




Article

Differentially Expressed Genes Associated with the Cabbage Yellow-Green-Leaf Mutant in the *ygl-1* Mapping Interval with Recombination Suppression

Xiaoping Liu [†], Hailong Yu [†], Fengqing Han, Zhiyuan Li, Zhiyuan Fang, Limei Yang, Mu Zhuang, Honghao Lv , Yumei Liu, Zhansheng Li, Xing Li and Yangyong Zhang ^{*}

Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Beijing 100081, China; 82101181120@caas.cn (X.L.); yuhailong@caas.cn (H.Y.); feng857142@163.com (F.H.); 82101181071@caas.cn (Z.L.); fangzhiyuan@caas.cn (Z.F.); yanglimei@caas.cn (L.Y.); zhuangmu@caas.cn (M.Z.); lvhonghao@caas.cn (H.L.); liuyumei@caas.cn (Y.L.); lizhansheng@caas.cn (Z.L.); xzdlixing@126.com (X.L.)

^{*} Correspondence: zhangyangyong@caas.cn; Tel.: +86-010-8210-8756

[†] These authors contributed equally to this work.

Received: 12 August 2018; Accepted: 20 September 2018; Published: 27 September 2018



Abstract: Although the genetics and preliminary mapping of the cabbage yellow-green-leaf mutant YL-1 has been extensively studied, transcriptome profiling associated with the yellow-green-leaf mutant of YL-1 has not been discovered. Positional mapping with two populations showed that the yellow-green-leaf gene *ygl-1* is located in a recombination-suppressed genomic region. Then, a bulk segregant RNA-seq (BSR) was applied to identify differentially expressed genes (DEGs) using an F₃ population (YL-1 × 11-192) and a BC₂ population (YL-1 × 01-20). Among the 37,286 unique genes, 5730 and 4118 DEGs were detected between the yellow-leaf and normal-leaf pools from the F₃ and BC₂ populations. BSR analysis with four pools greatly reduced the number of common DEGs from 4924 to 1112. In the *ygl-1* gene mapping region with suppressed recombination, 43 common DEGs were identified. Five of the DEGs were related to chloroplasts, including the down-regulated *Bo1g087310*, *Bo1g094360*, and *Bo1g098630* and the up-regulated *Bo1g059170* and *Bo1g098440*. The *Bo1g098440* and *Bo1g098630* genes were excluded by qRT-PCR. Hence, we inferred that these three DEGs (*Bo1g094360*, *Bo1g087310*, and *Bo1g059170*) in the mapping interval may be tightly associated with the development of the yellow-green-leaf mutant phenotype.

Keywords: cabbage; yellow-green-leaf mutant; recombination-suppressed region; bulk segregant RNA-seq; differentially expressed genes

1. Introduction

Yellow-green-leaf mutants have been extensively studied in many species, including *Arabidopsis thaliana* [1], barley [2], *Brassica napus* [3], rice [4–6], cabbage [7], and muskmelon [8]. Leaf color mutants are an ideal model for studying mechanisms of photosynthesis and light morphogenesis, since yellow-green-leaf mutants are commonly related to chlorophyll synthesis or degradation [9,10].

Chlorophyll is the most important pigment related to photosynthesis. In *Arabidopsis*, 27 genes involved in 15 steps in the pathway from glutamyl-tRNA to chlorophylls a and b have been identified. Leaf color mutants commonly result from blocking a portion of the chlorophyll synthesis pathway, such as the synthesis of 5-aminolevulinic acid (ALA) [11]. Runge et al. [12] isolated and classified some chlorophyll-deficient xantha mutants of *Arabidopsis thaliana* and found that some of the mutants were blocked at various steps of the chlorophyll pathway between ALA and protochlorophyllide (Pchlide), and the latter did not accumulate in the dark.

Bulked segregant analysis (BSA) is a powerful strategy that is commonly used in gene mapping [13]. Futschik and Schlötterer showed that sequencing of pools of samples from individuals are often more effective for Single Nucleotide Polymorphisms (SNP) discovery and provide more accurate allele frequency estimates [14]. Typically, two populations are used for BSA: a backcross (BC) population [15,16] and an F_2 population [17,18]. Mackay and Caligari [19] found that quantitative trait loci (QTLs) are more easily detected in BC populations than in F_2 populations.

In recent years, transcriptome analysis based on deep RNA sequencing (RNA-seq) has been used for the estimation of genome-wide gene expression levels [20,21]. Transcriptome sequencing encompasses mRNA transcript expression analysis. Combined RNA-seq analysis can be used for purposes such as novel transcript prediction, gene structure refinement, alternative splicing analysis, and SNP/InDel analysis [22]. Bulk segregant RNA-seq (BSR) has been applied to identify differentially expressed genes (DEGs) and trait-associated SNPs [23,24].

A yellow-green-leaf mutant (YL-1) was discovered in cabbage [10], and measurements of photosynthetic pigment contents, chloroplast ultrastructure, and chlorophyll fluorescence parameters indicated that YL-1 was deficient in its total chlorophyll content [10]. In a previous study, we mapped *ygl-1*, which controls the yellow-green-leaf phenotype, to chromosome C01 [7]. The linkage distance of the mapping interval was only 0.75 cM, but the physical distance in the reference genome TO1000 was ~10 Mb, indicating that recombination suppression existed in this interval. In this study, the recombination-suppressed region was identified by gene mapping. Two runs of BSR were performed using BC and F_3 populations, with the aim of obtaining DEGs associated with the yellow-green-leaf mutant.

2. Materials and Methods

2.1. Plant Materials

Group I: The F_2 , BC_1P_1 , and F_3 populations were constructed using as parents the yellow-green-leaf cabbage mutant YL-1 (P_1) and the normal green leaf cabbage inbred line 01-20 (P_2). The F_2 , BC_1P_1 population was employed for *ygl-1* mapping.

Group II: The BC_1P_1 and BC_2P_1 populations were constructed using as parents the mutant YL-1 (P_1) and the normal green leaf Chinese kale inbred line 11-192 (P_3) (Supplementary Figure S1). The BC_2P_1 population was employed for *ygl-1* mapping.

The F_3 population in group I and the BC_2 population in group II were used for RNA-seq analysis. All plant materials came from the Cabbage and Broccoli Research Group, the Institute of Vegetables and Flowers (IVF), and the Chinese Academy of Agricultural Sciences (CAAS).

2.2. Identification of Recombination Suppression in the *ygl-1* Gene-Mapping Interval

The sequences of 24 markers from the 02-12 reference genome (Supplementary Table S1) were aligned to chromosome C01 and the scaffold of the TO1000 reference genome [25] (Figure 1). Based on this alignment, we propose that possible assembly errors might exist in the 02-12 reference genome. Hence, InDel primers designed based on the TO1000 reference genome were applied for further mapping. The rates of recombination in the two populations were compared with the normal level in the cabbage genome (~600 kb/cM) to analyze the recombination-suppressed region.

