

# Integrated Analysis of Transcriptome and Metabolome Reveals New Insights into the Formation of Purple Leaf Veins and Leaf Edge Cracks in *Brassica juncea*

Kaijing Zhang <sup>1</sup>, Dekun Yang <sup>1</sup>, Yuchao Hu <sup>1</sup>, Martin Kagiki Njogu <sup>2</sup>, Jingjing Qian <sup>1</sup>, Li Jia <sup>3</sup>, Congsheng Yan <sup>3</sup>, Ziang Li <sup>4</sup>, Xing Wang <sup>4,\*</sup> and Liping Wang <sup>4,\*</sup>  
[mkagiki@chuka.ac.ke](mailto:mkagiki@chuka.ac.ke) <sup>2</sup>

<sup>1</sup> College of Agriculture, Anhui Science and Technology University, Fengyang 233100, China

<sup>2</sup> Department of Plant Science, Chuka University, Chuka P.O. Box 109-60400, Kenya

<sup>3</sup> Key Laboratory of Genetic Improvement and Ecophysiology of Horticultural Crop, Institute of Horticulture, Anhui Academy of Agricultural Sciences, Hefei 230001, China

<sup>4</sup> School of Landscape and Ecological Engineering, Hebei University of Engineering, Handan 056038, China

\* Correspondence: wangxing@hebeu.edu.cn (X.W.); wlp29@163.com (L.W.)

**Abstract:** Purple leaf veins and leaf edge cracks comprise the typical leaf phenotype of *Brassica juncea*; however, the molecular mechanisms and metabolic pathways of the formation of purple leaf veins and leaf edge cracks remain unclear. In this study, transcriptome and metabolome analyses were conducted to explore the regulation pathway of purple leaf vein and leaf edge crack formation based on four mustard samples that showed different leaf colors and degrees of cracking. The results showed genes with higher expression in purple leaf veins were mainly enriched in the flavonoid biosynthesis pathway. Integrating related genes and metabolites showed that the highly expressed genes of *ANS* (*BjuA004031*, *BjuB014115*, *BjuB044852*, and *BjuO009605*) and the excessive accumulation of dihydrokaempferol and dihydroquercetin contributed to the purple leaf veins by activating the synthetic pathways of pelargonidin-based anthocyanins and delphinidin-based anthocyanins. Meanwhile, “alpha-farnesene synthase activity” and “glucan endo-1, 3-beta-D-glucosidase activity” related to the adversity were mainly enriched in the serrated and lobed leaves, indicating that the environmental pressure was the dominant factor controlling the change in leaf shape. Overall, these results provided new insights into the regulation pathways for formation of purple leaf veins and leaf edge cracks, which could better accelerate the theoretical research on purple leaf vein color and leaf edge cracks in mustard.

**Keywords:** *Brassica juncea*; purple leaf vein; leaf edge cracks; multiomics; regulation pathway



**Citation:** Zhang, K.; Yang, D.; Hu, Y.; Njogu, M.K.; Qian, J.; Jia, L.; Yan, C.; Li, Z.; Wang, X.; Wang, L. Integrated Analysis of Transcriptome and Metabolome Reveals New Insights into the Formation of Purple Leaf Veins and Leaf Edge Cracks in *Brassica juncea*. *Plants* **2022**, *11*, 2229. <https://doi.org/10.3390/plants11172229>

Academic Editor: Alex Troitsky

Received: 18 July 2022

Accepted: 24 August 2022

Published: 28 August 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Mustard (*Brassica juncea* L.) is an annual herb that belongs to *Brassica* in the *Cruciferae* family that is cultivated and consumed mainly due to its edible stalks and leaves. In addition, it is an important oilseed and condiment crop [1]. Due to genetic variation, the cultivated species exhibit extensive morphological polymorphisms including the phenotypes of leaves, stems, roots, or seeds. These polymorphisms are the vital material basis for plant phenotypic research and directional breeding [2]. Among these polymorphisms, due to the important economic value, the study of leaf color and leaf shape has become the focus of researchers.

