Vegetable Research

ARTICLE

Open Access

https://doi.org/10.48130/vegres-0025-0020

Vegetable Research **2025**, 5: e027

Insertion of a large retrotransposon TNT fragment in the promoter region of *Brgl1* results in a gray leaf phenotype in pakchoi

Chuanhong Liu, Kai Guo, Bo Shang, Jie Ren, Chong Tan, Yun Zhang* and Zhiyong Liu*

Laboratory of Vegetable Genetics Breeding and Biotechnology, Department of Horticulture, Shenyang Agricultural University, No. 120 Dongling Road, Shenhe District, Shenyang 110866, China

* Corresponding authors, E-mail: zhangyun511@syau.edu.cn; liuzhiyong99@syau.edu.cn

Abstract

In plants, pentatricopeptide repeat (PPR) proteins perform post-transcriptional modifications on RNA and regulate photosynthesis. In this study, a natural gray leaf mutant, *M579*, was identified during the self-breeding process of the pakchoi (*Brassica campestris* L. ssp. *chinensis*) double haploid (DH) line '579'. The chlorophyll content and net photosynthetic rate of *M579* were lower than those of '579'. Mutant *M579* exhibited defective chloroplast development, characterized by a reduced number of grana stacks. Genetic analysis indicated that the gray leaf phenotype was controlled by a single pair of recessive nuclear genes, subsequently named *Brgl1*. Using bulked segregant analysis sequencing, *Brgl1* was initially mapped to chromosome A06. Further fine mapping using 1,877 recessive homozygous individuals from the F_2 population, together with a map-based cloning approach, localized *Brgl1* between markers SIN-33 and SIN-34. The target region spanned a physical distance of 8.069 kb and contained four genes. Further cloning revealed a 4,788 bp insertion in the promoter region of *BraA06g036440.3C* within the mapped interval, and this inserted fragment was identified as the retrotransposon TNT 1-94, suggesting that *BraA06g036440.3C* was the candidate gene for *Brgl1*. *BraA06g036440.3C* belonged to the P subfamily of the PPR gene family, encoding a hydrophilic, non-secretory protein with seven tandemly repeated PPR motifs. GUS activity analysis showed that *Brgl1* was primarily expressed in the leaves. The relative expression level of *Brgl1* was reduced in *M579* leaves. This study provides significant information regarding the molecular mechanisms underlying leaf color mutations in pakchoi.

Citation: Liu C, Guo K, Shang B, Ren J, Tan C, et al. 2025. Insertion of a large retrotransposon TNT fragment in the promoter region of *Brgl1* results in a gray leaf phenotype in pakchoi. *Vegetable Research* 5: e027 https://doi.org/10.48130/vegres-0025-0020

Introduction

Leaf color mutations are prevalent in the plant kingdom and have been observed in a diverse range of plant species, including *Arabidopsis thaliana*[1,2], maize[3], barley[4], Chinese cabbage[5], strawberry[6], and rice[7,8]. Leaf color mutations can be utilized to rapidly identify varietal purity in crop hybrid breeding, and to determine the mechanisms of photosynthesis, chlorophyll synthesis pathways, and gene regulatory networks in plants[9]. The occurrence of leaf color mutants is primarily attributed to mutations in chlorophyll synthesis genes, abnormalities in heme metabolism, and defective chloroplast development. Chlorophyll biosynthesis is tightly regulated by a network of proteins, and mutations in any of their encoding genes could potentially disrupt chlorophyll synthesis, leading to a decrease in total pigment content or an imbalance in pigment ratios within cells. In soybean, a mutation in the gene encoding the Mg-chelatase subunit (ChlI1a) resulted in reduced Mg-chelatase activity, which inhibited chlorophyll biosynthesis, consequently causing leaf chlorosis[10]. Mutations in the maize *cpx* gene, encoding the protochlorophyllidox oxidoreductase enzyme, which catalyzes chlorophyll biosynthesis, led to leaf yellowing[11]. Heme is metabolized to form the chromophore of photosensitive pigments[12]. The tomato mutants *Au* and *yg-2* experienced inhibited synthesis of linear tetrapyrrole chromophores, resulting in abnormal heme metabolism and subsequent heme accumulation feedback, leading to a lightening of their leaf color[13]. The chloroplast serves as the primary site for chlorophyll synthesis and storage in plants, while also being the site for photosynthesis. Therefore, abnormal chloroplast development is one of the contributing factors to plant leaf color mutations. The cucumber pale yellow leaf mutant *Vyl* was mapped, and the gene *CSVyl* was cloned, which had a mutation resulting in a significant delay in chloroplast development[14]. The mutant gene *ygl8*, localized and cloned in a rice yellow-green leaf mutant, had a significantly reduced chlorophyll content compared with that in the wild-type, possibly resulting from an indirect influence on chlorophyll synthesis through its impact on the grana stacking structure[15]. To date, many studies have focused on leaf yellowing mutations, with relatively limited investigations into the gray leaf phenotype.

Chloroplasts are semi-autonomous organelles that originated from a primary endosymbiotic event between a eukaryotic host and a photosynthetic cyanobacterium[16], which contain their own gene translation machinery and genome and are the sole organelles responsible for photosynthesis in algae and plants[17]. Chloroplast genes are categorized into three main functional classes: those involved in photosynthesis, those involved in gene expression, and those involved in the biosynthesis of fatty acids and other compounds (such as ATP synthase subunits, initiation factors, and NADH dehydrogenase subunits)[18]. Chloroplasts serve as valuable materials for research in plant physiology, biochemistry, and genetics.

Pentatricopeptide repeat (PPR) proteins constitute a large and diverse family of RNA-binding proteins found in eukaryotes, playing crucial roles in RNA splicing, maturation, stability, editing, and translation[19]. Based on the types and arrangements of PPR protein motifs, plant PPR proteins are classified into two major subclasses: P and PLS. The P subclass consists of classic 35-amino acid motifs and primarily participates in the transcriptional regulation of organellar genes. The PLS subclass, composed of classic (P), long (L), and short (S) motifs, mainly edit and modify the RNA transcripts of organellar

genes[20]. Several *PPR* genes have been reported to be associated with plant leaf color. The RNA interference (RNAi) lines of the P-type PPR protein ECD2 in *Arabidopsis thaliana* exhibited an albino cotyledon phenotype[21]. In rice, the *YLWS* gene, encoding a P-type PPR protein, was identified as the causative gene of the juvenile leaf white stripe (ylws) mutant[22]. A PLS-type PPR protein, GmPGL2, was identified in soybean, and *GmPGL2* mutant plants exhibited pale green leaves[23]. The maize *qKW9* mutant exhibited reduced ear size and kernel number. The *qKW9* gene encodes a PLS-type PPR protein involved in photosynthetic C-to-U editing of the chloroplastic *ndhB* gene[24]. In cabbage, the causative gene of the yellow-green leaf mutant *4036Y* was identified as *BoYgl-2*, which encodes a novel nuclear-targeted P-type PPR protein[25]. However, reports of *PPR* genes associated with leaf color in pakchoi (*Brassica campestris* L. ssp. *chinensis*) are limited.