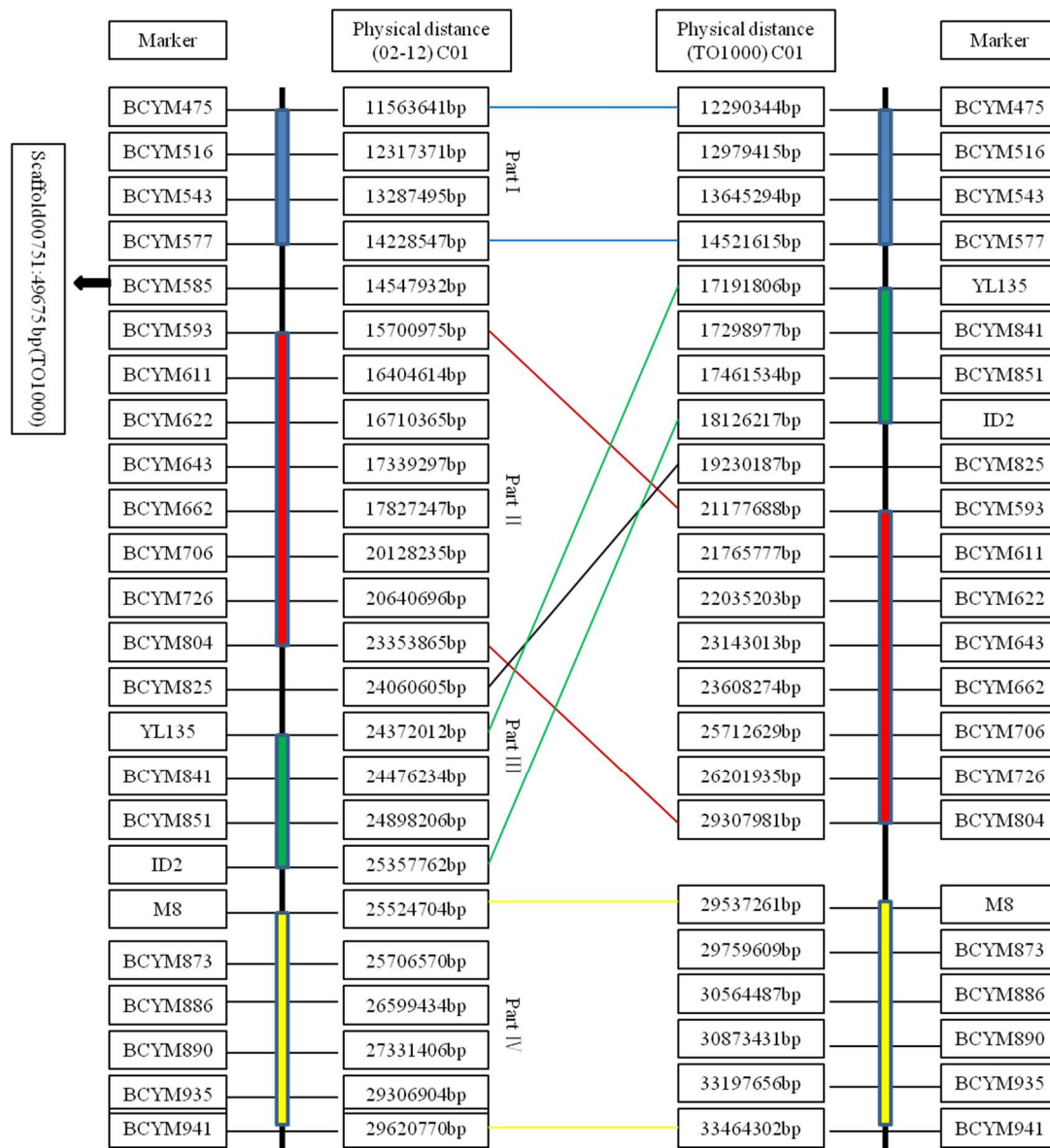


Figure 1. The physical distances of 24 InDel markers in the two reference genomes (02-12 and TO1000).

2.3. BSA, RNA Isolation, and Library Construction

Before RNA isolation, leaf samples from the two populations (the F₃ population in group I and the BC₂ population in group II) were harvested to prepare four bulk groups: Bulk F₃_yellow (consisting of equal amounts of leaf tissues from 20 yellow-green-leaf F₃ individuals), Bulk F₃_normal (20 normal-green-leaf F₃ individuals), BC₂_yellow (20 yellow-green-leaf BC₂ individuals), and BC₂_normal (20 normal-green-leaf BC₂ individuals).

Total RNA extraction was performed according to instructions of the manufacturer of the TIANGEN kit employed for extraction (Invitrogen, Carlsbad, CA, USA). RNA purity was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA), 1% formaldehyde gel electrophoresis, and a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

A total amount of 1 µg of RNA per sample was employed for RNA sample preparation. Sequencing libraries were generated using the NEBNext[®] Ultra[™] RNA Library Prep Kit for Illumina[®]

(Illumina, CA, USA) following the manufacturer's recommendations. The cDNA library products were sequenced in a paired-end flow cell using an Illumina HiSeq™ 2000 system.

3. Data Analysis

Reads containing adaptor sequences, low-quality reads (bases with more than 50% of quality scores ≤ 5), and unknown bases ($>5\%$ N bases) were removed from each dataset to obtain more reliable results, because such data negatively affect bioinformatics analyses. The sequencing reads were then aligned to the reference database for the *B. oleracea* genome (TO1000) (http://plants.ensembl.org/Brassica_oleracea/Info/Index) (accessed on 5 May 2017) [25] using HISAT [26]. Differential expression analysis to identify DEGs was performed using DESeq [27], with a threshold q value (or false discovery rate [FDR]) < 0.01 & $|\log_2(\text{fold change})| > 1$ for significant differential expression. DEGs were displayed using Circos v0.66 [28]. GO (<http://www.geneontology.org/>) (accessed on 7 May 2017) [29] enrichment analysis of the DEGs was implemented using GSeq, in which gene length bias was corrected. GO functional analysis provides GO functional classifications and annotations for DEGs. Various genes usually cooperate with each other to exercise their biological functions. A pathway-related database was therefore obtained based on Kyoto Encyclopedia of Genes and Genome (KEGG) results (<http://www.genome.jp/kegg/>) (accessed on 11 May 2017) [30].

Gene Expression Validation

DEGs associated with the yellow-green-leaf mutant were subjected to quantitative real-time RT-PCR (qRT-PCR) analysis. The primers designed according to the gene CDS sequences using DNAMAN are listed in Supplementary Table S6. Three technical replicates were performed for each gene. First-strand cDNA was synthesized using the PrimeScript™ RT reagent Kit (TAKARA BIO, Inc., Shiga, Japan). qRT-PCR was performed with the SYBR Premix Ex Taq™ Kit (Takara, Dalian, China) with the following cycling parameters: 95 °C for five min, followed by 40 cycles of 95 °C for 10 s and 55 °C for 30 s, with a final cycle of 95 °C for 15 s, 55 °C for 60 s, and 95 °C for 15 s. Relative transcription levels were analyzed using the $2^{-\Delta\Delta Ct}$ method [31]. qRT-PCR was performed in a BIO-RAD CFX96 system (Bio-Rad, Hercules, CA, USA), and the actin gene was employed as the internal control [32].