Leaf color is one of the easily recognizable morphological markers that influence consumer preference and market competitiveness. For leafy mustard cultivars, purple color of the leaf blade or leaf vein is a typical characteristic of great use in the study of germplasm resources. For example, purple leaves can be used as visible markers of plant transformation [3] and to rapidly identify seed purity in hybrid seed production at the seedling stage. Previous studies found that the formation of purple leaves is mainly

related to the accumulation of anthocyanins [4]. Anthocyanins are natural food colorants that play important roles in the life activities of animals or plants, such as in antioxidant activities [5] and in resistance to biotic and abiotic stresses [6]. Various genes involved in the biosynthesis pathway of anthocyanin have been widely explored in model plants including *Arabidopsis thaliana* [7,8], *Zea mays* [9], *Oryza sativa* [10], and so on. The ternary complex of MYB-bHLH-WD40 has been shown to play an important role in anthocyanin accumulation. In *Arabidopsis*, overexpression of the R2R3-MYB transcription factors (*PAP1*, *PAP2*, *MYB113*, and *MYB114*) increased anthocyanin accumulation [7], while in cauliflower, the upregulated expression of *BoMYB2* and *BoTT8* genes led to anthocyanin accumulation, resulting in purple leaves [11]. In contrast, the silent expression of the *BoMYBL2-1* gene resulted in the establishment of purple *Brassica oleracea* leaves [12]. In *Brassica* species, several candidate genes or loci related to purple color have been identified, such as *BrEGL3.1* and *BrEGL3.2* in *zicaitai* (*B. rapa* L. ssp. *chinensis* var. *purpurea*) [13], *BoDRF* gene in Chinese kale (*Brassica oleracea* var. *alboglabra*) [14], and so on. However, the regulatory mechanism of anthocyanin accumulation that results in purple leaf in *Brassica* vegetables remains unknown.

The diversity of leaf shape is one of the reasons for the diversity among species [15], and it affects the physiological functions of plants. Previous studies found that serrated or deeply cracked leaves increased the distance between leaves, which could increase heat dissipation and enhance disease defense [16]. The leaves of *Capsella rubella* have plasticity to temperature; the temperature is negatively related to the depth of leaf edge cracks, which means that the higher the temperature, the shallower the depth of cracks [17]. The difference in leaf shape in tomato also affects the flavor and quality of the fruits [18]. Leaf shape is not only regulated by genes [19–21], but is also influenced by the environment and hormones [22,23]. In *Arabidopsis thaliana*, the *CUP-SHAPED COTYLEDON1* (*CUC1*), *CUC2*, and *CUC3* genes regulate the formation of primitive shoot apex meristems and constrain the borders of cotyledons [24]. *CUC1* and *CUC2* are the target genes of small RNA *miR164*. The balance between *miR164* and *CUC2* determine the degree of leaf margin cracking [25]. The recessive mutation of *CUC2* leads to a decrease in the degree of leaf margin cracking, while the overexpression of *CUC2* leads to an increase in leaf margin cracking by destroying *miR164* [26]. Saddic found that the *LATE MERISTEM IDENTITY 1* (*LMI1*) gene plays an important role in the formation of a single serrated leaf [27]. The *REDUCED COMPLEXITY* (*RCO*) gene could inhibit local leaf growth and change leaf shape via negative regulation with multiple cytokinins (CK) [28]. Moreover, Gupta et al. found that the leaf morphology of different sweet potato varieties varied greatly in different geographical locations, which indicated that the environment also affected the leaf shape [29]. In *Brassica*, some candidate genes related to leaf edge cracks were reported, such as *BnaA10g26320D* and *BnaA10g26330D* in *Brassica napus* [30,31] and *BrLMI1* in *Brassica napus* [32]. However, there are few studies that reported the genes associated with leaf edge cracks in *Brassica juncea*, especially at the transcriptome level. Therefore, the molecular mechanism of the formation of leaf edge cracks in *Brassica juncea* requires further study.

In this study, different types of mustard leaf samples (purple vein, green vein, different degrees of leaf marginal dehiscence) were collected. Through newly developed multi-omics technology, including RNA sequencing (RNA-Seq) and metabolic assays, the complex biosynthesis pathways that result in purple leaf veins and different degrees of leaf edge cracks were dissected in *Brassica juncea*. The results provided novel knowledge in the understanding of the formation of purple leaf veins and different degrees of leaf edge cracking in *Brassica juncea*, which could be a solid theoretical basis for the breeding application.

## 2. Results

### 2.1. Comparison of Morphological Phenotypes among Mustards

Significant differences were observed in the morphological phenotypes among 44, 45, 65, and 66, especially in the leaf color and type. As shown in Figure 1, the leaves of 44 and 66 exhibited purple veins compared with those of 45 and 65, which had green veins. Furthermore, when comparing the entire leaves of 45, 44, 65, and 66, each showed

different degrees of leaf edge cracking: 44 exhibited a serrated leaf, while 65 and 66 showed a lobed leaf.