In this study, a novel gray leaf mutant *M579* of pakchoi was identified. The mutant exhibited reduced chlorophyll content and net photosynthetic rate, and impaired chloroplast development. Using BSA-seq and cloning, *Brgl1* (*BraA06g036440.3C*) was predicted as the candidate gene. An insertion of 4,788 bp was identified in its promoter region, and this inserted fragment was characterized as the retrotransposon TNT 1-94. *BraA06g036440.3C* belonged to the P subfamily of the *PPR* gene family and encoded a hydrophilic non-secretory protein. GUS activity analysis indicated that *Brgl1* was mainly expressed in leaves. The above findings provided insights into unraveling the molecular mechanisms underlying leaf color formation in pakchoi.

Materials and methods

Plant materials

The wild-type material used was pakchoi DH line '579', which was obtained from the Vegetable Genetics, Breeding, and Biotechnology Laboratory at the College of Horticulture, Shenyang Agricultural University (Shenyang, China). The gray leaf mutant *M579* was a natural mutant identified in the DH line '579' during the process of self-crossing breeding. All the materials were planted in the experimental base of the vegetable genetic breeding and biotechnology research group of Shenyang Agricultural University.

Investigation of agronomic characters

Two hundred seeds of '579' and *M579* were sown, respectively. When the third true leaf sprouted (September 20), the leaf length and width of the third true leaves of '579' and *M579*, as well as the plant width and height, were measured separately using a scale.

Three plants of '579' and *M579* were selected from the same period, and the leaves and roots of the plants were washed with water (taking care not to damage the root hairs). The water was wiped off the plant surface and roots using paper, the weight of each material was measured accurately using an electronic balance, and the fresh weight was recorded. The samples were then placed into paper bags, labeled, and put into an oven set at 65 °C for 24 h of drying. The dry weight was then recorded. The fresh and dry weights were measured every 5 d, five times in total.

Determination of the photosynthetic pigment content

Three plants of '579' and *M579* were used to measure and compare their photosynthetic pigment content. The extraction solution was prepared (acetone : absolute ethanol = 2:1). The sixth true leaves of the two groups of plants were selected, 0.2 g was weighed, the veins were removed, the leaves were cut into filaments using scissors, and then packed into 50 mL test tubes. Then, 20 mL of extraction solution was added to each test tube, which was placed in a dark room for 24 h. Ethanol acetone extract was used as the blank control. A spectrophotometer was used to measure the optical densities (ODs) of the samples at wavelengths of 663, 645, and 470 nm, respectively. Each sample was measured three times, and the average value was recorded. The content of chlorophyll a, chlorophyll b, and carotenoids was calculated based on a previously published method[26].

Determination of photosynthetic parameters

Three wild-type '579' and three mutant *M579* plants were selected, and the photosynthesis parameters were measured from 9:00–11:00 on a sunny morning. The sixth fully developed true leaf of each plant was taken as the measuring object, and the stomatal conductance (G_s), net photosynthetic rate (P_n), transpiration rate (T_r), and intercellular CO_2 concentration (C_i) of the plants were measured using an LI-6800 portable photosynthetic apparatus (Licor, Lincoln, NE, USA). Each material was measured three times, and the average value was calculated.

Scanning electron microscope observation

The sixth true leaf of wild-type '579' and mutant *M579* were selected for measurement. The leaves were rinsed twice with distilled water and dried. A segment with a diameter of 1 cm from the middle of the blade was taken out using a perforator, soaked in 100 mL of FAA fixative (70% ethanol : glacial acetic acid : formaldehyde = 90:5:5), and placed in a 4 °C refrigerator for 24 h. The fixed leaves were soaked in 50% alcohol solution, 75% alcohol solution, and 90% alcohol solution for 5 min respectively, then the leaves were immersed in 100% absolute ethanol three times, for 5 min each time to remove the residual water. Then, 50%, 75%, 90%, and 100% tertbutyl alcohol solution were used for gradient ethanol removal. The treated leaves were placed into a small metal beaker (pre-cooled at −80 °C in the refrigerator), and a small amount of 100% tert butyl alcohol (pre-melted at 65 °C in the incubator) was added. The leaves were then placed in a dryer for drying, and pasted on a clean sample table with double-sided adhesive tape. Vacuum spraying was carried out in an ion-sputtering instrument. Then, the leaf surface cells were detected using scanning electron microscopy (SEM) (Hitachi TM3030, Tokyo, Japan).

Ultrastructural observation of chloroplasts

The fourth true leaves of wild-type '579' and mutant *M579* were selected, and cut using scissors into 2–3 mm wide, 1 cm long filaments (avoiding the leaf veins during cutting). Sample processing and experimental methods were carried out according to a previous study[27]. The chloroplast ultrastructure was observed using an H-7700 transmission electron microscope (Hitachi, Japan).

Genetic characterization

The mutant *M579* was used as the female parent (P_2), and the microspore DH line '579' of pakchoi was used as the male parent (P_1). The F_1 population was obtained by crossing *M579* and '579'. Then, the BC_1 population was generated by backcrossing the F_1 population with the two parents. Finally, the F_2 population was obtained through the self-crossing of the F_1 population. The separation of gray leaf plants and green leaf plants was determined. The proportion of segregation of plant leaf color traits was detected using a χ^2 test.

BSA sequencing

For BSA-seq, DNA was extracted from 60 randomly selected green-leaf plants and 60 mutant gray-leaf plants from the F_2 generation to establish a green-leaf library and a mutant gray-leaf library. DNA of two parents and two pools was extracted using the modified cetyltrimethylammonium bromide method[28]. Detection

of the DNA concentration was carried out using agarose gel electrophoresis.

After a DNA purity test was performed, the DNA was broken into 350 bp fragments using a fragmentation kit. The TruSeq Library Construction Kit platform (Illumina, San Diego, CA, USA) was used to establish the DNA library. Then the low-quality fragments were processed, including fragment purification, repair of fragment ends, PCR amplification, and connection of polyA tails and connectors. Finally, the construction of the library was completed.

The constructed library was first roughly quantified using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and the library concentration was diluted to 1 ng/μL, and the insert size of the library was determined using Agilent 2100 spectrophotometry (Agilent, Santa Clara, CA, USA). After the insert size was determined, precise quantification of the effective concentration of the constructed library was calculated using qRT-PCR. The library was then pooled according to the effective concentration and target downstream data volume requirements and sequenced using the Illumina HiSeq TMPE150 platform.

To further determine the location of the gray leaf mutation gene, the clean reads obtained after filtering were compared with the *Brapa*_sequence_v3.0 reference genome. Single nucleotide polymorphism (SNP) and insertion/deletion (InDel) detection and annotation were performed to obtain the location and type of variation of the variant loci.

Fine mapping of the candidate gene

Through linkage analysis using InDel molecular markers, the candidate interval identified by BSA-seq was further narrowed. The InDel primers were designed using Primer Premier 5.0 software (Premier Biosoft, Palo Alto, CSA, USA), based on the *B. rapa* genome V3.0 database information. To narrow down the candidate interval within the F_2 population, InDel primers were employed to perform polymorphism screening on both parental lines and two mixed pools, using PCR amplification and agarose gel electrophoresis for primer screening and population linkage analysis. Genetic map distances were calculated from recombination values using Kosambi's mapping function.

Gene cloning and sequencing

The full-length DNA and the promoter (2,000 bp upstream of the transcription start site) of the candidate gene *BraA06g036440.3C* were amplified using a previously published method[29], followed by Sanger sequencing at GENEWIZ (Suzhou, China).

To analyze the sequence of the inserted fragment, all reads covering the promoter were screened according to the BSA-seq results, and a total of 20 reads were obtained. The analysis revealed that there are two types of reads: promoter sequences and insertion sequences of other regions. The inserted fragment was amplified and sequenced.