4. Results

4.1. Identification of the Recombination-Suppressed Region

In a previous study [7], we mapped *ygl-1*, which controls the yellow-green-leaf phenotype, to chromosome C01 using a population derived from YL-1 and 01-20. The *ygl-1* gene is flanked by the InDel markers ID2 and M8, and the interval between these two markers is 167 kb (C01: 25,357,762–25,524,704 bp) in the 02-12 reference genome.

However, these two markers are anchored to the TO1000 reference genome, in which the interval between ID2 (C01: 18,126,217 bp) and M8 (C01: 29,537,261 bp) is 11.41 Mb, which is approximately 680 times greater than the distance (167 kb) in the 02-12 reference genome. Then, 24 markers from the 02-12 reference genome (Supplementary Table S1) were aligned to chromosome C01 and the scaffold of the TO1000 reference genome (Figure 1). In the 02-12 reference genome, the physical interval between BCYM475 (11,563,641 bp) and BCYM941 (29,620,770 bp) could be divided into four parts [Part I: BCYM475 (11,563,641 bp) to BCYM577 (14,228,547 bp); Part II: BCYM593 (15,700,975 bp) to BCYM804 (23,353,865 bp); Part III: YL135 (24,372,012 bp) to ID2 (25,357,762 bp); and Part IV: BCYM873 (25,706,570 bp) to BCYM941 (29,620,770 bp)]. The physical locations of Part I and Part IV in the two reference genomes were parallel. However, the physical locations of Part II and Part III were opposite. The makers' order of linkage map was consistent with the physical map order of TO1000 reference genome but not 02-12 reference genome. Therefore, we proposed that an assembly error might exist in the 02-12 reference genome.

InDel primers designed based on the TO1000 reference genome were then applied for further mapping of the *ycl-1* gene. A total of 43 of the 62 pairs of InDel primers designed based on the TO1000 reference genome exhibited polymorphisms according to the F₃ population. The genetic distances of the 16 InDel markers are shown in Table 1 (the sequences of these 16 markers are provided in Supplementary Table S2). The *ycl-1* gene was flanked by the InDel markers T1-36 (18,069,792 kb) and T1-58 (29,537,314 kb), with genetic distances of 0.42 cM and 0.42 cM, respectively. The interval distance between the two markers was 11.47 Mb based on the TO1000 reference genome. In the mapping region, spanning a physical distance of 11.47 Mb with a genetic difference of only 0.84 cM, the recombination rate was almost twenty times lower than the normal level for the cabbage genome (~600 kb/cM), suggesting that recombination suppression existed in this region.

Table 1. Genetic distances of the InDel primers to the *ycl-1* in the two mapping populations.

Primers	YL-1 × 01-20	Primers	YL-1 × 11-192
	Genetic Distance (cM)		Genetic Distance (cM)
T2-3	9.21	T2-3	13.3
T2-5	6.90	-	-
T1-1	6.28	-	-
T1-14	4.39	T1-14	6.5
T1-18	3.97	T1-18	4.4
T1-26	2.51	T1-26	2.3
T1-28	1.46	T1-28	1.5
T1-30	1.05	T1-30	1.3
T1-34	0.63	T1-34	0.3
T1-36	0.42	T1-36	0.00
T1-58	0.42	T1-58	0.7
T2-6	0.42	T2-6	1.04
T2-10	0.63	T2-10	1.04
T2-14	0.63	T2-14	1.04
T2-16	3.14	T2-16	2.61
T2-18	5.02	T2-18	6.02

Another BC₂P₁ population, constructed with YL-1 and 11-192, was used to further identify recombination suppression. The *ycl-1* gene was flanked by InDel markers T1-34 (17,301,717 kb) and T1-58 (29,537,314 kb), with genetic distances of 0.3 cM and 0.7 cM, respectively. This result further demonstrated the existence of a recombination-suppressed region in the *ycl-1* mapping interval.

In a previous study [7], we showed that the region between markers the BCYM585 (14,547,932 bp) and BCYM825 (24,060,605 bp) exhibits recombination suppression. In Figure 1, the sequence of BCYM585 was aligned to an unanchored scaffold (Scaffold00751), and the sequence of BCYM825 was aligned to a physical distance of 19,230,187 bp based on the TO1000 reference genome. Part II was aligned between 21,177,688 bp and 29,307,981 bp based on the TO1000 reference genome. These results showed that the recombination-suppressed region observed between the markers T1-36 (18,069,792 kb) and T1-58 (29,537,314 kb) in this study was consistent with the recombination-suppressed region between the markers BCYM585 and BCYM825 identified in our previous study [7].

4.2. BSR Analysis, DEGs between the Yellow-Green-Leaf and Normal-Leaf Pools

BSR was applied to obtain DEGs using the F₃ segregated population constructed with YL-1 and 01-20 and the BC₂ population constructed with YL-1 and 11-192. A total of 339,481,468 reads were generated from the four cDNA libraries. Among these reads, 82,143,852 were obtained from BC_normal, 91,405,984 from BC_yellow, 86,447,180 from F_normal, and 79,484,452 from F_yellow. The GC contents of the sequences of the four libraries were all approximately 47%, and all Q30% scores (reads with average quality scores >30) were >90%, indicating that the accuracy and quality of the sequencing data were sufficient for further analysis. The sequenced reads were aligned to the *B. oleracea*

genome reference (TO1000) (http://plants.ensembl.org/Brassica_oleracea/Info/Index) (accessed on 5 May 2017). An overview of the sequencing process is shown in Supplementary Table S3. The density distribution and boxplot of all the genes exhibited similar patterns among the four samples, indicating that no bias occurred in the construction of the cDNA libraries (Supplementary Figure S2).

The number of DEGs identified between the yellow-green-leaf and normal-leaf samples is shown in Table 2 (Supplementary Figure S3). In the yellow-green-leaf pools, there were approximately 20% fewer down-regulated genes than up-regulated genes. In total, 5730 and 4118 (4924 on average) DEGs were detected between the yellow-green-leaf and normal-leaf pools for the F₃ and BC₂ populations. As shown in the Venn diagram presented in Figure 2, 1884 common DEGs were shared between the DEGs identified in BC₂_normal vs. BC₂_yellow and the DEGs identified in F₃_normal vs. F₃_yellow, representing approximately half of the total number of DEGs in either population. Cross-comparison showed that only 1112 DEGs (Supplementary Table S4) were common between yellow-leaf and normal-leaf bulks. Thus, BSR analysis using four pools greatly reduced the number of DEGs from 4924 to 1112.

Table 2. Numbers of DEGs between the yellow-leaf and normal-leaf samples.

