Characterization and Fine Mapping of a Yellow-Virescent Gene Regulating Chlorophyll Biosynthesis and Early Stage Chloroplast Development in Brassica napus

Chuanji Zhao,* Lijiang Liu,* Luqman Bin Safdar,* Meili Xie,* Xiaohui Cheng,* Yueying Liu,* Yang Xiang,* Chaobo Tong,* Jinxing Tu,† Junyan Huang,*1 and Shengyi Liu*

*Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences/The Key Laboratory of Biology and Genetic Improvement of Oil Crops, The Ministry of Agriculture and Rural Affairs, Wuhan 430062, China, †National Key Laboratory of Crop Genetic Improvement, National Center of Rapeseed Improvement, Huazhong Agricultural University, Wuhan 430070, China, ‡Guizhou Rapeseed Institute, Guizhou Academy of Agricultural Sciences, Guiyang 550008, P. R. China

ORCID IDs: 0000-0002-2100-2510 (C.Z.); 0000-0002-9544-8259 (L.B.S.)

ABSTRACT Chlorophyll biosynthesis and chloroplast development are crucial to photosynthesis and plant growth, but their regulatory mechanism remains elusive in many crop species. We isolated a Brassica napus yellow-virescent leaf (yvl) mutant, which exhibited yellow-younger-leaf and virescent-older-leaf with decreased chlorophyll accumulation and delayed chloroplast development. We mapped yvl locus to a 70-kb interval between molecular markers yvl-O10 and InDel-O6 on chromosome A03 in BC2F2 population using whole genome re-sequencing and bulked segregant analysis. The mutant had a ‘C’ to ‘T’ substitution in the coding sequence of BnaA03.CHLH, which encodes putative H subunit of Mg-protoporphyrin IX chelatase (CHLH). The mutation resulted in an imperfect protein structure and reduced activity of CHLH. It also hampered the plastid encoded RNA polymerase which transcribes regulatory genes of photosystem II and I. Consequently, the chlorophyll a/b and carotenoid contents were reduced and the chloroplast ultrastructure was degraded in yvl mutant. These results explain that a single nucleotide mutation in BnaA03.CHLH impairs PEP activity to disrupt chloroplast development and chlorophyll biosynthesis in B. napus.

KEYWORDS Brassica napus leaf color chlorophyll biosynthesis chloroplast development BSA-seq gene cloning RNA-seq

Chlorophyll biosynthesis is an important component of tetrapyrroles’ synthetic metabolism (Czarnecki and Grimm 2012), and disruption in this pathway can induce changes in pigments and leaf color, such as chlorotic leaves (Wu et al. 2007). Chlorotic leaves disrupt the photosynthetic capacity of plants resulting in damages to plant metabolism and ultimately yield products, and hence, several studies have evaluated the genetic mechanism underlying chlorotic leaves in crop plants (Yang et al. 2016; Ge et al. 2017). In a previous study, Gibson and colleagues reported that the conversion of protoporphyrin IX into Mg-protoporphyrin IX was possible when three genes BchH, BchD and BchI were expressed together, and mutations in these plant genes caused disruptions in chlorophyll biosynthesis (Gibson et al. 1995). Afterward, the various subunits of chlorophyll biosynthesis genes were studied in several crop plants, such as OsCHLH, OsCHLD in rice and GhCHLI in cotton (Jung et al. 2003; Ruan et al. 2017; Zhu et al. 2017a). Apart from its role in chlorophyll biosynthesis pathway, Mg-protoporphyrin IX chelatase is also involved in retrograde signaling. Retrograde signaling is a complex network that regulates chloroplast development during the early leaf stage and transfers information about the status of chloroplast to the nucleus to regulate gene expression (Pogson and Albrecht 2011; Hernández-Verdeja and Strand 2018). Earlier, Mg-protoporphyrin IX chelatase was suggested to be a plastid signal in retrograde signaling.
pathway that regulated the expression of photosynthesis-related nuclear genes (Mochizuki et al. 2001). A later study reported that Mg-protoporphyrin IX affected chlorophyll synthesis and diffusion, or shuttled chloroplast envelope to carry plastid signals into cell solutes (Barajas-Lopez Jde et al. 2013). CHLH gene coding the H subunit of Mg-chelatase is a key catalytic subunit in the production of Mg-protoporphyrin IX, and, therefore, an important component of chlorophyll biosynthesis and retrograde signaling (Moulin et al. 2008; Mochizuki et al. 2008).

Nuclear and chloroplast genes work in a coordinated manner to regulate chloroplast development, which is involved in the chloroplast and nuclear gene transcription, protein translation, processing modification, protein folding and degradation, thylakoid formation, and pigment synthesis (Beck 2001; Yu et al. 2013; Li et al. 2015). A cytoskeletal GTPase (FtsZ) is a component of plastid division machinery and is abundantly expressed in the early stage of chloroplast division (Vitha et al. 2001; TerBush et al. 2013). The transcription of plastid genes is primarily responsible for two RNA polymerases, the nuclear encoded RNA polymerase (NEP) and the plastid encoded RNA polymerase (PEP) (Shina et al. 2005; Zubo et al. 2011). The PEP complex is composed of four NEP-encoded core subunits (i.e., ropA, ropB, ropC1 and ropC2) and plays an important role in chloroplast maturity by producing many transcripts of photosynthesis genes (Borner et al. 2015).