Promoter activity analysis

The amplified promoter and pBWA(V)H-gus vector were digested with BsaI endonuclease. The target DNA fragment and vector fragment were obtained through gel recovery and ligated to obtain the vector carrying GUS (β-glucuronidase) reporter gene under the control of the isolated promoter. The vector was sequenced, verified, and introduced into *Agrobacterium tumefaciens* strain GV3101. Genetic transformation of wild-type *Arabidopsis* Col-0 was then carried out using the flower dipping method[30]. The collected seeds were cleaned and sown into MS solid Petri dishes containing 30 mg/L hygromycin. At 20–25 d after sowing, the viable positive plants were transplanted into the substrate and culture was continued. Subsequently, the DNA of leaves was extracted and tested using primers for the *GUS* gene. The primer sequences for the *GUS* gene are as follows: F: 5'-CGAACTGATTGACGTCATCG-3'; R: 5'-CAAGACCGGCAACAGG ATTCAATC-3'. Positive plants were selected for GUS tissue staining according to the instructions of the GUS activity kit (Coolaber Science & Technology Co., Ltd, Beijing, China).

Quantitative real-time reverse transcription PCR (qRT-PCR)

Total RNA was extracted from the leaves of '579' and *M579* and reverse transcribed into cDNA. The quantitative real-time PCR (qPCR) step of the qRT-PCR protocol was conducted using the cDNA as the template and Ultra SYBR Mixture dye (Kangwei Century, Beijing, China) on a QuantStudio 6 PCR system (Thermo Fisher Scientific, Shanghai, China). The *ACT* gene (*BraA10g027990.3C*) served as the internal control for gene expression[29]. Each experiment included three biological replicates and three technical replicates. Relative expression levels were calculated using the $2^{-\Delta\Delta C_T}$ method. The primer sequences were as follows: *ACT* (F: 5'-ATCTAC GAGGGTTATGCT-3'; R: 5'-CCACTGAGGACGATGTTT-3'); *Brgl1* (F: 5'-TCCCGTCTCCATCACTTCTACAT-3'; R: 5'-CTTCCACAAACAAACCGCA AA-3').

Bioinformatic analysis of the PPR gene family

The candidate gene *BraA06g036440.3C* is a member of the PPR gene family. The Pfam model PF01535 of PPR proteins was downloaded from the Pfam website (http://pfam.xfam.org/) to identify the PPR gene family of cabbage. Subcellular localization prediction of the identified PPR proteins was performed using the online tool Plant-mPLoc (www.csbio.sjtu.edu.cn). Analysis of the number of conserved structural domains of the identified PPR proteins was carried out using the Smart online tool (https://smart.embl-heidelberg.de/). Information on the location of the PPR gene on the chromosomes was extracted using TBtools. To analyze the evolutionary relationships among the PPR gene family members in cabbage, multiple sequence alignment of the screened PPR protein sequences was performed using MEGA5.0, and a phylogenetic tree was constructed using maximum likelihood estimation. Covariance analysis of the PPR gene family in cabbage was performed using TBtools.

Statistical data analysis

Data analysis was performed using SPSS software, and Duncan's Multiple Range Test (DMRT) was applied to compare the means ($p = 0.05$).

Results

Morphological characteristics and genetic analysis

The leaves of the wild-type '579' appeared bright green, while the leaves of the mutant *M579* appeared light gray (Fig. 1a). Otherwise, the leaves of '579' and *M579* had the same phenotypic characteristics. SEM observation of the leaf surfaces of '579' and *M579* revealed no differences in cell number, cell morphology, stomatal morphology, and stomatal number (Fig. 1b, c), indicating that the gray leaf mutation did not affect the epidermal cells.

The leaf length, leaf width, and plant height of *M579* were lower than those of '579', and the plant width of *M579* was higher than that of '579'. The overall differences in leaf length and plant height between '579' and *M579* were smaller during their growth. In ten measurements, the leaf length of *M579* was, at most, 12.03% lower and, at minimum, 2.40% lower than that of the control '579' (Fig. 2a1). Additionally, the plant height of *M579* was, at most, 16.7% lower and, at minimum, 1.8% lower than that of '579' (Fig. 2a2). The leaf width and plant width of '579' and *M579* differed significantly,

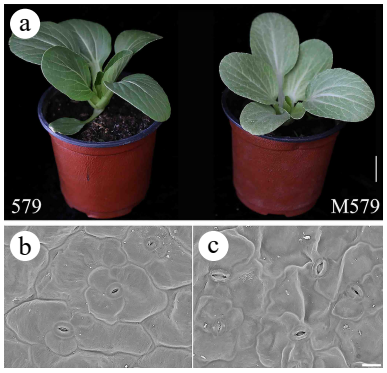

Fig. 1 Morphological characterization of wild-type '579' and mutant *M579*. (a) '579' with green leaves and *M579* with gray leaves. Bar = 2 cm. (b, c) Scanning electron microscopy observations of '579' and *M579*. Bar = 200 μm.

and among the 10 measurements, the leaf width of *M579* was, at most, 22% lower and, at minimum, 6.67% lower than that of the control '579' (Fig. 2a3). Furthermore, the plant width of *M579* was, at most, 33.84% lower and, at minimum, 10.71% lower than that of '579' (Fig. 2a4). Statistical analysis of the first four measurements showed that the plant width of '579' was higher than that of *M579* (4.12%, 8.12%, 7.45%, and 18.9%, respectively), and the maximum difference in plant width was in the eighth measurement, which was 33.84% lower than that of *M579*.

We found that the fresh and dry weights of *M579* were always lower than those of '579' during the whole growth cycle (five measurements). The fresh weight of the mutant was lower than that of the wild-type by 57.85% and 60.71% in the first two measurements, and the difference decreased gradually as the plants grew, with the latter three measurements being 23.20%, 28.01%, and 23.08% lower than those of '579', respectively (Fig. 2b1). The dry weights of *M579* were 39.4%, 7.84%, 20.52%, 31.80%, and 18.18% lower than those of '579', respectively (Fig. 2b2). This indicated that the gray leaf mutation affects the plant growth, development, and the accumulation of dry matter.

The F_1 plants of the mutant *M579* and '579' all had green leaves, and in the F_2 population, 1535 had green leaves and 465 had gray leaves, in line with the ratio of 3:1. The results indicated that the gray leaf phenotype of mutant *M579* was controlled by one pair of recessive nuclear genes (Table 1).

Photosynthetic pigment content, parametric analysis, and chloroplast structure

The content of all photosynthetic pigments in *M579* was lower than that in '579'. Compared with '579', the total chlorophyll content (CHl) of *M579* decreased by 54.28%, the content of chlorophyll a (Chla) decreased by 55.2%, the content of chlorophyll b (Chlb) decreased by 51.43%, and the content of carotenoid (Car) decreased by 41.18% (Fig. 2c).

To investigate the effect of the gray leaf mutation on plant photosynthesis, the photosynthetic parameters of '579' and *M579* were measured. The net photosynthetic rate (P_n), stomatal conductance (G_s), and transpiration rate (T_r) of *M579* were lower than those of '579', with decreases of 60.98%, 50%, and 33.75%, respectively, compared with those in '579'. The intercellular CO_2 concentration (C_i) of *M579* was 26.87% higher than that of '579', indicating that the reduced photosynthetic pigments affected its net photosynthetic rate. The relationship between the net photosynthetic rate and intercellular C_i correlates positively; therefore the result indicated that the CO_2 utilization rate of mutant *M579* decreased, resulting in an increased intercellular C_i (Table 2).