	No. of DEGs	No. of Up-Regulated DEGs	Percentage (%)	No. of Down-Regulated DEGs	Percentage (%)
BC ₂ _normal vs. BC ₂ _yellow	4118	2384	58	1734	42
BC ₂ _normal vs. F ₃ _yellow	8009	4894	60	3315	40
F ₃ _normal vs. F ₃ _yellow	5730	3226	56	2504	44
F ₃ _normal vs. BC ₂ _yellow	5405	2844	53	2561	47

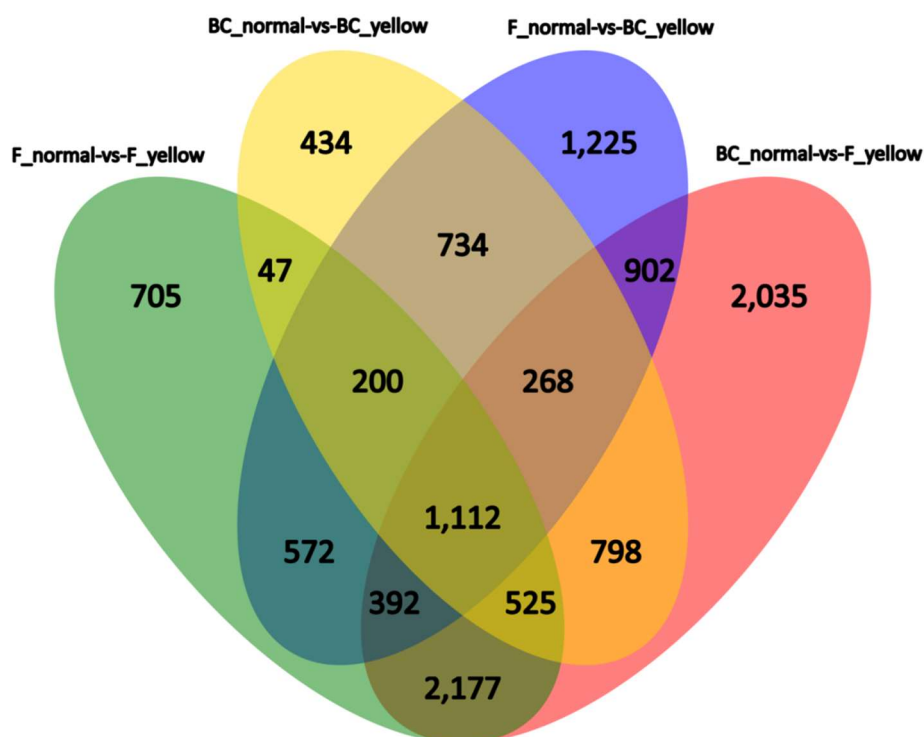


Figure 2. Venn diagram showing the numbers of overlapping and nonoverlapping DEGs (FDR < 0.01 and fold change > 2.0 or < −2.0) in the indicated segments from normal-leaf samples and yellow-leaf samples.

These 1112 DEGs were assigned into three Gene Ontology (GO) classes: biological process, cellular component, and molecular function. Thirty of the most significantly enriched of GO terms are shown in Figure 3, including “carbohydrate binding”, “sequence-specific DNA binding transcription factor activity”, “receptor activity”, “brassinosteroid sulfotransferase activity”, “unfolded protein binding” and “protein phosphatase inhibitor activity” under GO molecular functions and “endoplasmic reticulum lumen”, “plant-type cell wall”, “cytoplasm”, “vacuolar membrane”, “apoplast”, and “nucleus” under GO cellular components. Seventeen biological function or functional groups were enriched in the GO biological process category. In certain biological functions, genes play roles by interacting with each other, and KEGG pathway analysis helps provide an in-depth understanding of the biological functions of genes. A total of 1112 DEGs were annotated in the KEGG database, and 117 KEGG pathways were assigned. These 117 pathways were divided into three levels. Level one included “genetic information processing”, “metabolism”, “cellular processes”, “organismal systems”, and “environmental information processing.” The nineteen terms in level two are shown in Figure 4.

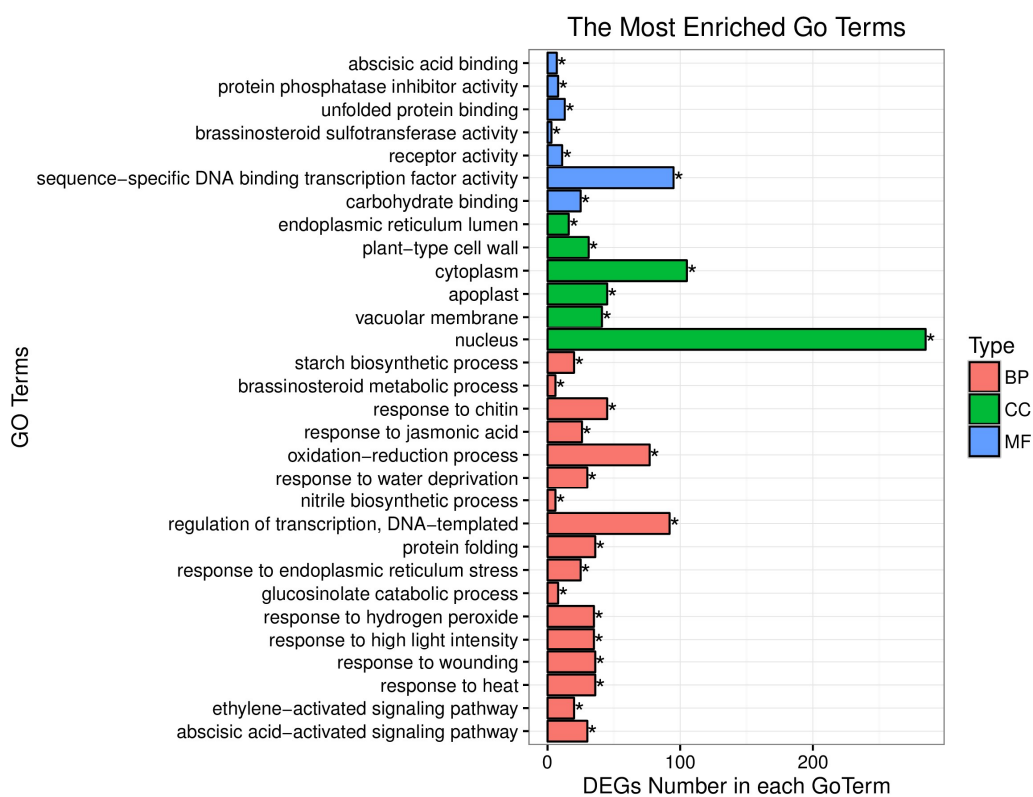


Figure 3. The thirty top GO assignments of 1112 DEGs. Blue: molecular function, green: cellular component, and red: biological process. The Y-axis represents the GO Term; the X-axis represents the number of DEGs for each GO Term. “*” indicates significant enrichment of the GO Term.



Figure 4. The top-20 enriched KEGG pathways of the 1112 DEGs. The Y-axis represents the pathway term; the X-axis represents the rich factor. The sizes of the points represent different DEG numbers, such that the bigger the point, the greater the DEG number. The colors represent different Q-values.

