after NGS re-sequencing, large F<sub>2</sub> population validation, and transcriptome sequencing indicated that the SNP was homozygous. The leaves of some horticultural plants are golden yellow, which could be explained by the modification of this CHLH

in these species. The color change can be achieved by mutating a single amino acid, which makes it a valuable target. The resources developed in this study provide a basis for further studies of gene expression and functional genomics in Chinese cabbage.

**Acknowledgements** The authors are grateful to the editors and the anonymous reviewers for their valuable comments.

**Funding information** This work was supported by the National Key Research and Development Program of China (2016YFD0101701) and National Natural Science Foundation of China (31772298).

## References

- Abdi H (2007) Bonferroni and Šidák corrections for multiple comparisons. In: Salkind N (ed) Encyclopedia of Measurement and Statistics. Thousand Oak, pp 103–107
- Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Muluneh T, Innan H, Cano L, Kamoun S, Terauchi R (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30:174–178
- Arnon DI (1949) Copper enzymes in isolated chloroplasts: polyphenoloxidase in *Beta vulgaris*. *Plant Physiol* 24:1–15
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. *Nat Genet* 25:25–29
- Beale SI (2005) Green genes gleaned. *Trends Plant Sci* 10:309–312
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 57:289–300
- Braumann I, Stein N, Hansson M (2014) Reduced chlorophyll biosynthesis in heterozygous barley magnesium chelatase mutants. *Plant Physiol Biochem* 78:10–14
- Chen H, Cheng ZJ, Ma XD, Wu H, Liu YL, Zhou KN, Chen YL, Ma WW, Bi JC, Zhang X, Guo XP, Wang JL, Lei CL, Wu FQ, Lin QB, Liu YQ, Liu LL, Jiang L (2013) A knockdown mutation of *YELLOW-GREEN LEAF2* blocks chlorophyll biosynthesis in rice. *Plant Cell Rep* 32:1855–1867
- Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster strain w<sup>1118</sup>; iso-2; iso-3*. *Fly (Austin)* 6:80–92
- Colombo N, Emanuel C, Lainez V, Maldonado S, Prina AR, Bömer T (2008) The barley plastome mutant CL2 affects expression of nuclear and chloroplast housekeeping genes in a cell-age dependent manner. *Mol Gen Genomics* 279:403–414
- Dai M, Thompson RC, Maher C, Contreras-Galindo R, Kaplan MH, Markovitz DM, Omenn G, Meng F (2010) NGSQC: cross-platform quality analysis pipeline for deep sequencing data. *BMC Genomics* 11:S7