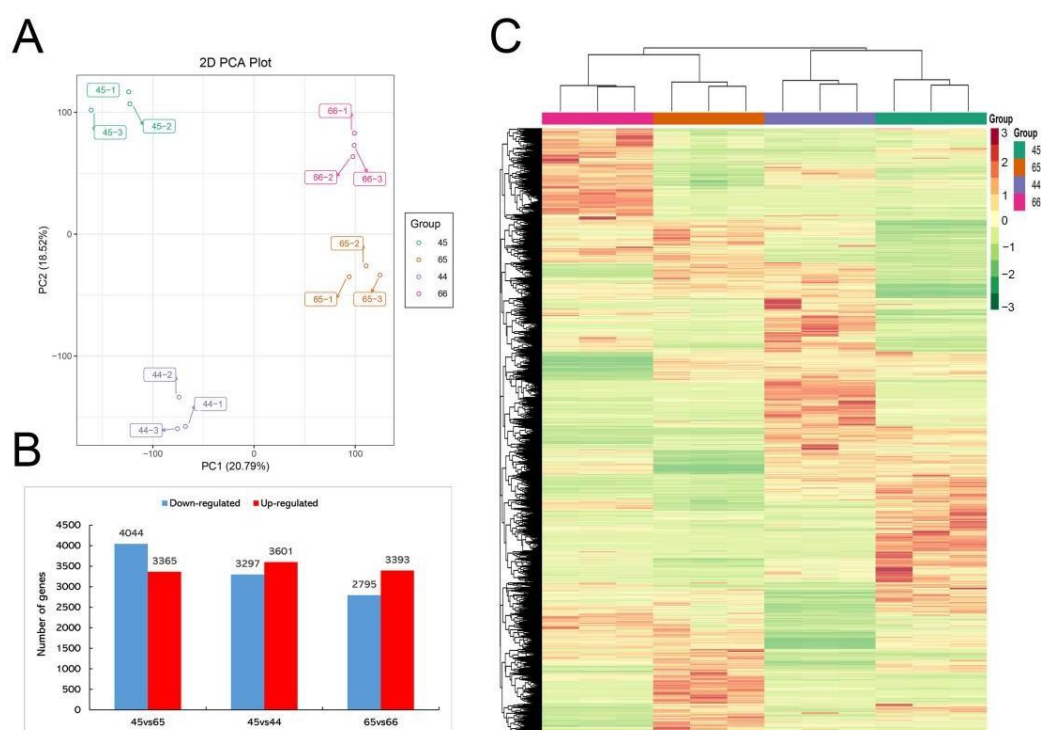
**Figure 1.** The morphological phenotypes of the four mustard cultivars. (A) The phenotype of each single leaf. (B) Overall plant phenotypes: 45 has green veins and an entire leaf; 44 shows purple veins and a serrated leaf; 65 shows green veins and a lobed leaf; 66 exhibits purple veins and a lobed leaf.

### 2.2. Identification of Differential Expression of Genes with RNA-Seq

A total of 48.69–58.05 million raw reads were obtained via Illumina sequencing of the mustard leaf samples. After removing reads containing adapters or poly-N and low-quality reads, clean reads were obtained with Q30 percentages of 90.47–92.77% and GC percentages of 47.37–47.86% (Table 1). The transcriptome data of different mustard leaf samples were separated using PCA; PC1 and PC2 were 20.79% and 18.52%, respectively (Figure 2A). One group of samples was clearly distinguished from the other three groups of samples and three repeated samples in each group were gathered together, indicating that the samples had good reproducibility and that there were significant differences in gene expression levels among different groups. There were 7409, 6898, and 6188 DEGs in the three comparison groups of 45 vs. 65, 45 vs. 44, and 65 vs. 66, respectively (Figure 2B). To clearly exhibit the differences among different groups, the FPKM values of DEGs among different groups were extracted for a hierarchical cluster analysis (Figure 2C). Those DEGs among different groups may contribute to the different phenotypes of mustard to a great extent.