Brassica napus (rapeseed) is a widely cultivated oil crop, mainly for its oil-rich seeds as it is the third largest vegetable oil source for humans. Although the chlorophyll biosynthesis genes have been well studied in other crops, such as A. thaliana and O. sativa, limited information is available for the regulatory mechanism of photosynthetic genes and yellowish leaf phenotype in rapeseed. A few leaf color mutants are reported (Wang et al. 2016b; Zhao et al. 2013) and only one gene, BnaC07.HO1, which encodes a heme oxygenase, has been cloned in B. napus (Zhu et al. 2017b). Apart from the involvement of YVL genes in photosynthetic machinery, yvl phenotype also plays a vital role as an indicator in cross breeding. During the crop production of hybrid rapeseed, the purity of male sterile line had decreased due to the influence of environmental factors or the mixing of maintainer lines in sterile lines, which further affected the purity of hybrid seeds. If the recessive leaf color marker traits, which are not only different from the normal green leaves but also well-developed in seedling stage and have little effect on yield, are introduced into sterile lines, false hybrids can be detected through leaf color at seedling stage (Zhao et al. 2010). However, most leaf color mutants lead to a functional disruption in chloroplasts and subsequently abnormal growth, sometimes even lethal effects on plants, which pose serious challenges to crop yield (Yang et al. 2016; Ge et al. 2017; Hsieh et al. 2017).

In this study, we isolated a mutant yvl with ethyl methanesulfonate (EMS) mutagenesis and successfully cloned the causal gene BnaA03.yvl with BSA-seq method. BnaA03.yvl caused a yellow-virescent leaf phenotype with chlorophyll deficient and disruptions in chloroplast structure. Also, mutations in BnaA03.yvl caused variation in transcript abundance levels of photosynthesis pathway genes. The results of present study explain the regulatory framework of chlorophyll biosynthesis pathway with BnaA03.yvl being a major contributor to the yellow-virescent leaf phenotype. Furthermore, the yvl mutant can also be used as a potential breeding marker in B. napus.

**MATERIALS AND METHODS**

**Plant materials and growth conditions**

A yellow-virescent leaf (yvl) mutant of *Brassica napus* was isolated from ZS9 by treating with 0.5% EMS for 16 h. Two F2 populations, which had 927 and 346 individuals respectively, were generated from reciprocal crossing of yvl with ZS9. The F1 hybrid, yvl·ZS9, was three times backcrossed to yvl. One BC1F3 population of 159 lines and the two F2 populations were used for genetic analysis and BSA-seq. About 2200 progeny from self-crossing of BC2F1 were used for fine mapping. Twenty BC1F3 plants and a diversity panel of 629 oilseed accessions (Table S1), including ZS9 and yvl, were used to validate the substitution site in *BnaA03.YVL*. The 629 rapeseed accessions were worldwide collected and most of them came from China. All the plants were grown in fields located in the Hubei Province in central China.

**Chlorophyll determination**

For pigment extraction, eight-leaf staged yvl mutant and ZS9 were sampled. The second and fifth leaves from the top were used as the younger and older leaves comprising about 30 mg of fresh weight. The leaves were then soaked in an 80% (v/v) acetone and ethanol mixture solution at 25° in dark for 48 h. Chlorophyll a (Chl a), chlorophyll b (Chl b) and carotenoid (Car) were measured with UV-spectrophotometer (ANALYTIKJENA), according to the method outlined by Arnon (Arnon 1949).

**Transmission electron microscopy analysis**

The collected younger and older leaves were fixed in a phosphate buffer containing 2.5% (w/v) glutaraldehyde overnight at 4°, and then further fixed in 1% OsO4 at 4° for 3 h. Afterward, samples were successively distilled three times with saline phosphate buffer for 60 min, dehydrated with a series of gradient alcohol, treated with acetone and embedded in epoxy resins and polymerized at 60°. The samples were stained with uranyl acetate and examined using Tecnai G2 20 TWIN transmission electron microscope (FEI, U.S.A.).

**Whole genome re-sequencing and BSA-seq analysis**

High-quality genomic DNA (gDNA) was isolated from fresh leaves by using a Hi-DNA secure Plant Kit (TIANGEN) and quantified to equal concentrations. For the rapid mapping and delimitation of the yvl locus, two parental lines and two extreme pools (i.e., G-pool and Y-pool, each consisting of 30 individuals) were selected for the ZS9-type and yvl-type leaf colors from the F2 population of yvl·ZS9. Samples were prepared and subjected to WGS using Illumina HiSeq2500 platform and 125 bp paired-end reads were generated with an insert size of around 350 bp. Burrows-Wheeler Aligner (BWA) was used to align the clean reads of the two parents and two pools against the reference genome of *B. napus* (Li and Durbin 2009; Chalhoub et al. 2014). Alignment files were converted to Binary Alignment/Map (BAM) files using the Sequence Alignment/Map (SAM) tools software (Li et al. 2009). For all samples, SNP calling was performed by using the Unified Genotype function of the Genome Analysis Toolkit (GATK) software (McKenna et al. 2010).

To identify candidate genomic region responsible for yvl, we compared the ‘SNP-index’ between the G- and Y-pools. The SNP-index is estimated as a proportion of reads aligned to a position with a variant nucleotide different from the reference sequence. These positions were filtered out; SNP indexes in both pools were less than 0.3 and SNP sequencing depths were less than 7. The ΔSNP-index was calculated by subtracting the SNP-index of the G-pool from the Y-pool. A sliding window approach with 1 Mb window size and 10 kb step size as the default settings was used to measure the average distribution of all SNP-index of SNPs mapped
across 19 chromosomes of B. napus in the given genomic interval. The distribution pattern of SNPs and InDels between ZS9 and yvl, as well as the SNP-index graphs for the G- and Y-pools, and corresponding ΔSNP-index graphs were plotted and showed by circular pictures.

**Fine mapping of the yvl locus**

According to the WGS data, homozygous SNPs and InDels as polymorphic molecular markers were first identified by the alignment of reads from the parental lines to the reference genome using SAM tools. In order to confirm whether SNPs were homozygous and possibly polymorphic between the two parents, the primers (Table S2) were designed to amplify approximately 400-2000 bp nucleic acid sequences containing variations of SNPs and sequence the PCR products by the Sanger method to confirm positive polymorphisms (Wang et al. 2017). InDel primers (Table S2) were designed based on the flanking sequences of identified InDel positions. All of the primers were designed using Primer 3.0 (version 0.4.0) online software (http://bioinfo.ut.ee/primer3-0.4.0/).