Chloroplasts are the main site of photosynthesis. The effect of gray leaf mutation on chloroplasts was observed by TEM (Fig. 2d). Inside the chloroplasts, '579' had a well-developed thylakoid membrane system with clear, well-arranged grana lamella and obvious starch grain; however, *M579* had a low number of grana and the small number of thylakoid stacks were irregular.

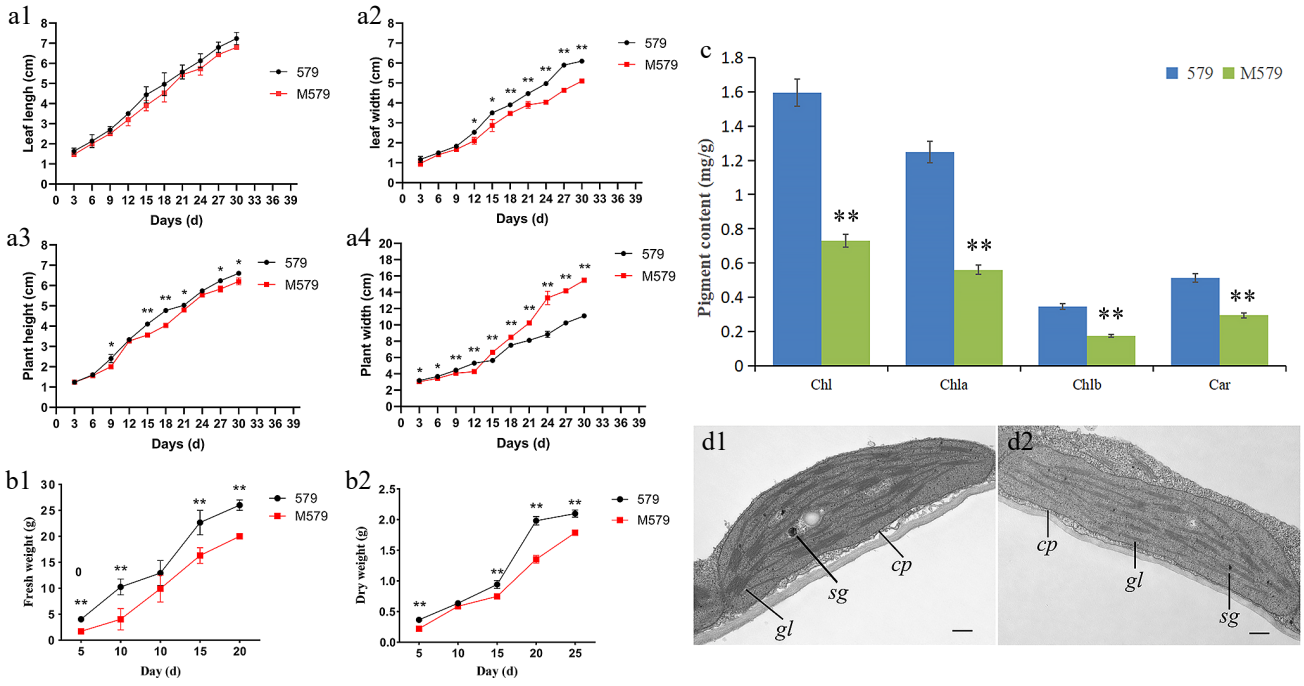

Fig. 2 Agronomic trait analysis of '579' and *M579*. (a) Growth potential curves of '579' and *M579*: (a1) leaf length, (a2) leaf width, (a3) plant height, and (a4) plant width. (b1) Fresh weight of '579' and *M579*. (b2) Dry weight of '579' and *M579*. (c) Photosynthetic pigment content of '579' and *M579*. (d) Transmission electron microscopy (TEM) analysis of chloroplasts in (d1) '579', and (d2) *M579*.

Table 1. Genetic analysis of the gray leaf mutant *Brgl1* in pakchoi.

Generation	Total	Green leaf	Gray leaf	Segregation ratio	χ^2	$\chi^2_{0.05}$
P_1 ('579')	23	23	0			
P_2 (*M579*)	23	0	23			
F_1 ($P_1 \times P_2$)	200	200	0			
F_1' ($P_2 \times P_1$)	136	136	0			
F_2	2,000	1,535	465	3.3:1	0.81	3.841

Table 2. '579' and *M579* photosynthetic parameters.

Material	Photosynthetic rate (mol/m^2/s)	Stomatal conductance (mol/m^2/s)	Transpiration rate (mol/m^2/s)	CO_2 concentration (mol/m^2/s)
579	6.38 ± 0.10**	0.06 ± 0.01**	1.60 ± 0.09*	221.33 ± 2.08**
M579	2.49 ± 0.06	0.03 ± 0.07	1.06 ± 0.17	302.67 ± 1.53

* Significantly different at p = 0.05 by the t-test. ** Significantly different at p = 0.01 by the t-test.

Preliminary mapping of *Brgl1*

After sequencing and subsequent filtering, 18,982,906,500 bp, 23,326,581,900 bp, 30,115,148,700 bp, and 31,095,782,700 bp of clean base pairs were obtained for '579', *M579*, the green leaf pool, and the gray leaf pool, respectively. These reads aligned to the reference genome at proportions of 97.44%, 97.21%, 96.88%, and 97.95%, respectively. A total of 3,200,779 SNP sites and 742,291 Indel sites were identified. A distribution map was generated to visually represent the SNP-index distribution across the chromosomes of the mutant progeny (Fig. 3a). Windows exceeding the threshold at the 95% confidence level were selected as candidate intervals, identifying the target region for the gray leaf trait within the 1–23 Mb region on chromosome A06 (Fig. 3b).

Fine mapping of *Brgl1*

To further narrow down the genetic location of *Brgl1*, 38 pairs of Indel marker primers were designed targeting the 23 Mb candidate interval on chromosome A06 (Supplementary Table S1). A total of 1,877 gray leaf F_2 plants were screened, and seven polymorphic markers were identified: SIN-7, SIN-13, SIN-22, SIN-27, SIN-33, SIN-34, and SIN-35 (Fig. 4a). The mapping results indicated that the markers SIN-7, SIN-13, SIN-22, SIN-27, and SIN-33 were located on one side of *Brgl1*, while SIN-34 and SIN-35 were on the other side. The *Brgl1* gene was ultimately localized to the region between markers SIN-33 and SIN-34 (Fig. 4b). A total of three recombinants were observed between the SIN-34 marker and *Brgl1*, corresponding to a genetic distance of 0.16 cM (Fig. 4b). Two recombinants were detected between the SIN-35 marker and *Brgl1*, corresponding to a genetic distance of 0.11 cM (Fig. 4b). The physical distance between the two markers was determined to be 8.069 kb (Supplementary Table S2).

Cloning of the candidate gene *Brgl1* and its promoter

Based on the reference genome information, a total of four genes were identified within the mapped interval. *BraA06g036440.3C*, which encodes a PPR superfamily protein, was selected for further analysis due to the involvement of some PPR family genes in leaf color development. The remaining three genes were annotated with functions unrelated to leaf color. The full-length sequences of *BraA06g036440.3C* in the wild type '579' and the mutant *M579* were compared, revealing no differences (Supplementary Fig. S1), indicating that the gray leaf trait in mutant *M579* is not due to mutations in the full-length gene.