**Table 1.** Summary of RNA-Seq data from different types of mustard.

Samples	Raw Reads	Clean Reads	Q30 (%)	GC Content (%)	Unique Mapped
44-1	52,764,138	51,838,932	90.47	47.65	41,712,627 (80.47%)
44-2	57,898,070	56,684,510	92.36	47.71	46,074,374 (81.28%)
44-3	48,686,530	47,839,794	92.77	47.59	38,895,911 (81.30%)
45-1	59,411,784	58,048,268	92.23	47.75	47,126,857 (81.19%)
45-2	59,141,800	57,661,630	92.14	47.56	45,972,841 (79.73%)
45-3	52,856,618	51,892,220	92.01	47.64	41,742,258 (80.44%)
65-1	59,045,222	57,784,616	92.21	47.37	46,641,623 (80.72%)
65-2	52,647,580	51,702,172	91.46	47.61	41,756,075 (80.76%)
65-3	57,451,232	56,518,720	92.24	47.45	45,650,258 (80.77%)
66-1	54,027,590	53,162,134	92.64	47.86	42,963,193 (80.82%)
66-2	56,428,334	55,492,792	92.28	47.8	44,569,467 (80.32%)
66-3	51,296,644	50,462,792	92.09	47.62	40,610,737 (80.48%)

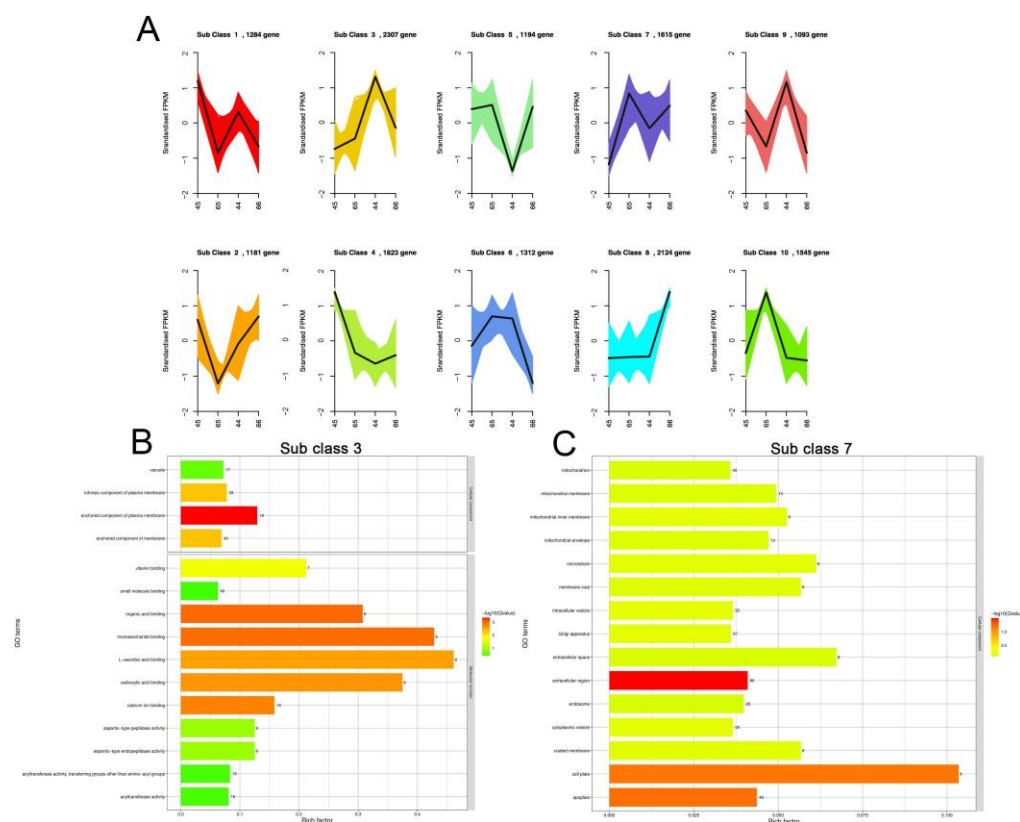


**Figure 2.** Multivariate statistical analysis of genes from four mustards. **(A)** PCA score plot of genes of the four mustard samples; the  $x$ -axis represents the first principal component and the  $y$ -axis represents the second principal component. **(B)** The numbers of DEGs in different sample combinations. **(C)** The hierarchical clustering analysis and heatmaps for DEGs.

### 2.3. K-Means Cluster and Enrichment Analyses of DEGs Associated with the Formation of Different Phenotypes

Genes with similar expression patterns may play a similar function. To understand gene-expression patterns among different mustard leaf samples, the K-means clustering algorithm was used. The results of the K-means cluster analysis showed that the DEGs were divided into 10 subclasses. Combined with the phenotypes of different samples, subclasses 3 and 7 were selected for further study. The genes in subclasses 3 and 7 totaled 2307 and 1615, respectively (Table S1). Compared with the green vein samples (45 and 65), the genes in subclass 3 in the purple leaf vein samples (44 and 66) showed higher expression trends (Figure 3A), indicating that these genes may play a regulatory role in the formation of purple leaf veins. Similarly, in subclass 7, compared with the entire leaf sample (45), the genes in the serrated leaf sample (44) and lobed leaf samples (65 and 66)

showed higher expression trends (Figure 3A), indicating that the high expression levels of these genes resulted in different degrees of leaf edge cracking.