To obtain the yvl locus, 73 F2 and 544 BC2F2 individuals exhibiting the mutant phenotype derived from crossing yvl and ZS9 were constructed as mapping populations. The total genomic DNA was isolated from fresh young leaves of rapeseed seedlings using the cetyltriethylammnonium bromide (CTAB) mini-prep method (Murray and Thompson 1980).

Amplified fragments of InDel markers were subjected to electrophoresis on high resolution agarose gel and SNPs genotyping of mapping population were visualized with either the Sanger method sequence or Kompetitive Allele Specific PCR (KASP), a high throughput SNP detection technology (Semagn et al. 2013). The Laboratory of the Government Chemist (LGC) Group was employed to design the KASP primers, which matched the SNP alleles containing two different forward primers labeled by FAM and HEX with fluorophores and one same reverse primer (Table S2).

**Gene cloning and multi-sequence alignment**

Genomic DNA of 14 open reading frames (ORFs) in the candidate region were amplified in ZS9 and yvl using gene-specific primers (Table S2) by using a PrimeSTAR GXL Kit (Takara, Japan) and cloned into the PBI121 Vector using ClonExpress II One Step Cloning Kit C112-01 (Vazyme). The recombinant plasmids were transformed into DH5α competent cells for sequencing. Multiple sequence alignment was performed with Clustal Omega online software (http://www.ebi.ac.uk/Tools/msa/clustalo/) and visualized by Sequence Manipulation Suite: Color Align Conservation (http://www.bioinformatics.org/sms2/color_align_cons.html).

**Phylogenetic and synteny analysis**

The amino acid sequences of the CHLH homologous proteins were retrieved through the Brassica Database (http://brassicadb.org/brad/index.php) and the plant genomics resource Phytozome v12.1 (https://phytozome.jgi.doe.gov/pz/portal.html) by searching with the total amino acid sequences of CHLH. The multiple sequence alignments were
conducted by ClustalX 1.83 (Thompson et al. 1997). Phylogenetic trees were constructed using MEGA 5.1 software based on the neighbor-joining (NJ) method with 1000 bootstrap trials and 111 random seed (Tamura et al. 2011).

Extract the extended upstream and downstream 30 kb of the orthologous regions covering the CHLH genes from the reference genome in Arabidopsis, B. napus, B. rapa and B. oleracea. Perform blast alignments between each species to obtain a synteny region. The visualization of synteny region was exported from Circos software.

RNA extraction, RNA analysis and qRT-PCR
Total RNA was extracted from the freshly prepared younger and older leaves at the eight-leaf stage using the RNA prep pure plant kit (TIANGEN). The RNA samples were diluted to 10 ng μl⁻¹ and analyzed by Agilent 2100 (Agilent, U.S.A.). The RNA 6000 Nano Total RNA Analysis Kit (Agilent, U.S.A.) was used for analysis. Each RNA sample (2 μg) for quantitative RT-PCR (qRT-PCR) was reverse transcribed using PrimeScript RT reagent Kit with gDNA Eraser (Takara, Japan) following the manufacturer’s instructions. The BnaActin gene (GenBank: AF111812.1) was used as an internal control. The qRT-PCR amplification was carried out in a CFX Connect Real-time PCR system (Bio-Rad, U.S.A.) by using the SYBR Green Real-time PCR Master Mix (Bio-Rad, U.S.A.) in 20 μl reaction mixture. Primers used in qRT-PCR are listed in Table S2. The 2⁻ΔΔCT method was used to analyze the relative quantification of gene expression (Livak and Schmittgen 2001). The data were expressed as the mean ± SD (three biological repeats, each with three technical repeats).

RNA-seq analysis
Total RNA samples were extracted from the younger and older leaves of ZS9 and yvl. The time of sample collection for RNA-seq was in accordance with the determination of chlorophyll content, chloroplast ultra-structure. The transcriptomes were sequenced on an Illumina HiSeq2500 platform (Illumina, U.S.A.). Comparing 218.8 million clean reads of the samples with the B. napus reference (http://www.genoscope.cns.fr/brassicanapus/data/), total 147.7 million reliable mapped reads were obtained for RNA-seq analysis (Figure S1). The cluster method was used to calculate the distance in biological repeats. Levels of gene expression were calculated using the fragments per kilo base transcript per million reads (FPKM) method (Trapnell et al. 2010). The significance level of differentially expressed genes (DEGs) was determined using p-values < 0.05 and |log2 (Fold Change) | > 1.

Statistical analysis
All statistical analyses in this study were performed using Student’s t-test. Each diagram was annotated with the number of biological repeats in each experiment. Values were considered as significantly different with the threshold of *P < 0.05 and **P < 0.01.

Data availability
Figure S1 describes RNA-seq analysis of ZS9 and yvl. Figure S2 shows polymorphic molecular markers. Figure S3, S4 and S5 illustrate sequence alignment of three genes in which SNP variation occurred. Figure S6 and S7 illustrate amplification and digestion by Bln I of ORF10. Figure S8 illustrates the domain of CHLH protein in multiple species. Figure S9 shows FPKM value of ORF10. Figure S10 illustrates the expression of chlorophyll synthesis genes. Table S1 lists the information of 629 rapeseed accessions. Table S2 lists all the primer sequence in this study. Table S3 lists the information of 629 rapeseed accessions. Table S4 describes the comparison of agronomic traits between ZS9 and yvl mutant. Table S5 describes chlorophyll contents of younger and older leaves between ZS9 and yvl. Table S5 shows the statistics of whole genome re-sequencing in
Table 1 Segregation of F2 and BC1F1 populations

| Cross          | ZS9×yvl | yvl×ZS9 | (yvl×ZS9)×yvl |
|---------------|---------|---------|---------------|
| Number of ZS9-type plants | 719 | 261 | 75 |
| Number of yvl-type plants | 208 | 85 | 84 |
| Total plants | 927 | 346 | 159 |
| ZS9-type/yvl-type | 3.45 | 3.07 | 0.89 |
| Excepted      | 3.1 | 3.1 | 1:1 |
|χ² | 3.11 | 0.02 | 0.40 |

a,b ZS9-type plants and yvl-type plants were determined by visual inspection. c χ² < χ²₀.₀₅ = 3.84, P > 0.05 was considered statistically significant.