The promoter regions of *BraA06g036440.3C* in '579' and *M579* were cloned and sequenced. The results revealed an insertion of a 4,788 bp fragment in the promoter of *M579*, relative to '579' (Fig. 5a; Supplementary Fig. S2). The inserted sequence was located at position −246 bp of the gene's translation start site, between positions 24,495,043 and 24,495,044 on chromosome A06. BLAST analysis showed a high degree of homology between the insertion fragment and the retrotransposon TNT 1-94. During transposition, this type of transposon first transcribes DNA to form mRNA, which is then reverse transcribed into cDNA using the mRNA as a template, and is finally inserted into genomic DNA to complete transposition. This type of transposon can amplify portions of the genome and cause changes in gene structure. The insertion of TNT 1-94 in the mutant *M579* caused an abnormal promoter structure of *BraA06g036440.3C*, which was likely to be responsible for the gray leaf mutant phenotype. Therefore, *BraA06g036440.3C* was predicted to be a strong candidate gene for *Brgl1*.

In plants, long terminal repeat retrotransposons are mainly divided into two superfamily groups, Copia and Gypsy. TNT 1-94, a member of the Copia superfamily (Fig. 5b), is a long terminal repeat (LTR) retrotransposon consisting of target site repeats (TSRs), identical 5' and 3' LTRs, a TGCA box, a primer binding site (PBS), a polypurine tract (PPT), and a protein-coding region. The latter encodes gag and pol proteins, including key enzymes essential for autonomous transposition.

Expression pattern analysis of *Brgl1*

To investigate the expression pattern of *Brgl1*, a vector containing the GUS gene driven by the *Brgl1* promoter was introduced into *Arabidopsis*. The results indicated that the promoter was active in the leaves of *Arabidopsis* (Fig. 6a–c).

The relative expression levels of *Brgl1* in '579' and *M579* were further analyzed using qRT-PCR. The results indicated *Brgl1* expression was significantly reduced in the mutant *M579*, suggesting that the promoter mutation of *Brgl1* affects its expression (Fig. 6d).

Bioinformatic analysis of the PPR gene family

BraA06g036440.3C belongs to the P-subfamily of the PPR gene family. Studies on PPR proteins have shown that they are mainly located in organelles, such as chloroplasts and mitochondria, and

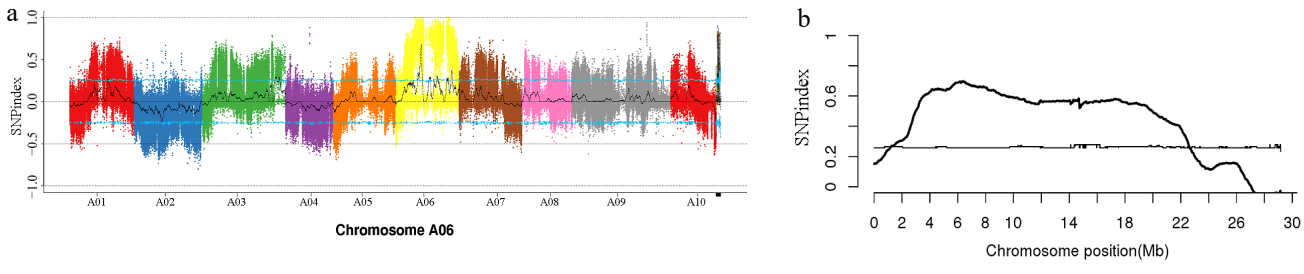

Fig. 3 Single nucleotide polymorphism (SNP) index mapping chart. (a) SNP index distribution on the chromosomes of pakchoi. (b) The candidate gene for the gray leaf trait was mapped to chromosome A06.

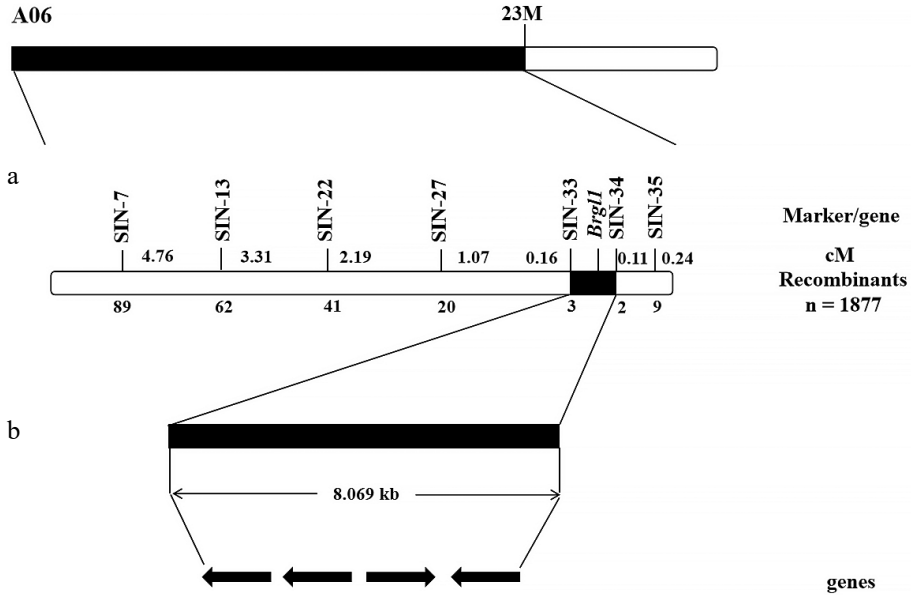

Fig. 4 The genetic and physical maps of *Brgl1*. (a) *Brgl1* was mapped between the molecular markers SIN-33 and SIN-34. (b) The 8.069 kb region between SIN-33 and SIN-34 containing four genes.

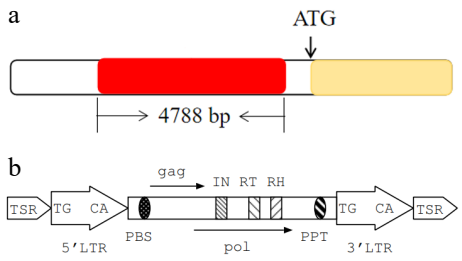

Fig. 5 Gene structural features. (a) Structural features of the *Brgl1* gene and the insertion fragment. The yellow box represents the *Brgl1* gene, the red box represents the insertion fragment. (b) Structural features of transposon TNT 1-94.

participate in the expression and regulation of organellar genes. The Pfam model PF01535 for PPR proteins was downloaded from the Pfam website to identify the PPR gene family. A total of 485 PPR genes were identified, including 226 members of the P subfamily and 259 members of the PLS subfamily. The subcellular localization prediction results showed that 174 PPR proteins were localized to multiple locations, including chloroplasts, mitochondria, and the nucleus; 210 PPR proteins were localized only to chloroplasts; 53 PPR proteins were localized only to the nucleus; and 30 PPR proteins were localized only to mitochondria.

The structural characteristics and motif distribution of genes within the PLS (Supplementary Fig. S3) and P (Supplementary Fig. S4) subfamilies were analyzed. Genes in both subfamilies exhibited variations in exon number, with some genes containing over 20 exons, while others possessed fewer, indicating potential functional divergence and evolutionary adaptations. A total of 20 distinct motifs (Motif 1 to Motif 20) were identified, showing significant differences in motif composition and abundance among the genes. Notably, Motifs 1, 2, and 3 were found to be highly conserved across the majority of PPR genes in both the PLS and P subfamilies, suggesting their crucial roles in the PPR gene family.