- Deng LC, Qin P, Liu Z, Wang GL, Chen WL, Tong JH, Xiao LT, Tu B, Sun YT, Yan W, He H, Tan J, Chen XW, Wang YP, Li SG, Ma BT (2017) Characterization and fine-mapping of a novel premature leaf senescence mutant *yellow leaf and dwarf 1* in rice. *Plant Physiol Biochem* 111:50–58
- Dong H, Fei GL, Wu CY, Wu FQ, Sun YY, Chen MJ, Ren YL, Zhou KN, Cheng ZJ, Wang JL, Jiang L, Zhang X, Guo XP, Lei CL, Su N, Wang H, Wan JM (2013) A rice *virescent-yellow leaf* mutant reveals new insights into the role and assembly of plastid caseinolytic protease in higher plants. *Plant Physiol* 162:1867–1880
- Fekih R, Takagi H, Tamiru M, Abe A, Natsume S, Yaegashi H, Sharma S, Sharma S, Kanzaki H, Matsumura H, Saitoh H, Mitsuoka C, Utsushi H, Uemura A, Kanzaki E, Kosugi S, Yoshida K, Cano L, Kamoun S, Terauchi R (2013) MutMap+: genetic mapping and mutant identification without crossing in rice. *PLoS One* 8:e68529
- Fodje MN, Hansson A, Hansson M, Olsen JG, Gough S, Willows RD, Al-Karadaghi S (2001) Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase. *J Mol Biol* 311:111–122
- Gao ML, Hu LL, Li YH, Weng YQ (2016) The chlorophyll-deficient *golden leaf* mutation in cucumber is due to a single nucleotide substitution in *CsChlI* for magnesium chelatase I subunit. *Theor Appl Genet* 129:1961–1973
- Hao N, Du Y, Li H, Wang C, Wang C, Gong S, Zhou S, Wu T (2018) *CsMYB36* is involved in the formation of yellow green peel in cucumber (*Cucumis sativus* L.). *Theor Appl Genet* 131:1659–1669
- Ihnatowicz A, Siwinska J, Meharg AA, Carey M, Koomneef M, Reymond M (2014) Conserved histidine of metal transporter AtNRAMP1 is crucial for optimal plant growth under manganese deficiency at chilling temperatures. *New Phytol* 202: 1173–1183
- Imamura T, Takagi H, Miyazato A, Ohki S, Mizukoshi H, Mori M (2018) Isolation and characterization of the betalain biosynthesis gene involved in hypocotyl pigmentation of the allotetraploid *Chenopodium quinoa*. *Biochem Biophys Res Commun* 496:280–286
- Jensen PE, Gibson LC, Hunter CN (1999) ATPase activity associated with the magnesium- protoporphyrin IX chelatase enzyme of *Synechocystis* PCC6803: evidence for ATP hydrolysis during Mg<sup>2+</sup> insertion, and the MgATP-dependent interaction of the ChlI and ChlD subunits. *Biochem J* 339: 127–134
- Jiao Y, Burrow G, Gladman N, Acosta-Martinez V, Chen J, Burke J, Ware D, Xin Z (2018) Efficient identification of causal mutations through sequencing of bulked F<sub>2</sub> from two allelic bloomless mutants of *Sorghum bicolor*. *Front Plant Sci* 8:2267
- Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamaniishi Y (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36:D480–D484
- Kusaba M, Ito H, Morita R, Iida S, Sato Y, Fujimoto M, Kawasaki S, Tanaka R, Hirochika H, Nishimura M, Tanaka A (2007) Rice NON-YELLOW COLORING1 is involved in light-harvesting complex II and grana degradation during leaf senescence. *Plant Cell* 19:1362–1375
- Kusumi K, Sakata C, Nakamura T, Kawasaki S, Yoshimura A, Iba K (2011) A plastid protein NUS1 is essential for build-up of the genetic system for early chloroplast development under cold stress conditions. *Plant J* 68:1039–1050
- Lee S, Lee SC, Byun DH, Lee DY, Park JY, Lee JH, Lee HO, Sung SH, Yang TJ (2014) Association of molecular markers derived from the *BrCRTISO1* gene with prolycopene-enriched orange-colored leaves in *Brassica rapa*. *Theor Appl Genet* 127:179–191
- Li H, Durbin R (2009) Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25:1754–1760
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009a) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25:2078–2079
- Li RQ, Yu C, Li YR, Lam TW, Yiu SM, Kristiansen K, Wang J (2009b) SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25:966–1967
- Mamidi S, Lee RK, Goos JR, McClean PE (2014) Genome-wide association studies identifies seven major regions responsible for iron deficiency chlorosis in soybean (*Glycine max*). *PLoS One* 9:e107469
- Mochizuki N, Tanaka R, Grimm MT, Moulin M, Smith AG, Tanaka A, Terry MJ (2010) The cell biology of tetrapyrroles: a life and death struggle. *Trends Plant Sci* 15:488–498
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNASeq. *Nat Methods* 5:621–628
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8:4321–4325
- Pogson BJ, Albrecht V (2011) Genetic dissection of chloroplast biogenesis and development: an overview. *Plant Physiol* 155: 1545–1551
- Reid JD, Hunter CN (2002) Current understanding of the function of magnesium chelatase. *Biochem Soc Trans* 30:643–645
- Shi DY, Zheng X, Li L, Lin WH, Xie WJ, Yang JP, Chen SJ, Jin WW (2013) Chlorophyll deficiency in the maize *elongated mesocotyl2* mutant is caused by a defective heme oxygenase and delaying grana stacking. *PLoS One* 8:e80107
- Sozer O, Komenda J, Ughy B, Domonkos I, Laczkó-Dobos H, Malec P, Gombos Z, Kis M (2010) Involvement of carotenoids in the synthesis and assembly of protein subunits of photosynthetic reaction centers of *synechocystis* sp. PCC 6803. *Plant Cell Physiol* 51:823–835
- Stenbaek A, Jensen PE (2010) Redox regulation of chlorophyll biosynthesis. *Phytochemistry* 71(8–9):853–859
- Takagi H, Uemura A, Yaegashi H, Tamiru M, Abe A, Mitsuoka C, Utsushi H, Natsume S, Kanzaki H, Matsumura H, Saitoh H, Yoshida K, Cano LM, Kamoun S, Terauchi R (2013) MutMap-Gap: whole-genome resequencing of mutant F<sub>2</sub> progeny bulk combined with *de novo* assembly of gap regions identifies the rice blast resistance gene *Pii*. *New Phytol* 200:276–283
- Takagi H, Tamiru M, Abe A, Yoshida K, Uemura A, Yaegashi H, Obara T, Oikawa K, Utsushi H, Kanzaki E, Mitsuoka C, Natsume S, Kosugi S, Kanzaki H, Matsumura H, Urasaki N, Kamoun S, Terauchi R (2015) MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nat Biotechnol* 33:445–449
- Tanaka R, Tanaka A (2007) Tetrapyrrole biosynthesis in higher plants. *Plant Biol* 58:321–346
- Tanaka R, Kobayashi K, Masuda T (2011) Tetrapyrrole metabolism in *Arabidopsis thaliana*. *Arabidopsis Book /Am Soc Plant Biol* 9:e0145