4.3. DEGs Involved in *B. oleracea* Chlorophyll Synthesis

The chlorophyll a, chlorophyll b, and total chlorophyll contents of the yellow-green-leaf mutant YL-1 were significantly lower than those of wild-type plants over the entire growth period [10]. Among the 1112 identified DEGs, 18 DEGs related to chlorophyll were clustered, which are shown in Supplementary Figure S4, including nine down- and nine up-regulated DEGs. These 18 DEGs were distributed among different chromosomes. Among the nine chromosomes, there were more DEGs on C01, C03, and C06 than on the other chromosomes (Supplementary Figure S5). In the 11.47 Mb recombination suppression region, two genes *Bo1g088040* (homologous gene *AT1G58290*, *HEMA1*) and *Bo1g098190* (homologous gene *AT1G61520*, *LCA3*) were related to chlorophyll according to the annotations, but there were not DEGs among these four pools by transcriptomics analysis and semi-quantitative PCR. Besides, no sequence variation was detected in the CDS region of these two genes of YL-1, compared with the sequences of 01-20, 11-192, and reference genome TO1000.

DEGs located in the *ysl-1* mapping interval with recombination suppression were selected for further analysis. In the BC_normal vs. BC_yellow comparison, 82 DEGs were found in the 11.47 Mb genomic region, with 45 DEGs being down-regulated and 37 being up-regulated. In the F_normal vs. F_yellow comparison, 105 DEGs were found in the 11.47 Mb genomic region, with 47 DEGs being down-regulated and 58 being up-regulated. Among these four pools, 43 common DEGs were present, with 20 DEGs being down-regulated and 23 being up-regulated (Supplementary Table S5). According to the annotations, five of these genes were related to chloroplasts (Table 3), including the down-regulated genes *Bo1g087310*, *Bo1g094360*, and *Bo1g098630* and the up-regulated genes *Bo1g059170* and *Bo1g098440*. These five genes were applied in qRT-PCR and RT-PCR analyses of the three parents (01-20, YL-1, 11-192). The relative normalized expression of these five genes is shown in Figure 5. The primers of qRT-PCR were supplied on Supplementary Table S6. Based on the relative normalized expression, it can be observed that the expression of *Bo1g059170*, *Bo1g087310*,

and *Bo1g094360* genes was consistent with the results of BSR, whereas the relative expression of the *Bo1g098440* and *Bo1g098630* genes differed from the results of BSR. We inferred that these two genes' transcription levels were irrelevant to the yellow-green-leaf trait. In the other three genes that related to chloroplasts, *Bo1g087310* (homologous gene *AT1G56340*, Calreticulins-1) plays important roles in calcium ion binding, plant growth, and plant height [33]. *Bo1g059170* (homologous gene *AT3G51420*) is involved in strictosidine synthase activity and plant defense [34], and *Bo1g094360* (homologous gene *AT3G08840*) functions in D-alanine-D-alanine ligase activity (Table 3) [35]. Hence, we inferred that these three candidate genes (*Bo1g094360*, *Bo1g087310*, and *Bo1g059170*) may be responsible for the development of the yellow-green-leaf mutant phenotype.

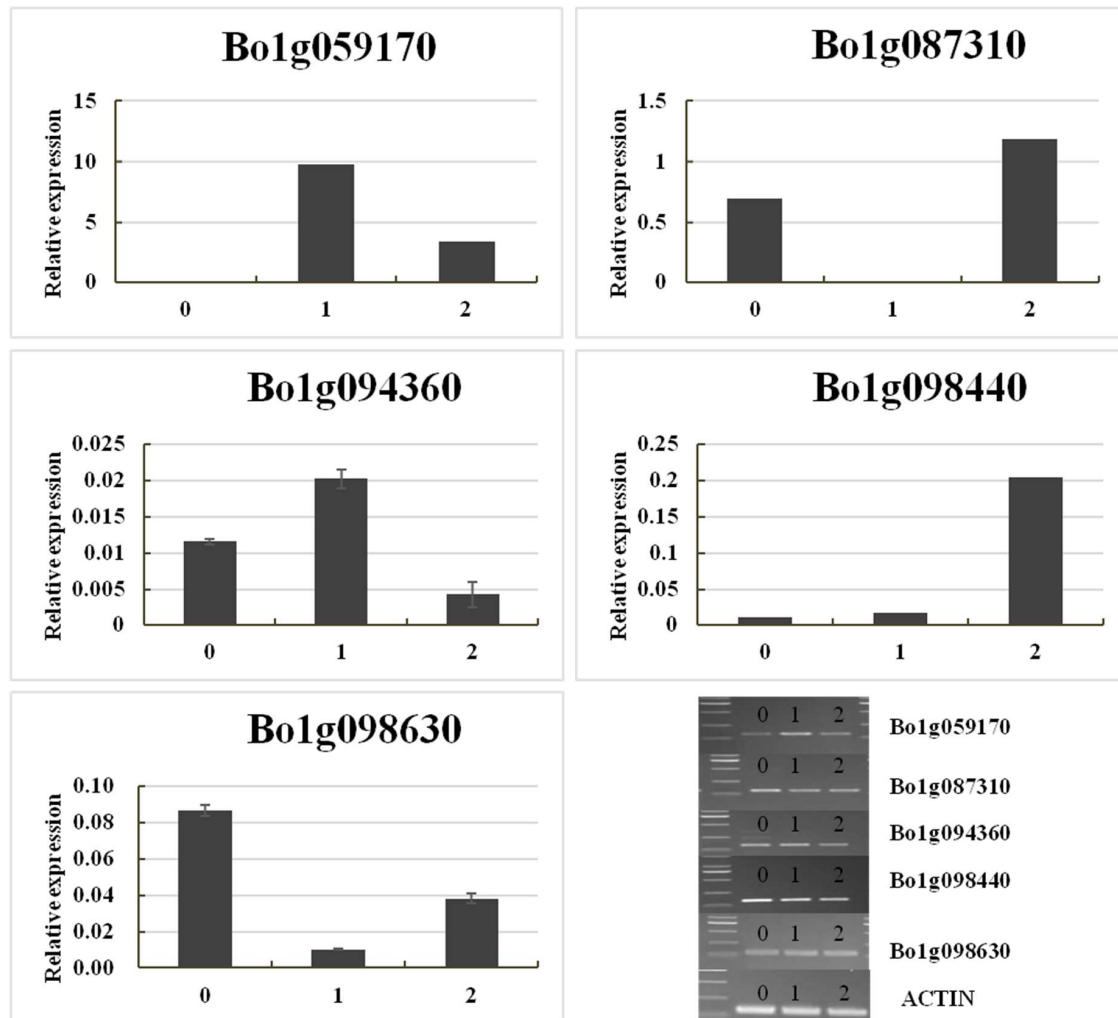


Figure 5. qRT-PCR and RT-PCR validation of transcripts of five DEGs associated with the yellow-green-leaf mutant. 0: the parent 01-20, 1: Mutant YL-1, 2: the parent 11-192.

Table 3. DEGs related to chloroplasts in the recombination-suppressed region.