To investigate the regulatory mechanisms of the *PPR* gene family, 2,000 bp upstream of the ATG start codon of each gene was selected as the promoter sequence for cis-regulatory element analysis. The results revealed that the promoter regions of both the PLS (Supplementary Fig. S5) and P (Supplementary Fig. S6) subfamilies were enriched with a diverse array of cis-regulatory elements, particularly exhibiting high complexity and diversity in hormone response and stress response elements. The presence of these specific elements suggests the potential crucial roles of *PPR* genes in growth and development as well as in response to abiotic stresses.

Using TBtools, we extracted the positional information of PPR genes in 10 chromosomes. The results indicated that 482 sequences were located on the chromosomes, with an average of 48.2 PPR genes per chromosome. However, the overall distribution of genes showed a large difference in the number of PPR gene families on each chromosome. The number of genes on chromosomes A06 and A09 was significantly higher than the average, with A09 having the highest number of PPR genes (n = 79), while chromosomes A04 and A08 contained only 22 and 26 PPR genes, respectively, which is significantly lower than the average. The distribution of different subfamilies on each chromosome also varies significantly. On chromosome A03, the P subfamily accounts for 64.71% of the total gene number, while on chromosome A05, the PLS subfamily accounts for 70.91% of the total number (Table 3).

To analyze the evolutionary relationship among members of the PPR gene family, multiple sequence alignment was performed on selected PPR proteins. The results showed that the PPR genes could be divided into the P and PLS subfamilies (Fig. 7). Co-linearity analysis of the PPR gene family using TBtools revealed a strong overall co-linear relationship among PPR members, indicating the preservation of a high degree of similarity among family members during the evolutionary process.

Discussion

Previous studies have isolated and identified a variety of leaf color mutants, which can be classified into different categories, including albino mutants, yellow mutants, striped mutants, light greening mutants, and zebra mutants[31]. This study identified a gray leaf mutant and cloned the mutated gene *Brgl1* (*BraA06g036440.3C*). *Brgl1* belongs to the P subfamily of the PPR gene family and encodes a hydrophilic non-secretory protein. This study successfully cloned the gray leaf mutation gene *Brgl1* for the first time, contributing to

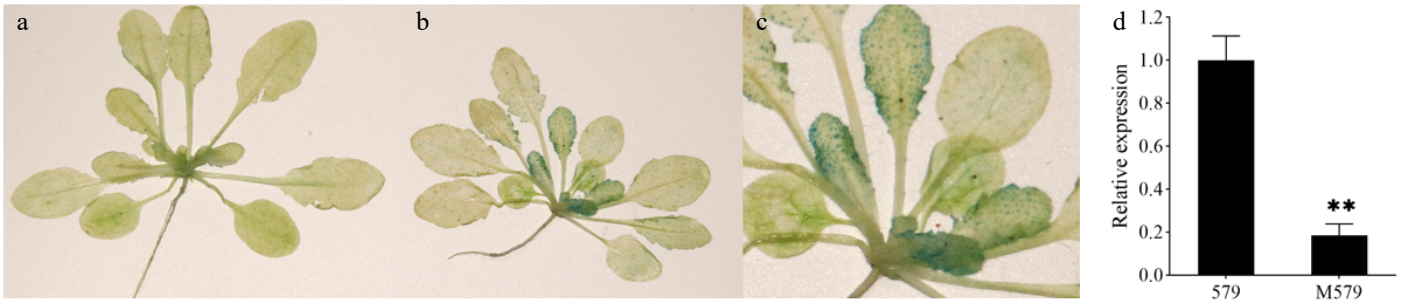

Fig. 6 Expression patterns of the *Brgl1* gene. (a) GUS staining of the control in *Arabidopsis*. (b) GUS-stained positive plants in *Arabidopsis*. (c) Close-up of GUS-stained positive plants in *Arabidopsis*. (d) qRT-PCR analysis of *Brgl1* in the leaves of '579' and *M579*.

Table 3. The distribution of PPR family genes across the ten chromosomes of pakchoi.

Chromosome number	Number of P subgroups	Number of PLS subgroups	Total number of PPR genes
A01	25	23	58
A02	20	10	30
A03	18	33	51
A04	10	12	22
A05	16	39	55
A06	38	35	73
A07	26	20	46
A08	10	16	26
A09	45	44	79
A10	16	26	42
Total	224	258	482

determining the molecular mechanisms underlying leaf color formation in pakchoi.

PPR proteins are involved in the regulation of chloroplast development and the processing and maturation of chloroplast mRNA. Mutations in certain PPR genes result in alterations in leaf color phenotypes. In the rice white-striped leaf mutant *Wsl4*, the striped leaf phenotype and reduced chlorophyll content were caused by defects in chloroplast development, with the mutated gene *wsl4* identified to encode a P-type PPR protein[32]. A *PPR* gene, *GhYGL1d*, which is essential for plastid development, was identified in cotton, and its mutation resulted in a yellow-green leaf phenotype[33]. In rice, a PLS-type PPR protein, PGL12, was identified, and mutations in the *pgl12* gene resulted in a pale green leaf phenotype in the mutant[34]. A pale green leaf mutant, *Gmpgl2*, was also identified in soybean[23]. In maize, the mutated gene *PPR8522*, which encodes a P-type PPR protein, was identified, and its mutation resulted in the albino seedling phenotype in the *emb8522* mutant[35]. Research on the *PPR* gene family has revealed that PPR proteins are predominantly localized in organelles, including mitochondria and chloroplasts[36]. The specific mechanism by which PPR proteins in mitochondria affect leaf color formation remains unclear. An EMS-mutagenized rice screen identified a suppressor, *ospus1*, which encodes a PPR protein localized in the mitochondria; its mutation can restore the albino leaf phenotype[37]. A P-type PPR protein named *BoYgl-2* has been identified in cabbage. The *BoYgl-2* knockout mutant exhibited yellow-green leaves, reduced chlorophyll content, and abnormal chloroplast development. Studies have shown that *BoYgl-2* participated in C-to-U editing of chloroplast RNA, regulating the expression of multiple genes and affecting chloroplast development and chlorophyll biosynthesis[25]. The PPR protein Maker00002998 may be involved in the RNA editing of the chloroplast gene *rps14*, providing new insights into the role of PPR proteins in the mechanism of leaf albinism in *Kandelia obovata* propagules[38]. This study identified the mutated gene *Brgl1* in the gray leaf mutant *M579* of pakchoi. GUS activity analysis indicated that *Brgl1* was expressed in the leaves.