- Terry MJ, Kendrick RE (1999) Feedback inhibition of chlorophyll synthesis in the phytochrome chromophore-deficient *aurea* and *yellow-green-2* mutants of tomato. *Plant Physiol* 119: 143–152
- Wachsman G, Modliszewski JL, Valdes M, Benfey PN (2017) A SIMPLE pipeline for mapping point mutations. *Plant Physiol* 174:1307–1313
- Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet* 43: 1035–1039
- Wang H, Zhang Y, Sun L, Xu P, Tu R, Meng S, Wu W, Anis GB, Hussain K, Riaz A, Chen D, Cao L, Cheng S, Shen X (2018) *WBI*, a regulator of endosperm development in Rice, is identified by a modified MutMap method. *Int J Mol Sci* 19: 2159
- Wu ZM, Zhang X, He B, Diao LP, Sheng SL, Wang JL, Guo XP, Su N, Wang LF, Jiang L, Wang CM, Zhai HQ, Wan JM (2007) A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. *Plant Physiol* 145:29–40
- Xu J, Yang J, Wu Z, Liu H, Huang F, Wu Y, Carrie C, Narsai R, Murcha M, Whelan J, Wu P (2013) Identification of a dual-targeted protein belonging to the mitochondrial carrier family that is required for early leaf development in rice. *Plant Physiol* 161:2036–2048
- Xu L, Wang C, Cao W, Zhou S, Wu T (2018) *CLAVATA1*-type receptor-like kinase *CsCLAVATA1* is a putative candidate gene for dwarf mutation in cucumber. *Mol Gen Genomics* 293:1393–1405
- Zhang H, Liu LL, Cai MH, Zhu SS, Zhao JY, Zheng TH, Xu XY, Zeng ZQ, Niu J, Jiang L, Chen SH, Wan JM (2015) A point mutation of magnesium chelatase *OsCHL1* gene dampens the interaction between CHLI and CHLD subunits in rice. *Plant Mol Biol Report* 33:1975–1987
- Zhou KN, Ren YL, Lv J, Wang YH, Liu F, Zhou F, Zhao SL, Chen SH, Peng C, Zhang X, Guo XP, Cheng ZJ, Wang JL, Wu FQ, Jiang L, Wan JM (2012) *Young Leaf Chlorosis 1*, a chloroplast-localized gene required for chlorophyll and lutein accumulation during early leaf development in rice. *Planta* 237:279–292
- Zhou S, Hu ZL, Zhu MK, Zhang B, Deng L, Pan Y, Chen GP (2013) Biochemical and molecular analysis of a temperature-sensitive albino mutant in kale named “white dove”. *Plant Growth Regul* 71:281–294
- Zou T, Xiao Q, Li WJ, Luo T, Yuan GQ, He ZY, Liu MX, Li Q, Xu PZ, Zhu J, Liang YY, Deng QM, Wang SQ, Zheng AP, Wang LX, Li P, Li SC (2017) *OsLAP6/OsPKS1*, an orthologue of *Arabidopsis* *PKSA/LAP6*, is critical for proper pollen exine formation. *Rice* 10:53

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.