Gene ID ^a	Physical Distance (TO1000)	F_Normal _b	F_Yellow _b	BC_Normal _b	BC_Yellow _b	Diff ^c	A.T. Annotation ^d
Bo1g087310	C1:25381300-25383803	1837.98	156.85	1920.64	287.42	Down	Calreticulins-1, response to oxidative stress, response to cadmiumion, response to salt stress, calciumion homeostasis;
Bo1g094360	C1:27829353-27834745	48.65	10.53	29.70	2.04	Down	D-alanine-D-alanine ligase activity
Bo1g098630	C1:29261755-29263303	4002.89	475.81	1119.81	125.36	Down	GPT2: glucose-6-phosphate/ phosphate translocator 2
Bo1g059170	C1:18110687-18112080	167.35	828.45	277.80	858.19	Up	SSL4: strictosidine synthase-like 4
Bo1g098440	C1:29037892-29038492	129.41	285.27	120.70	427.05	Up	Protein of unknown function, DUF538

^a Five *B. oleracea* DEGs related to chloroplasts (reference genome TO1000). ^b Expression levels in the four samples.

^c Differential regulation: up-regulation and down-regulation. ^d GO annotations for seven Bo to AT best-hit genes obtained from The Arabidopsis Information Resource (TAIR).

5. Discussion

5.1. Efficiency of BSR in DEG Detection

BSA (an efficient method for rapidly identifying markers linked to mutant phenotypes) combined with RNA-seq has been performed to map important agronomic traits at the transcription level in some species, such as catfish [23], onion [36] maize [37], Chinese cabbage [38], Chinese wheat cultivar [39], polyploid wheat [40], etc. Using BSR, Kim et al. [35] identified the candidate gene, AcPMS1, which is involved in DNA mismatch repair, for the fertility restoration of cytoplasmic male sterility in onions. Ramirez-Gonzalez et al. [24] mapped *Yr15* to a 0.77-cM interval in hexaploid wheat using a segregated F₂ population through BSR. In the present study, RNA-seq analysis of four bulks detected only 1112 common DEGs between the four pools (4924 on average), which can reduce the number of genes related to the phenotype. Therefore, BSR was further demonstrated to be an efficient method for analyzing the genes associated with the yellow-green-leaf mutant phenotype.

5.2. DEGs Analysis Associated with the Yellow-Green-Leaf in a Recombination-Suppressed Region via RNA-Seq

In recent years, the fine mapping of important agronomic traits in *Brassica* has developed rapidly [41–43]. Some yellow leaf color genes have been mapped in *Brassica* crops. A mutation responsible for chlorophyll deficiency in *Brassica juncea* was mapped between amplified fragment length polymorphism (AFLP) markers EA4TG4 and EA7MC1, with genetic distances of 33.6 and 21.5 cM, respectively [44]. In *B. napus*, Wang et al. [45] mapped the *CDE1* locus to a 0.9 cM interval of chromosome C08, and Zhu et al. [3] mapped a chlorophyll-deficient mutant between the markers BnY5 and CB10534, which are closely linked to the chlorophyll deficiency gene *BnaC.YGL*, with genetic distances of 3.0 and 3.2 cM on C06, respectively. Gene mapping for the above leaf color mutant was based on normal recombination in the segregated population. Recombination suppression was reported in many species, such as tomato [46], barley [47], petunia [48], *Populus* [49], hexaploid wheat [50], and buffelgrass [51]. In this study, we identified a large recombination suppression region spanning ~11 Mb on C01. However, recombination rate of *Brassica oleracea* C01 in previous studies seemed to be normal. The genetic map was constructed based on *Brassica oleracea* re-sequencing data; the C01 linkage groups spanned 97.59 cM, with an average distance of 1.15 cM between neighboring loci; and no recombination suppression was found [52]. Lv et al. (2016) [53] constructed a high-density genetic map while describing a comprehensive QTL analysis of key agronomic traits of cabbage. On C01, twelve markers existed between the markers Indel481 (17,365,179 bp) and

Indel14 (28,513,070 bp), which showed recombination was observed to be normal at the 17.3–28.5 Mb. In the present study, recombination suppression was observed at C01: 18,069,792–29,537,314 bp in the mapping of *ygl-1* gene using the population constructed from YL-1 and 01-20. Moreover, a recombination-suppressed region was identified in the same area while mapping *ygl-1* using another population constructed from YL-1 and 11-192. These two populations have one same parent YL-1. Therefore, we speculated that the suppression of recombination may be due to the YL-1 mutant.

In the recombination-suppressed region, it is difficult to identify candidate genes using fine mapping. Some research has revealed genes related to the phenotype by RNA-seq, such as *Fhb1* in wheat [54] and *BPH15* in rice [55]. In the *ygl-1* gene-mapping interval, a total of 10478 SNPs and Indels, with 455 genes, were identified in the recombination-suppressed region, including 78 genes related to chloroplasts. Comparison of the two bulk RNA-seq groups showed that only 43 genes were common DEGs, only five of which were related to chloroplasts. Furthermore, three of these five genes' expression by qRT-PCR were consistent with the results of BSR. Therefore, BSA combined with RNA-seq was able to greatly reduce the number of DEGs, demonstrating that this method is an effective alternative for identifying candidate genes in a recombination-suppressed region.

5.3. Assembly Error in the Reference Genome

Brassica oleracea reference genome sequencing was completed in 2014 [25,56]. However, the 02-12 reference genome assemblies have been woefully incomplete, and some assembly errors have been identified in recent studies. For example, Lee et al. [47] revised 27 v-blocks, 10 s-blocks, and several other blocks in the 02-12 reference genome assembly during the mapping of clubroot resistance QTLs through genotyping-by-sequencing. The purple leaf gene (*BoPr*) in the ornamental kale was mapped on an unanchored scaffold by Liu et al. (2017) [57]. In a previous study [7], we identified possible assembly errors in the 02-12 reference genome. According to the comparison of marker positions with the TO1000 reference, the physical locations of Part II and Part III in the 02-12 reference genome likely represent assembly errors (Figure 1). The makers' order of linkage map was consistent with the physical map order of TO1000 reference genome. All the results showed that the TO1000 reference genome is reliable. These results will contribute to the improvement of the cabbage genome.

6. Conclusions

In conclusion, we mapped the yellow-green-leaf gene *ygl-1* on a recombination-suppressed genomic region by two populations. Bulk segregant RNA-seq (BSR) was applied to identify differentially expressed genes using two segregate populations. BSR analysis with four pools greatly reduced the number of common DEGs from 4924 to 1112. Eighteen DEGs related to chlorophyll were clustered. In the *ygl-1* gene mapping region with suppressed recombination, 43 common DEGs were identified. Five of the genes were related to chloroplasts; the *Bo1g098440* and *Bo1g098630* genes were excluded by qRT-PCR. Hence, *Bo1g059170*, *Bo1g087310*, and *Bo1g094360* in the mapping interval may be tightly associated with the development of the yellow-green-leaf mutant phenotype. Further studies on these genes may reveal the molecular mechanism of yellow-green-leaf formation in *B. oleracea*.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/1422-0067/19/10/2936/s1>.

Author Contributions: X.L. and H.Y. developed the F₂ and BC₁ populations and wrote and revised the manuscript. H.Y., F.H., Z.L., and X.L. isolated the samples, performed the marker assays, and analyzed the marker data. F.H., Z.F., L.Y., M.Z., H.L., Y.L., Z.L., and Y.Z. conceived the study and critically reviewed the manuscript. All authors read and approved the final manuscript.