Variations in promoter sequences could potentially affect the expression of plant genes[39]. A yellow-green young leaf mutant (*ygl*) was identified in *Brassica napus*, which was characterized by chlorophyll deficiency, caused by the insertion of a non-coding RNA into

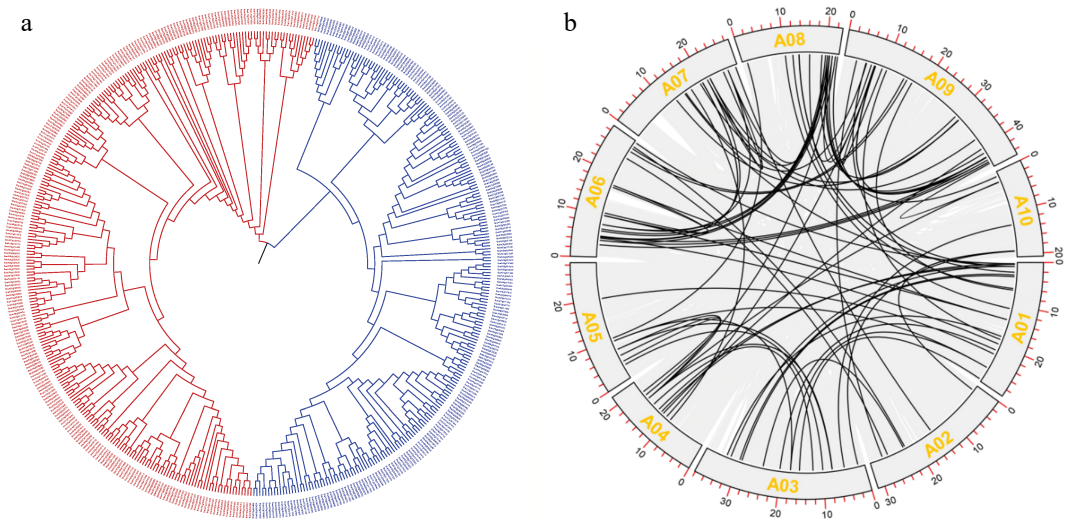

Fig. 7 Evolutionary analysis of PPR proteins. (a) Phylogenetic tree of PPR proteins. (b) Co-linearity analysis of PPR proteins.

the promoter of *BnaC07.HO1*[40]. A green leaf mutant, *9110Gt*, was identified in cucumber, characterized by young yellow leaves. A single nucleotide change in the promoter region of the mutant gene *v-1* led to significantly lower *v-1* expression in the true leaves of the mutant compared with that in the wild-type[41]. A 30-bp deletion in the promoter of *Bra024218* in Chinese cabbage led to leaf yellowing in the mutant *pem*, accompanied by delayed chloroplast development, reduced chlorophyll content, and decreased photosynthetic capacity[42]. In this study, a large insertion of 4,788 bp in the promoter region of the gene *Brgl1* resulted in the gray leaf phenotype observed in the mutant *M579*.

Transposable elements (TEs) are non-coding DNA sequences that constitute a significant portion of eukaryotic genomes. They can self-replicate and are distributed throughout the genome. TEs have also been known as 'jumping genes' because of their capacity to relocate within the genome. In higher plant genomes, transposon sequences constitute a significant portion, with most being inactive. Activation under certain conditions could lead to rapid transposition and genome expansion[43]. In maize, the insertion of a 51 bp Popin transposon within *ZmCAO1* severely disrupted its transcription, leading to the yellow-green leaf phenotype of the *ygl* mutant[44]. The insertion of a Tourist-like miniature inverted repeat transposable element within the *TCP4* promoter in *japonica* rice (*TCP4*$^{M+}$) transcriptionally suppressed *TCP4* expression[45]. The insertion of a transposable element within the promoter of the E2 ubiquitin conjugase gene *OsUBC12* in rice significantly enhanced its transcription, thereby improving the low-temperature germination ability of rice[46]. In this study, the insertion of long terminal repeat retrotransposon TNT 1-94 within the promoter of the candidate gene *Brgl1* in the gray leaf mutant *M579* disrupted the promoter structure. This structural disruption was the cause of the significant downregulation of *Brgl1* gene expression.

In conclusion, this study identified a natural mutant, *M579*, with gray leaves, which showed decreased chlorophyll content and net photosynthetic rate, as well as chloroplast developmental defects. *Brgl1*, encoding a member of the P subfamily of the PPR protein family, was identified as a candidate gene based on the insertion of transposon TNT 1-94 into its predicted promoter. This study provided a foundation to further determine the molecular mechanisms underlying gray leaf formation in pakchoi.

Conclusions

This study identified a grey leaf mutant of pakchoi, designated as *M579*. Using BSA-seq and fine mapping, the candidate gene was identified as *Brgl1* (*BraA06g036440.3C*). The insertion of a 4,788 bp fragment in the *Brgl1* promoter was identified as the cause of the gray leaf phenotype in *M579*. Our findings lay the foundation for understanding the formation of grey leaves in pakchoi.

Author contributions

The authors confirm contributions to the paper as follows: data analysis: Liu C, Guo K; drafted the manuscript: Liu C; materials creation, and performing the experiments: Liu C, Guo K, Shang B, Tan C, Zhang Y; directed the whole study including designing experiments and revising the manuscript: Ren J, Liu Z. All authors reviewed the results and approved the final version of the manuscript.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Acknowledgments

This work was supported by the National Key Research and Development Program (2022YFF1003004 and 2023YFD1200101), the National Natural Science Foundation of China (32272736 and 32472762), and the Earmarked Fund for CARS-23.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/vegres-0025-0020)

Dates

Received 6 January 2025; Revised 15 March 2025; Accepted 22 April 2025; Published online 5 August 2025

References

1. Carol P, Stevenson D, Bisanz C, Breitenbach J, Sandmann G, et al. 1999. Mutations in the *Arabidopsis* gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. *The Plant Cell* 11:57−68
2. Kumar AM, Söll D. 2000. Antisense HEMA1 RNA expression inhibits heme and chlorophyll biosynthesis in *Arabidopsis*. *Plant Physiology* 122:49−56
3. Lonosky PM, Zhang X, Honavar VG, Dobbs DL, Fu A, et al. 2004. A proteomic analysis of maize chloroplast biogenesis. *Plant Physiology* 134:560−74
4. Rudoi AB, Shcherbakov RA. 1998. Analysis of the chlorophyll biosynthetic system in a chlorophyll b-less barley mutant. *Photosynthesis Research* 58:71−80
5. Zhao Y, Huang S, Zhang M, Zhang Y, Feng H. 2021. Mapping of a pale green mutant gene and its functional verification by allelic mutations in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Frontiers in Plant Science* 12:699308
6. Ma YY, Shi JC, Wang DJ, Liang X, Wei F, et al. 2023. A point mutation in the gene encoding magnesium chelatase I subunit influences strawberry leaf color and metabolism. *Plant Physiology* 192:2737−55
7. Oster U, Tanaka R, Tanaka A, Rüdiger W. 2000. Cloning and functional expression of the gene encoding the key enzyme for chlorophyll b biosynthesis (CAO) from *Arabidopsis thaliana*. *The Plant Journal* 21:305−10
8. Lee S, Kim JH, Yoo ES, Lee CH, Hirochika H, et al. 2005. Differential regulation of chlorophyll a oxygenase genes in rice. *Plant Molecular Biology* 57:805−18
9. Zhang T, Dong X, Yuan X, Hong Y, Zhang L, et al. 2022. Identification and characterization of *CsSRP43*, a major gene controlling leaf yellowing in cucumber. *Horticulture Research* 9:uhac212
10. Campbell BW, Mani D, Curtin SJ, Slattery RA, Michno JM, et al. 2015. Identical substitutions in magnesium chelatase paralogs result in chlorophyll-deficient soybean mutants. *G3 Genes|Genomes|Genetics* 5:123−31
11. Williams P, Hardeman K, Fowler J, Rivin C. 2006. Divergence of duplicated genes in maize: evolution of contrasting targeting information for enzymes in the porphyrin pathway. *The Plant Journal* 45:727−39
12. Tanaka R, Tanaka A. 2007. Tetrapyrrole biosynthesis in higher plants. *Annual Review of Plant Biology* 58:321−46
13. Fölsche V, Großmann C, Richter AS. 2022. Impact of porphyrin binding to GENOMES UNCOUPLED 4 on tetrapyrrole biosynthesis *in planta*. *Frontiers in Plant Science* 13:850504
14. Song M, Wei Q, Wang J, Fu W, Qin X, et al. 2018. Fine mapping of CsVYL , conferring virescent leaf through the regulation of chloroplast development in cucumber. *Frontiers in Plant Science* 9:432
15. Zhu X, Guo S, Wang Z, Du Q, Xing Y, et al. 2016. Map-based cloning and functional analysis of YGL8, which controls leaf colour in rice (*Oryza sativa*). *BMC Plant Biology* 16:134