BSA-seq study. Table S6 and S7 respectively list the functional annotation of candidate genes and genes transcribed by PEP and NEP. Supplemental material available at figshare: https://doi.org/10.25387/g3.12625796.

RESULTS

Phenotypic characterization of yvl mutant

The yvl mutant was obtained from EMS-induced B. napus ZS9. The younger leaves of mutant plant were clearly distinguished by a yellowish leaf color phenotype (Figure 1A, B). With plants growing, the younger leaves remained yellowish at the cotyledon stage and gradually became virescent toward maturity (Figure 1C, D). At maturity, the agronomically important traits showed only slight but non-significant differences between the mutant and wild type plants. Plant height and branch initiation height were slightly shorter in mutant plants but the differences were statistically insignificant (Figure 1F and Table S3).

Physiological and anatomical features of yvl mutant

Since the yellowish (chlorotic) leaves are associated with disruptions of chlorophyll biosynthesis, we estimated the chlorophyll content of younger and older leaves from the yvl mutant and ZS9 plants to determine the difference in pigment accumulation of the two. We observed that chlorophyll a (Chl a), chlorophyll b (Chl b) and carotenoid (Car) levels were remarkably lower in both younger and older leaves of yvl than in ZS9 (Figure 2A and Table S4). In the older leaves, the pigment levels seemed to recover partially but were still significantly lower in than in the wild type (Figure 2A and Table S4). This would suggest a disruption of chlorophyll biosynthesis pathway throughout the plant development.

To further investigate the effects of mutation of the chloroplast, we observed the ultrastructure of the chloroplast in both mutant and wild type under Transmission electron microscopy (TEM). The chloroplasts in ZS9 contained well-developed lamellar structures equipped with normally stacked grana and thylakoids regardless of younger and older leaves (Figure 2F-I). However, the structure of chloroplasts in yvl younger leaves was anomalously shaped with several vesicles instead of thylakoids (Figure 2B, C). The density and shape of chloroplasts of older leaves in yvl were almost similar with ZS9 leaves, but again, the granum was accumulated with abnormally developed lamellar structures composed of thylakoid membranes and lacuna (Figure 2D, E). The results suggested that the mutation(s) had clearly disturbed the chloroplast development and chlorophyll biosynthesis, which can ultimately affect the photosynthetic capacity of the plants.

Rapid delimitation of a candidate genomic region by BSA-seq

For genetic analysis of the yvl locus, two crosses ZS9×yvl and yvl×ZS9 were developed. The reciprocal F1 plants from the two crosses exhibited ZS9-like phenotypes (Figure 1E). The F2 segregation pattern fitted a 3:1 ratio of ZS9- to yvl-type plants (χ² < χ²₀.₀₅ = 3.84; P > 0.05) (Table 1). In addition, the BC1 progenies showed an expected Mendelian inheritance ratio 1:1 (χ² < χ²₀.₀₅ = 3.84; P > 0.05) (Table 1). These data indicated that the phenotype of yvl mutant was controlled by a single recessive nuclear gene.

As for the BSA-seq, 268.83 million paired-end reads of two parental lines (133.73 million reads for ZS9, and 135.10 million reads for yvl) were generated with 97.29% and 98.43% mapping rate, 19.84 x and 23.49 x sequencing depth and 91.43% and 92.73% genome coverage, respectively (Table S5). Similarly, the G- and Y-pools had an alignment of 30.43 and 26.96 x sequencing depth, and 95.99% and 94.34% genome coverage, respectively (Table S5). The statistics results suggested that the WGS data were reliable and could be used for subsequent mutation detection and BSA-seq analysis.

Figure 3 BSA-seq approach applied for mapping genomic region. (A) The innermost two circles (a and b) respectively denoted genome-wide densities of SNPs and InDels between ZS9 and yvl based on WGS data. The outermost circles denoted the different physical size of 19 chromosomes. The green (c) and yellow (d) circles displayed the SNP-index of yvl-type leaf color pool and ZS9-type leaf color pool respectively. The red (e) circle showed ΔSNP-index. (B) SNP index plot of yvl-type leaf color pool (top), ZS9-type leaf color pool (middle) and ΔSNP-index plot (bottom) of chromosome A03. The significant genomic region (1.0–4.0 Mb) was highlighted in shaded yellow color.
Based on the results of genotyping and filtration, the homozygous 182,078 SNPs and 91,553 InDels polymorphic markers were obtained from the two parental lines. The density of SNPs and InDels in 

revealed that the variations were evenly distributed throughout the whole genome (Figure 3A). The visualized SNP-index plots of G- and Y-pools were delineated, followed by the depiction of D

SNP-index. The SNP-index plots were similar across the entire genome for the two pools, except for a single region on chromosome A03 ranging 1.0–4.0 Mb (Figure 3A, B). Therefore, this region on chromosome A03 was considered as a unique candidate region of the yvl locus (Figure 3B).

**Map-based cloning of the yvl gene**

According to the data of WGS, eleven positive molecular markers, including 5 SNP and 4 InDel markers in the candidate region, as well as 2 SNP flanking markers, were developed (Figure S2). Utilizing the developed markers and 72 F₂ homozygous yvl-type individuals, the yvl locus was narrowed down to 591-kb on chromosome A03 between marker yvl-3 and InDel-9 (Figure 4A). Further mapping using 544 BC₂F₂ yvl-type individuals and linked markers indicated that the yvl locus was restricted to a 70-kb region flanked by yvl-O10 and InDel-O6 (Figure 4A).