Funding: This work was financially supported by grants from the National Natural Science Foundation of China (31572141), the Major State Research Development Program (2016YFD0101702), the Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS), and the earmarked fund for the Modern Agro-Industry Technology Research System, China (nycytx-35-gw01).

Acknowledgments: This work was performed in the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Beijing 100081, People's Republic of China. The work reported here

was performed in the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Beijing 100081, China.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carol, P.; Stevenson, D. Mutations in the *Arabidopsis* gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. *Plant Cell*. **1999**, *11*, 57–68. [[CrossRef](#)] [[PubMed](#)]
2. Svensson, J.T.; Crosatti, C. Transcriptome analysis of cold acclimation in barley albina and xantha mutants. *Plant Physiol.* **2006**, *141*, 257–270. [[CrossRef](#)] [[PubMed](#)]
3. Zhu, L.; Zeng, X. Genetic characterisation and fine mapping of a chlorophyll-deficient mutant (*BnaC.ygl*) in *Brassica napus*. *Mol. Breed.* **2014**, *34*, 603–614. [[CrossRef](#)]
4. Chen, H.; Cheng, Z. A knockdown mutation of YELLOW—GREEN LEAF2, blocks chlorophyll biosynthesis in rice. *Plant Cell Rep.* **2013**, *32*, 1855–1867. [[CrossRef](#)] [[PubMed](#)]
5. Li, C.; Hu, Y. Mutation of FdC2 gene encoding a ferredoxin-like protein with C-terminal extension causes yellow-green-leaf phenotype in rice. *Plant Sci.* **2015**, *238*, 127–134. [[CrossRef](#)] [[PubMed](#)]
6. Ma, X.; Sun, X. Map-based cloning and characterization of the novel yellow-green-leaf gene *ys83* in rice (*Oryza sativa*). *Plant Physiol. Biochem.* **2016**, *111*, 1–9. [[CrossRef](#)] [[PubMed](#)]
7. Liu, X.; Yang, C. Genetics and fine mapping of a yellow-green-leaf gene (*ygl-1*) in cabbage (*Brassica oleracea* var. capitata L.). *Mol. Breed.* **2016**, *36*, 1–8. [[CrossRef](#)]
8. Whitaker, T.W. Genetic and Chlorophyll Studies of a Yellow-Green Mutant in Muskmelon. *Plant Physiol.* **1952**, *27*, 263–268. [[CrossRef](#)] [[PubMed](#)]
9. Zhong, X.M.; Sun, S.F. Research on photosynthetic physiology of a yellow-green mutant line in maize. *Photosynthetica* **2015**, *53*, 1–8. [[CrossRef](#)]
10. Yang, C.; Zhang, Y.Y. Photosynthetic Physiological Characteristics and Chloroplast Ultrastructure of Yellow Leaf Mutant YL-1 in Cabbage. *Acta Hort. Sin.* **2014**, *41*, 1133–1144.
11. Ladygin, V.G. Spectral features and structure of chloroplasts under an early block of chlorophyll synthesis. *Biophysics* **2006**, *51*, 635–644. [[CrossRef](#)]
12. Runge, S.; Cleve, B.V.; Lebedev, N.; Armstrong, G.; Apel, K. Isolation and classification of chlorophyll-deficient xantha mutants of *Arabidopsis thaliana*. *Planta* **1995**, *197*, 490–500. [[CrossRef](#)] [[PubMed](#)]
13. Chantret, N.; Sourdille, P. Location and mapping of the powdery mildew resistance gene MIRE and detection of a resistance QTL by bulked segregant analysis (BSA) with microsatellites in wheat. *Theor. Appl. Genet.* **2000**, *100*, 1217–1224. [[CrossRef](#)]
14. Futschik, A.; Schlötterer, C. The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics* **2010**, *186*, 207–218. [[CrossRef](#)] [[PubMed](#)]
15. Zeng, F.; Yi, B.; Tu, J.; Fu, T. Identification of AFLP and SCAR markers linked to the male fertility restorer gene of pol, CMS (*Brassica napus* L.). *Euphytica* **2009**, *165*, 363–369. [[CrossRef](#)]
16. Wang, Y.; Thomas, C.E.; Dean, R.A. Genetic mapping of a Fusarium wilt resistance gene (*Fom-2*) in melon (*Cucumis melo* L.). *Mol. Breed.* **2000**, *6*, 379–389. [[CrossRef](#)]
17. Subudhi, P.K.; Borkakati, R.P.; Virmani, S.S.; Huang, N. Molecular mapping of a thermosensitive genetic male sterility gene in rice using bulked segregant analysis. *Genome* **1997**, *40*, 188–194. [[CrossRef](#)] [[PubMed](#)]
18. Cheema, K.K.; Grewal, N.K. A novel bacterial blight resistance gene from *Oryza nivara* mapped to 38 kb region on chromosome 4L and transferred to *Oryza sativa* L. *Gen. Res.* **2008**, *90*, 397–407. [[CrossRef](#)] [[PubMed](#)]
19. Mackay, I.J.; Caligari, P.D. Efficiencies of F₂ and backcross generations for bulked segregant analysis using dominant markers. *Crop Sci.* **2000**, *40*, 626–630. [[CrossRef](#)]
20. Zhang, G.; Guo, G.W. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Res.* **2010**, *20*, 646–654. [[CrossRef](#)] [[PubMed](#)]
21. Song, H.K.; Hong, S.E. Deep RNA Sequencing Reveals Novel Cardiac Transcriptomic Signatures for Physiological and Pathological Hypertrophy. *PLoS ONE* **2012**, *7*, e35552. [[CrossRef](#)] [[PubMed](#)]
22. Jarvie, T.; Harkins, T. Transcriptome sequencing with the Genome Sequencer FLX system. *Nat. Methods* **2008**, *5*. [[CrossRef](#)]