16. Mullet JE. 1993. Dynamic regulation of chloroplast transcription. *Plant Physiology* 103:309–13
17. Moreira D, Le Guyader H, Philippe H. 2000. The origin of red algae and the evolution of chloroplasts. *Nature* 405:69–72
18. Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. 1999. Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Research* 6:283–90
19. Liu X, Lan J, Huang Y, Cao P, Zhou C, et al. 2018. WSL5, a pentatricopeptide repeat protein, is essential for chloroplast biogenesis in rice under cold stress. *Journal of Experimental Botany* 69:3949–61
20. Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, et al. 2004. Genome-wide analysis of *Arabidopsis pentatricopeptide* repeat proteins reveals their essential role in organelle biogenesis. *The Plant Cell* 16:2089–103
21. Wang X, An Y, Qi Z, Xiao J. 2021. PPR protein Early Chloroplast Development 2 is essential for chloroplast development at the early stage of *Arabidopsis* development. *Plant Science* 308:110908
22. Lan J, Lin Q, Zhou C, Liu X, Miao R, et al. 2023. Young Leaf White Stripe encodes a P-type PPR protein required for chloroplast development. *Journal of Integrative Plant Biology* 65:1687–702
23. Feng X, Yang S, Zhang Y, Cheng Z, Tang K, et al. 2021. GmPGL2 , encoding a pentatricopeptide repeat protein, is essential for chloroplast RNA editing and biogenesis in soybean. *Frontiers in Plant Science* 12:690973
24. Huang J, Lu G, Liu L, Raihan MS, Xu J, et al. 2020. The kernel size-related quantitative trait locus qKW9 encodes a pentatricopeptide repeat protein that aaffects photosynthesis and grain filling. *Plant Physiology* 183:1696–709
25. Zhang B, Wu Y, Li S, Ren W, Yang L, et al. 2024. Chloroplast C-to-U editing, regulated by a PPR protein BoYgl-2, is important for chlorophyll biosynthesis in cabbage. *Horticulture Research* 11:uhae006
26. Holm G. 1954. Chlorophyll mutations in barley. *Acta Agriculturae Scandinavica* 4:457–71
27. Zhao H, Yu L, Huai Z, Wang X, Ding G, et al. 2014. Mapping and candidate gene identification defining BnChd1-1, a locus involved in chlorophyll biosynthesis in *Brassica napus*. *Acta Physiologiae Plantarum* 36:859–70
28. Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8:4321–26
29. Liu C, Song G, Wang N, Huang S, Gao Y, et al. 2021. A single SNP in Brcer1 results in wax deficiency in Chinese cabbage (*Brassica campestris* L. *pekinensis*). *ssp. Scientia Horticulturae* 282:110019
30. Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* 16:735–43
31. Jung KH, Hur J, Ryu CH, Choi Y, Chung YY, et al. 2003. Characterization of a rice chlorophyll-deficient mutant using the T-DNA gene-trap system. *Plant and Cell Physiology* 44:463–72
32. Wang Y, Ren Y, Zhou K, Liu L, Wang J, et al. 2017. *WHITE STRIPE LEAF4* encodes a novel P-type PPR protein required for chloroplast biogenesis during early leaf development. *Frontiers in Plant Science* 8:1116
33. He P, Wu S, Jiang Y, Zhang L, Tang M, et al. 2019. GhYGL1d, a pentatricopeptide repeat protein, is required for chloroplast development in cotton. *BMC Plant Biology* 19:350
34. Chen L, Huang L, Dai L, Gao Y, Zou W, et al. 2019. PALE-GREEN LEAF12 encodes a novel pentatricopeptide repeat protein required for chloroplast development and 16S rRNA processing in rice. *Plant and Cell Physiology* 60:587–98
35. Sosso D, Canut M, Gendrot G, Dedieu A, Chambrier P, et al. 2012. PPR8522 encodes a chloroplast-targeted pentatricopeptide repeat protein necessary for maize embryogenesis and vegetative development. *Journal of Experimental Botany* 63:5843–57
36. Wang X, An Y, Xu P, Xiao J. 2021. Functioning of PPR proteins in organelle RNA metabolism and chloroplast biogenesis. *Frontiers in Plant Science* 12:627501
37. Zu X, Luo L, Wang Z, Gong J, Yang C, et al. 2023. A mitochondrial pentatricopeptide repeat protein enhances cold tolerance by modulating mitochondrial superoxide in rice. *Nature Communications* 14:6789
38. Xu C, Wang JC, Sun L, Zhuang LH, Guo ZJ, et al. 2024. Genome-wide identification of pentatricopeptide repeat (PPR) gene family and multi-omics analysis provide new insights into the albinism mechanism of *Kandelia obovata* propagule leaves. *Plant, Cell & Environment* 47:5498–510
39. Wang Y, Mai W, Liang C, Zhang M. 2003. Advances on studies of plant promoters. *Acta Botanica Boreali-occidentalia Sinica* 23:2040–48
40. Zhu L, Yang Z, Zeng X, Gao J, Liu J, et al. 2017. Heme oxygenase 1 defects lead to reduced chlorophyll in *Brassica napus*. *Plant Molecular Biology* 93:579–92
41. Miao H, Zhang S, Wang M, Wang Y, Weng Y, et al. 2016. Fine mapping of virescent leaf gene v-1 in cucumber (*Cucumis sativus* L.). *International Journal of Molecular Sciences* 17:1602
42. Li X, Huang S, Liu Z, Hou L, Feng H. 2019. Mutation in EMB1923 gene promoter is associated with chlorophyll deficiency in Chinese cabbage (*Brassica campestris* ssp. *pekinensis*). *Physiologia Plantarum* 166:909–20
43. Naito K, Zhang F, Tsukiyama T, Saito H, Hancock CN, et al. 2009. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* 461:1130–34
44. Li Q, Zhou S, Liu W, Zhai Z, Pan Y, et al. 2021. A chlorophyll a oxygenase 1 gene ZmCAO1 contributes to grain yield and waterlogging tolerance in maize. *Journal of Experimental Botany* 72:3155–67
45. Zhang H, Zhang J, Xu P, Li M, Li Y. 2024. Insertion of a miniature inverted-repeat transposable element into the promoter of OsTCP4 results in more tillers and a lower grain size in rice. *Journal of Experimental Botany* 75:1421–36
46. Zhang C, Wang H, Tian X, Lin X, Han Y, et al. 2024. A transposon insertion in the promoter of OsUBC12 enhances cold tolerance during Japonica rice germination. *Nature Communications* 15:2211