Based on the annotations of B. napus genome database (http://www.genoscope.cns.fr/brassicanapus/data/), this region comprised of 14 putative genes with annotated functions (Figure 4A and Table S6). The alignment showed no sequence variations in the coding region between ZS9 and yvl of other ORFs except ORF5, ORF10 and ORF14 (Figure S3-S5). The SNP variations that occurred in ORF5 and ORF14 were synonymous mutations. Whereas, the mutation at SNP marker yvl-8, which co-segregated with the yvl phenotype, was a T substitution C in BnaA03g04440D, generating a premature stop codon and a Cleaved Amplified Polymorphism Sequences (CAPS) marker of the restriction enzyme Bln I, yvl and ZS9 are the control in the two gels respectively. The grouping of gels were cropped from different parts of the same gel and the original full-length gel was displayed in Figure S7. The marker used in (B) and (D) was 1kb ladder marker.

---

**Figure 4** Map-based cloning of the yvl gene. (A) The yvl locus was initially mapped to a 3.0 Mb of chromosome A03 based on BSA-seq. Finally the yvl locus was limited to a 70-kb region linked with yvl-O10 and InDel-O6 using 73 F₂ and 544 BC₂F₂ mutant phenotype individuals. (B) Gene structure of BnaA03.yvl and confirmation the mutation site by Sanger sequencing and enzyme digestion analysis. The positions of ATG and TAG are the start and stop codons. Black boxes and white box represent exons and 3’ UTR region, respectively. Green shade and red box indicates the ‘C’ to ‘T substitution and codon change. Verify the change with CASP marker by Bln I digestion. The original full-length gel was displayed in Figure S6. (C) SNP variation among 629 rapeseed accessions and 20 BC₃ F₁ plants. Red, green, and blue dots represent three types SNP variation (T:T, T:C, and C:C). Black dots are the negative control. (a) yvl mutant and 10 yvl-type BC₃F₁ plants; (b) 10 ZS9-type BC₃F₁ plants; (c) 628 rapeseed accessions. (D) 10 yvl-type (the bottom) and 10 ZS9-type (the top) of BC₂F₂ plants digested by Bln I, yvl and ZS9 are the control in the two gels respectively. The grouping of gels were cropped from different parts of the same gel and the original full-length gel was displayed in Figure S7. The marker used in (B) and (D) was 1kb ladder marker.
We genotyped 10 homozygous yvl-type and 10 heterozygous ZS9-type individuals derived from BC3F1 were genotyped by this causal SNP substantiating the allelic variation (Figure 4C, D). The causal SNP polymorphism within ORF10 was found to be perfectly consistent with the normal leaf color among the 628 accessions including ZS9 (Figure 4C and Table S1), suggesting that this site maybe highly conservative in *B. napus*. Taken together, these results suggest that BnaA03.CHLH is the most likely candidate gene responsible for the yvl phenotype.

**Characterization of BnaA03.CHLH**

To detect the homologous genes of BnaA03.CHLH in *B. napus* and their evolutionary relationship, a phylogenetic tree was constructed including Brassicaceae species and four other representative species. Phylogenetic analysis clearly distinguished monocotyledons (*Oryza sativa* and *Zea mays*) from dicotyledons (*Glycine max*, *Arabidopsis thaliana*, *B. rapa*, *B. oleracea* and *B. napus*). In Brassicaceae species, the evolutionary relationship of CHLH conforms with U’s triangle (Figure 5A). At the chromosomal level, the syntenic regions harboring CHLH genes in *Arabidopsis*, *B. rapa*, *B. oleracea* and *B. napus* were delineated and the result of syntenic relationship was in accordance with the evolutionary relationship (Figure 5B).

It was reported that CHLH consists of two functional regions, the N-terminal region and cage-like assembly region, which are respectively composed of domains I-II and domains III-VI, as well as domain VI is a tight α-helix bundle in C-terminus (Chen et al. 2015). The sequence alignment of CHLH in monocotyledons (*Oryza sativa*, *Zea mays*) and dicotyledons (*Glycine max*, *Arabidopsis thaliana*, *B. rapa*, *B. oleracea* and *B. napus*) showed that the mutant SNP in domain VI was conserved (Figure S8). And then, we analyzed the protein structure prediction of CHLH in an online software Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/). The conformation of...
CHLH protein in ZS9 showed that the last α-helix of C-terminus was from 1360 (Glu) to 1378 (Glu) in the amino acid sequence; however, all amino acids were deleted after 1371 (Ser) in yvl, resulting in a deformed/truncated CHLH protein in yvl (Figure 5C).

To investigate the expression pattern of BnaA03.CHLH in detail, its transcription levels in the root, stem, leaf, flower and silique tissues of ZS9 and yvl were analyzed by qRT-PCR. The expression level of BnaA03.CHLH was highest in the leaf, followed by flower, silique, root and stem in a descending order, which could be understandable due to an abundance of chloroplasts in the green plant tissues either in ZS9 or yvl (Figure 5D). In addition, it expressed higher in all tissues of ZS9 than in yvl (Figure 5D). Apart from qRT-PCR, we further analyzed the expression pattern of BnaA03.CHLH using RNA-seq data of Zhongshuang11 tissues (including root, stem, leaf, bud, silique, sepal, ovule, pericarp and callus), previously published by our group (Li et al. 2019; Yao et al. 2020). The transcript abundance level of BnaA03.CHLH in leaf was consistent with the relative expression patterns observed in qRT-PCR (Figure S9). Thus, the expression pattern of BnaA03.CHLH showed consistency with the leaf color changes, further validating the candidacy of this gene for mutant phenotype.

**Mutation in BnaA03.CHLH degraded chloroplast structure and chlorophyll biosynthesis by limiting PEP activity**

Chlorophyll content and chloroplast development were critically affected in yvl mutant. From RNA-seq data, we observed that the expression of class I plastid genes, which are transcribed by PEP was considerably reduced in the mutant plants, especially at the initial stages of development (Figure 6A and Table S7). The regulatory expression of other plastid and chlorophyll synthesis related genes was changed either slightly or insignificantly (Figure 6B and Figure S10). This indicated that the PEP activity may have been impaired to a greater extent, which led to the impairment of chloroplast development and chlorophyll degradation in yvl. Similar with these results, the decreased expression of class I plastid genes and increased expression of class III plastid genes was previously reported in WP1 mutant, where the regulation of rRNA had also reduced significantly, especially in younger leaves (Wang et al. 2016c). Therefore, we also tested the rRNA levels in both younger and older leaves. rRNA analysis showed that there was no significant difference in the contents of 16S, 18S, 23S and 25S rRNA (Figure 7A). The expression levels of rRNA and ribosomal protein coding genes were slightly higher in yvl (Figure 7B, C). Therefore, we concluded that the chloroplast ribosome was not affected in the mutant, and the degradation of chloroplast ultrastructure occurred due to reduced PEP activity.