23. Wang, R.; Sun, L. Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. *BMC Genomics* **2013**, *14*, 929–939. [[CrossRef](#)] [[PubMed](#)]
24. Ramirez-Gonzalez, R.H.; Segovia, V. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotech. J.* **2014**, *13*, 613–624. [[CrossRef](#)] [[PubMed](#)]
25. Parkin, I.A.; Koh, C. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biol.* **2014**, *15*, R77. [[CrossRef](#)] [[PubMed](#)]
26. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–359. [[CrossRef](#)] [[PubMed](#)]
27. Anders, S.; Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **2010**, *11*, R106. [[CrossRef](#)] [[PubMed](#)]
28. Krzywinski, M.; Schein, J. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645. [[CrossRef](#)] [[PubMed](#)]
29. Ashburner, M.; Ashburner, M. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)] [[PubMed](#)]
30. Kanehisa, M.; Goto, S. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)] [[PubMed](#)]
31. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
32. Guo, J.; Zhang, Y. Transcriptome sequencing and de novo analysis of a recessive genic male sterile line in cabbage (*Brassica oleracea* L. var. capitata). *Mol. Breed.* **2016**, *36*, 117–119. [[CrossRef](#)]
33. Piippo, M.; Allahverdiyeva, Y. Chloroplast-mediated regulation of nuclear genes in *Arabidopsis thaliana* in the absence of light stress. *Physiol. Gen.* **2006**, *25*, 142–152. [[CrossRef](#)] [[PubMed](#)]
34. Sohani, M.M.; Schenk, P.M.; Schultz, C.J.; Schmidt, O. Phylogenetic and transcriptional analysis of a strictosidine synthase-like gene family in *Arabidopsis thaliana*, reveals involvement in plant defence responses. *Plant Biol.* **2009**, *11*, 105–117. [[CrossRef](#)] [[PubMed](#)]
35. Jyothi, T.; Duan, H.; Liu, L.; Schuler, M.A. Bicistronic and fused monocistronic transcripts are derived from adjacent loci in the *Arabidopsis* genome. *RNA* **2005**, *11*, 128–138.
36. Kim, S.; Kim, C.W.; Park, M.; Choi, D. Identification of candidate genes associated with fertility restoration of cytoplasmic male-sterility in onion (*Allium cepa* L.) using a combination of bulked segregant analysis and RNA-seq. *Theor. Appl. Genet.* **2015**, *128*, 2289–2299. [[CrossRef](#)] [[PubMed](#)]
37. Liu, C.; Zhou, Q.; Dong, L.; Wang, H.; Liu, F.; Weng, J.; Li, X.; Xie, C. Genetic architecture of the maize kernel row number revealed by combining QTL mapping using a high-density genetic map and bulked segregant RNA sequencing. *BMC Genomics* **2016**, *17*, 915. [[CrossRef](#)] [[PubMed](#)]
38. Huang, Z.; Peng, G.; Liu, X.; Deora, A.; Falk, K.C.; Gossen, B.D.; McDonald, M.R.; Yu, F. Fine Mapping of a Clubroot Resistance Gene in Chinese Cabbage Using SNP Markers Identified from Bulk Segregant RNA Sequencing. *Front. Plant Sci.* **2017**, *8*, 1448–1459. [[CrossRef](#)] [[PubMed](#)]
39. Wang, Y.; Xie, J.; Zhang, H. 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.* **2017**, *130*, 2191–2201. [[CrossRef](#)] [[PubMed](#)]
40. Trick, M.; Adamski, N.M.; Mugford, S.G.; Jiang, C.C.; Febrer, M.; Uauy, C. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biol.* **2012**, *12*, 14–19. [[CrossRef](#)] [[PubMed](#)]
41. Lei, S.; Yao, X. Towards map-based cloning: Fine mapping of a recessive genic male-sterile gene (*BnMs2*) in *Brassica napus* L. and syntenic region identification based on the *Arabidopsis thaliana* genome sequences. *Theor. Appl. Genet.* **2007**, *115*, 643–651. [[CrossRef](#)] [[PubMed](#)]
42. Shimizu, M.; Pu, Z.J. Map-based cloning of a candidate gene conferring Fusarium yellows resistance in *Brassica oleracea*. *Theor. Appl. Genet.* **2015**, *128*, 119–130. [[CrossRef](#)] [[PubMed](#)]
43. Liang, J.; Ma, Y. Map-based cloning of the dominant genic male sterile *Ms-cd1* gene in cabbage (*Brassica oleracea*). *Theor. Appl. Genet.* **2016**, *5*, 1–9. [[CrossRef](#)] [[PubMed](#)]
44. Tian, Y.; Huang, Q. Inheritance of chlorophyll-deficient mutant *L638-y* in *Brassica juncea* L. and molecular markers for chlorophyll deficient gene *gr1*. *J. Northwest A F Univ.* **2012**, *12*, 17–19.

45. Wang, Y.; He, Y. Fine mapping of a dominant gene conferring chlorophyll-deficiency in *Brassica napus*. *Sci. Rep.* **2016**, *6*, 314–319. [[CrossRef](#)] [[PubMed](#)]
46. Sherman, J.D.; Stack, S.M. Two-dimensional spreads of synaptonemal complexes from Solanaceous plants. VI. Highresolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics* **1995**, *141*, 683–708. [[PubMed](#)]
47. Wei, F.; Gobelman-Werner, K. The Mla (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* **1999**, *153*, 1929–1948. [[PubMed](#)]
48. Ten, H.R.; Robbins, T.P. Localization of T-DNA Insertions in Petunia by Fluorescence in Situ Hybridization: Physical Evidence for Suppression of Recombination. *Plant Cell* **1996**, *8*, 823–830.
49. Stirling, B.; Newcombe, G. Suppressed recombination around the MXC3 locus, a major gene for resistance to poplar leaf rust. *Theor. Appl. Genet.* **2001**, *103*, 1129–1137. [[CrossRef](#)]
50. Neu, C.; Stein, N.; Keller, B. Genetic mapping of the Lr20-Pm1 resistance locus reveals suppressed recombination on chromosome arm 7AL in hexaploid wheat. *Genome* **2002**, *45*, 737–744. [[CrossRef](#)] [[PubMed](#)]
51. Jessup, R.W.; Burson, B.L. Disomic Inheritance, Suppressed Recombination, and Allelic Interactions Govern Apospory in Buffelgrass as Revealed by Genome Mapping. *Crop Sci.* **2002**, *42*, 1688–1694. [[CrossRef](#)]
52. Lee, J.; Izzah, N.K. Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Res.* **2015**, *14*, S113. [[CrossRef](#)] [[PubMed](#)]
53. Lv, H.; Wang, Q. Whole-Genome Mapping Reveals Novel QTL Clusters Associated with Main Agronomic Traits of Cabbage (*Brassica oleracea* var. *capitata* L.). *Front. Plant Sci.* **2016**, *7*, 989–999. [[CrossRef](#)] [[PubMed](#)]
54. Schweiger, W.; Schweiger, W.; Steiner, B.; Vautrin, S.; Nussbaumer, T.; Siegwart, G.; Zamini, M.; Jungreithmeier, F.; Gratl, V.; Lemmens, M.; et al. Suppressed recombination and unique candidate genes in the divergent haplotype encoding *Fhb1*, a major Fusarium head blight resistance locus in wheat. *Theor. Appl. Genet.* **2016**, *129*, 1607–1623. [[CrossRef](#)] [[PubMed](#)]
55. Lv, W.; Du, B. BAC and RNA sequencing reveal the brown planthopper resistance gene *BPH15*, in a recombination cold spot that mediates a unique defense mechanism. *BMC Genomics* **2014**, *15*, 674–679. [[CrossRef](#)] [[PubMed](#)]
56. Liu, S.; Liu, Y.; Yang, X.; Tong, C.; Edwards, D.; Parkin, I.A. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* **2014**. [[CrossRef](#)] [[PubMed](#)]
57. Liu, X.P.; Gao, B.Z.; Han, F.Q.; Fang, Z.Y.; Yang, L.M.; Zhuang, M.; Lv, H.H.; Liu, Y.M.; Li, Z.S.; Cai, C.C.; et al. Genetics and fine mapping of a purple leaf gene, BoPr, in ornamental kale (*Brassica oleracea* L. var. *acephala*). *BMC Genomics* **2017**, *18*, 230–239. [[CrossRef](#)] [[PubMed](#)]