**DISCUSSION**

Yellow leaf mutants are often associated with disruptions of photosynthetic pathway, including the regulatory network of chlorophyll biosynthesis and chloroplast development genes. In the present study, we isolated a yellow-virescent leaf mutant in rapeseed and identified the causal mutant allele on chromosome A03 (BnaA03.CHLH).
CHLH is a multifunctional gene involved in chlorophyll biosynthesis and the retrograde signaling pathway in higher plants (Adhikari et al. 2011; Chen et al. 2017). A previous study of CHLH mutant allele in Arabidopsis reported that mutations in the central protein domains (I to VI) resulted in reduction of Mg-chelatase activity and plastid signaling (Ibata et al. 2016). Here, we observed that the mutation in BnaA03.CHLH domain VI led to yellowish leaf color phenotype in young leaves and the mutation site was conservative in 629 rapeseed accessions (Figure 4B, C). The protein 3D structure showed that the last α-helix of the C-terminal region was incomplete in yvl mutant (Figure 5C), resulting in a yellowish phenotype. Mutants with yellowish leaf phenotypes are generally known to have degraded chlorophyll synthesis, which halts photosynthetic machinery of cells, ultimately leading to yield losses or cell death in severe cases. We observed reductions of chlorophyll content and chloroplast development predominantly in the younger leaves (Figure 2). In the older leaves of yvl mutant, the regulation of genes associated with chlorophyll synthesis, chloroplast development and light harvest chlorophyll a/b-binding (LHCA and LHCB) was not significantly inhibited (Figure 6A, C and Figure S10). This resulted in a delayed greenness in mutants, but the chlorophyll a/b and carotenoid contents were still considerably lower as compared to ZS9.

Delayed greenness has also been observed previously in other crops, for example, in rice, many leaf color mutants with yellowish younger leaves and greenish older leaves are reported (Wang et al. 2016a; Wan et al. 2015; Gong et al. 2014). The genes controlling leaf color in most of these mutants were identified to be essential for early chloroplast development, which is reasonable to explain the change of leaf color from yellow to green (Wu et al. 2007). Similarly, in this study, one assumption could be that the mutations in BnaA03.CHLH did not completely halt the chlorophyll synthesis process but slowed it down, and as the plants grew older, the chlorophyll synthesis recovered and a greenish leaf color phenotype was restored. Another plausible and rather interesting reason for the re-emergence of phenotype in older leaves could be associated to the gene expression patterns in polyploid genomes. In polyploid crops such as rapeseed, homologous expression bias and effects of dosage balance on gene expression are common phenomena (Yoo et al. 2014; Wang et al. 2006; Rapp et al. 2009; Stupar et al. 2007; Wu et al. 2018). For BnaA03.CHLH gene, we found a homologous gene on C-subgenome.

Figure 7 The rRNA analysis and expression levels of chloroplast ribosomal protein gene. (A) rRNAs analysis by Agilent 2100. (B) The expression levels of rRNAs in younger and older leaves of ZS9 and yvl. (C) Different expression of 58 chloroplast ribosomal protein genes according to RNA-seq data. The graph shows the log2 ratio of transcript levels in yvl compared with ZS9.
(Figure 5A, B). It could be assumed that redundant functional activity of the homologous gene BnaA03.CHLH restored the virescent leaf color phenotype in older leaves when the function of BnaA03.CHLH was halted due to the mutations in central protein domains. This could be studied further thoroughly in the future studies to understand the gene expression behavior of polyploid genomes.

Chlorophyll biosynthesis and chloroplast development are complex pathways and their regulatory mechanisms involve a multitude of genes and gene networks. Disruption in these pathways, even by point mutations, bring about changes in their entire regulatory networks, as has been reported in several studies in past (Wang et al. 2016c; Ge et al. 2017). Previous study of Wang and coworkers associated such phenotype changes in rice with an inhibitory expression of class-I plastid genes, which are transcribed by PEP. They further observed that the ribosomal proteins and rRNA levels were reduced in mutants, which resulted in degradation of chloroplast ultrastructure and chlorophyll biosynthesis (Wang et al. 2016c). We did not observe any considerable changes in the rRNA levels in chloroplast, but the class-I plastid genes including the regulatory genes of photosystem I and II were substantially reduced in the mutant plants. Since photosystem II is light harvesting complex and the initiating point of photosynthetic mechanism, reduced abundance of its regulatory transcripts caused disruptions in the entire pathway. Therefore, we conclude that a single nucleotide mutation in BnaA03.CHLH results in an imperfect PEP activity causing a reduced chlorophyll accumulation and a deformed chloroplast in rapeseed.

ACKNOWLEDGMENTS

This research was supported by the National Key Research and Development Program of China (2016YFD0101007, 2018YFD0200904, 2016YFD100305, 2018YFE0108000), the National Natural Science Foundation of China (31770250), Natural Science Foundation of Hubei Province (2017CFB628),the Earmarked Fund for China Agriculture Research System (CARS-12), the Agricultural Science and Technology Innovation Program (ASTIP) of Chinese Academy of Agricultural Sciences, the Agricultural Scientific and Technological Research Projects of Guizhou province ([2019] 2397).

LITERATURE CITED

Adhikari, N. D., J. E. Frolich, D. D. Strand, S. M. Buck, D. M. Kramer et al., 2011 GUN4-porphyrin complexes bind the ChlH/GUN5 subunit of Mg-Chelatase and promote chlorophyll biosynthesis in Arabidopsis. Plant Cell 23: 1449–1467. https://doi.org/10.1105/tpc.110.082503

Arnon, D. I., 1949 Copper enzymes in isolated chloroplasts: polyphenoloxidase. Plant Physiol. 24: 1–15. https://doi.org/10.1104/pp.24.1.1

Barajas-Lopez Ida, D., N. E. Blanco, and A. Strand, 2013 Plastid-to-nucleus communication, signals controlling the running of the plant cell. Biochim. Biophys. Acta 1833: 425–437. https://doi.org/10.1016/j.bbamrc.2012.06.020

Beck, C. F., 2001 Signaling pathways in chloroplast-to-nucleus communi-

cation. Protoplast 152: 175–182. https://doi.org/10.1007/s00709-001-0056

Borner, T., A. Y. Aleyunkova, Y. O. Zabo, and V. K. Kunetsov, 2015 Chloroplast RNA polymerases: Role in chloroplast biogenesis. Biochim. Biophys. Acta 1847: 761–769. https://doi.org/10.1016/j.bba bio.2015.02.004

Chalhoub, B., F. Denoue, S. Liu, I. A. P. Parkin, H. Tang et al., 2014 Plant genetics Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345: 950–953. https://doi.org/10.1126/science.1253435

Chen, S. T., N. Y. He, J. H. Chen, and F. Q. Guo, 2017 Identification of core subunits of photosystem II as action sites of HSP21, which is activated by the GUN5-mediated retrograde pathway in Arabidopsis. Plant J. 89: 1106–1118. https://doi.org/10.1111/pj.13447

Chen, X., H. Fu, Y. Fang, X. Wang, S. Zhao et al., 2015 Crystal structure of the catalytic subunit of magnesium chelatase. Nat. Plants 1: 15125. https://doi.org/10.1038/nplants.2015.125

Czarnecki, O., and B. Grimm, 2012 Post-translational control of tetrapyrrole biosynthesis in plants, algae, and cyanobacteria. J. Exp. Bot. 63: 1675–1687. https://doi.org/10.1093/jxb/err437

Ge, C., L. Wang, W. Ye, L. Wu, Y. Cui et al., 2017 Single-point Mutation of an Histidine-aspartic Domain-containing Gene involving in Chloroplast Ribosome Biogenesis Leads to White Fine Stripe Leaf in Rice. Sci. Rep. 7: 3298. https://doi.org/10.1038/s41598-017-03327-2

Gibson, L. C., R. D. Willows, C. G. Kannangara, D. von Wettstein, and C. N. Hunter, 1995 Magnesium-protoporphyrin chelatase of Rhodobacter sphaeroides: reconstitution of activity by combining the products of the bchH,-I, and -D genes expressed in Escherichia coli. Proc. Natl. Acad. Sci. USA 92: 1941–1944. https://doi.org/10.1073/pnas.92.6.1941

Gong, X., Q. Su, D. Lin, Q. Fang, J. Xu et al., 2014 The rice OsV4 encoding a novel pentatricopeptide repeat protein is required for chloroplast development during the early leaf stage under cold stress. J. Integr. Plant Biol. 56: 400–410. https://doi.org/10.1111/jipb.12138

Hernández-Verdeja, T., and A. Strand, 2018 Retrograde Signals Navigate the Path to Chloroplast Development. Plant Physiol. 176: 967–976. https://doi.org/10.1104/pp.17.01299

Hsieh, W. Y., J. C. Liao, H. T. Wang, T. H. Hung, C. C. Tseeng et al., 2017 The Arabidopsis thiamin-deficient mutant pale green1 lacks thiamin monophosphate phosphatase of the vitamin B1 biosynthesis pathway. Plant J. 91: 145–157. https://doi.org/10.1111/pj.13552

Ibata, H., A. Nagatani, and N. Mochizuki, 2016 CHLH/GUN5 Function in Tetrapyrrole Metabolism Is Correlated with Plastid Signaling but not ABA Responses in Guard Cells. Front Plant Sci 7: 1650. https://doi.org/10.3389/fpls.2016.01650

Jung, K. H., J. Hur, C. H. Ryu, Y. Choi, Y. Y. Chung et al., 2003 Characterization of a Rice Chlorophyll-Deficient Mutant Using the T-DNA Gene-Trap System. Plant Cell Physiol. 44: 463–472. https://doi.org/10.1093/pcp/pcg064

Li, C., Y. Hu, R. Huang, X. Ma, Y. Wang et al., 2015 Mutation of FdC2 gene encoding a ferredoxin-like protein with C-terminal extension causes yellow-green leaf phenotype in rice. Plant Sci. 238: 127–134. https://doi.org/10.1016/j.plantsci.2015.06.010

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan et al., 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Y., C. Dong, M. Hu, Z. Bai, C. Tong et al., 2019 Identification of Flower-Specific Promoters through Comparative Transcriptome Analysis in Brassica napus. Int. J. Mol. Sci. 20: 5949. https://doi.org/10.3390/ijms20235949

Lívak, K. J., and T. D. Schmittgen, 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT Method. Methods 25: 402–408. https://doi.org/10.1016/S1046-8727(01)00262-6

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis et al., 2010 The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297–1303. https://doi.org/10.1101/gr.107524.110

Mochizuki, N., J. A. Brusslan, R. Larkin, A. Nagatani, and J. Chory, 2001 Arabidopsis genomes uncoupled 5 (GUNS) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc. Natl. Acad. Sci. USA 98: 2053–2058. https://doi.org/10.1073/pnas.98.4.2053

Mochizuki, N., R. Tanaka, A. Tanaka, T. Masuda, and A. Nagatani, 2008 The steady-state level of Mg-protoporphyrin IX is not a determinant of plastid-to-nucleus signaling in Arabidopsis. Proc. Natl. Acad. Sci. USA 105: 15184–15189. https://doi.org/10.1073/pnas.0802425105

Moynihan, A., C. M. McCormac, M. J. Terry, and A. G. Smith, 2008 Tetrapyrrole profiling in Arabidopsis seedlings reveals that retrograde plastid nuclear signaling is not due to Mg-protoporphyrin IX
accumulation. Proc. Natl. Acad. Sci. USA 105: 15178–15183. https://doi.org/10.1073/pnas.0803054105
Murray, M. G., and W. F. Thompson. 1980 Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8: 4321–4325. https://doi.org/10.1093/nar/8.19.4321
Pogson, B. J., and V. Albrecht. 2011 Genetic dissection of chloroplast biogenesis and development: an overview. Plant Physiol. 155: 1545–1551. https://doi.org/10.1109/pp.1170365
Rapp, R. A., J. A. Udall, and J. F. Wendel. 2009 Genomic expression dominance in allopolyploids. BMC Biol. 7: 18. https://doi.org/10.1186/1741-7007-7-18
Ruan, B., Z. Gao, J. Zhao, B. Zhang, A. Zhang et al. 2017 The rice YGL gene encoding an Mg²⁺-chelatase ChiD subunit is affected by temperature for chlorophyll biosynthesis. J. Plant Biol. 60: 314–321. https://doi.org/10.1007/s12374-016-0596-0
Semagn, K., R. Babu, S. Hearne, and M. Olsen. 2013 Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol. Breed. 33: 1–14. https://doi.org/10.1007/s11032-013-9917-x
Shiina, T., Y. Tsunoyama, Y. Nakahira, and M. S. Khan. 2005 Plastid RNA polymerases, promoters, and transcription regulators in higher plants. Int. Rev. Cytol. 244: 1–515. https://doi.org/10.1007/s10321018-1551
Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei et al., 2011 MEGAS: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739. https://doi.org/10.1093/molbev/msr121
TerBush, A. D., Y. Yoshida, and K. W. Osteryoung. 2013 FtsZ in chloroplast division: structure, function and evolution. Curr. Opin. Cell Biol. 25: 461–470. https://doi.org/10.1016/j.jcb.2013.04.006
Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997 The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882. https://doi.org/10.1093/nar/25.24.4876
Trappin, C. B., A. Williams, G. Pertea, A. Mortazavi, G. Kwan et al., 2010 Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28: 511–515. https://doi.org/10.1038/nbt.1621
Vitha, S., R. S. McAndrew, and K. W. Osteryoung. 2001 FtsZ ring formation at the chloroplast division site in plants. J. Cell Biol. 153: 111–120. https://doi.org/10.1083/jcb.153.1.111
Wan, C., C. Li, X. Ma, Y. Wang, C. Sun et al., 2015 GRY79 encoding a putative metallo-beta-lactamase-trihelix chimera is involved in chloroplast development at early seedling stage of rice. Plant Cell Rep. 34: 1353–1363. https://doi.org/10.1007/s00207-015-1792-y
Wang, J., L. Tian, H. S. Lee, N. E. Wei, H. Jieng et al., 2006 Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. Genetics 172: 507–517. https://doi.org/10.1534/genetics.105.047894
Wang, L., C. Wang, Y. Wang, M. Niu, Y. Ren et al., 2016a WSL3, a component of the plastid-encoded plastid RNA polymerase, is essential for early chloroplast development in rice. Plant Mol. Biol. 92: 581–595. https://doi.org/10.1007/s11103-016-0533-0
Wang, Y., Y. He, M. Yang, J. He, P. Xu et al., 2016b Fine mapping of a dominant gene conferring chlorophyll-deficiency in Brassica napus. Sci. Rep. 6: 31419. https://doi.org/10.1038/srep31419
Wang, Y., C. Wang, M. Zheng, J. Lyu, Y. Xu et al., 2016c WHITE PANICLE1, a Val-tRNA Synthetase Regulating Chloroplast Ribosome Biogenesis in Rice, Is Essential for Early Chloroplast Development. Plant Physiol. 170: 2110–2123. https://doi.org/10.1093/pp.15.01949
Wang, Y., J. Xie, H. Zhang, B. Guo, S. Ning et al., 2017 Mapping stripe rust resistance gene YrZH22 in Chinese wheat cultivar Zhoumai 22 by bulked segregant RNA-Seq (BSR-Seq) and comparative genomics analyses. Theor. Appl. Genet. 130: 2191–2201. https://doi.org/10.1007/s00122-017-4250-0
Wu, J., L. Lin, M. Xu, P. Chen, D. Liu et al., 2018 Homoeolog expression bias and expression level dominance in resynthesized allopolyploid Brassica napus. BMC Genomics 19: 586. https://doi.org/10.1186/s12864-016-4966-5
Wu, Z., X. Zhang, B. He, L. Diao, S. Sheng et al., 2007 A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. Plant Physiol. 145: 29–40. https://doi.org/10.1104/pp.107.100321
Yao, S., F. Liang, R. A. Gill, J. Huang, X. Cheng et al., 2020 A global survey of the transcriptome of the allopolyploid Brassica napus based on single molecule long-read isoform sequencing and illumina-based RNA-seq data. Plant J. (Preprint posted April 9, 2020). https://doi.org/10.1111/tpj.14754
Yoo, M. J., X. Liu, J. C. Pires, P. S. Soltis, and D. E. Soltis. 2014 Nonadditive gene expression in polyploids. Annu. Rev. Genet. 48: 485–517. https://doi.org/10.1146/annurev-genet-120113-092159
Yu, Q. B., Y. Lu, Q. Ma, T. T. Zhao, C. Huang et al., 2013 TAC7_, an essential component of the plastid transcriptionally active chromosome complex, interacts with FLN1, TAC10, TAC12 and TAC14 to regulate chloroplast gene expression in Arabidopsis thaliana. Physiol. Plant. 148: 408–421. https://doi.org/10.1111/j.1399-3054.2012.01718.x
Zhu, L., Z. Yang, X. Zeng, J. Gao, J. Liu et al., 2017 Mapping stripe rust resistance gene YrZH22 in Chinese wheat cultivar Zhoumai 22 by bulked segregant RNA-Seq (BSR-Seq) and comparative genomics analyses. Theor. Appl. Genet. 130: 2191–2201. https://doi.org/10.1007/s00122-017-4250-0
Communicating editor: J. Ma