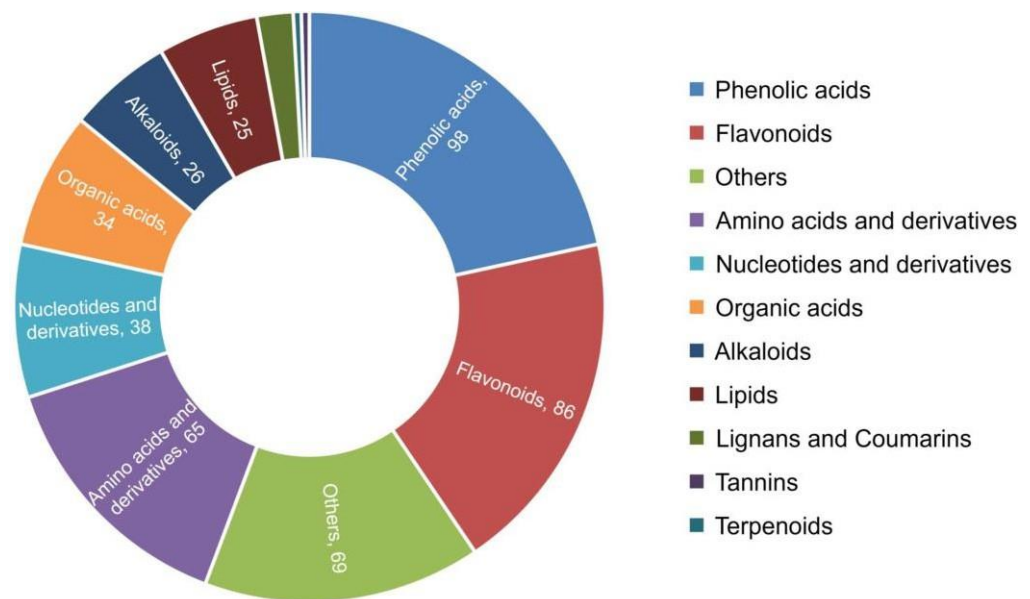


**Figure 3.** K-means cluster analysis and functional GO pathway classification of the DEGs. (A) Dynamics of DEGs among four samples. The bold black line represents the average pattern of all DEGs in each class, the different colored lines represent different classes, and each line represents the standardized FPK over three biological replicates. (B) GO-enriched analysis of DEGs in subclass 3. (C) GO-enriched analysis of DEGs in subclass 7.

To identify the functions of DEGs in subclasses 3 and 7, a Gene Ontology (GO)-enriched analysis was performed (Figure 3B,C). Based on the GO-enriched analysis, in subclass 3, the most significantly enriched molecular functions (MFs) were “L-ascorbic acid binding”, “calcium ion binding”, “acyltransferase activity”, and “small molecule binding”. The most significantly enriched cellular components (CCs) were “endoplasmic reticulum”, “intrinsic component of membrane”, and “cell periphery”. There was no biological process (BP) enriched. In subclass 7, only the CCs were significantly enriched, including “isomerase activity”, “extracellular region”, “cell plate”, “apoplast”, “extracellular space”, and so on. All of the GO enrichment results are shown in Tables S2 and S3. To further explore the biological pathways of formation of different mustard leaf samples, the DEGs in subclasses 3 and 7 were annotated based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The results showed that in subclass 3, the flavonoid biosynthesis, flavone and flavonol biosynthesis, and phenylalanine metabolism were the most significantly changed pathways (Table S4), which are the typical metabolic pathways that form plant color. In subclass 7, the pathways of carbohydrate metabolism, oxidative phosphorylation, energy metabolism, and citrate cycle (TCA cycle) were significantly enriched. In particular, the pathways enriched in environmental adaptation may also contribute to the change in leaf shape between 45 and the other three mustard leaf samples (44, 65, 66) (Table S5).

To further explore the key genes controlling leaf color, we compared and analyzed the DEGs among the four samples. We focused on the genes highly expressed in the 45 vs. 44

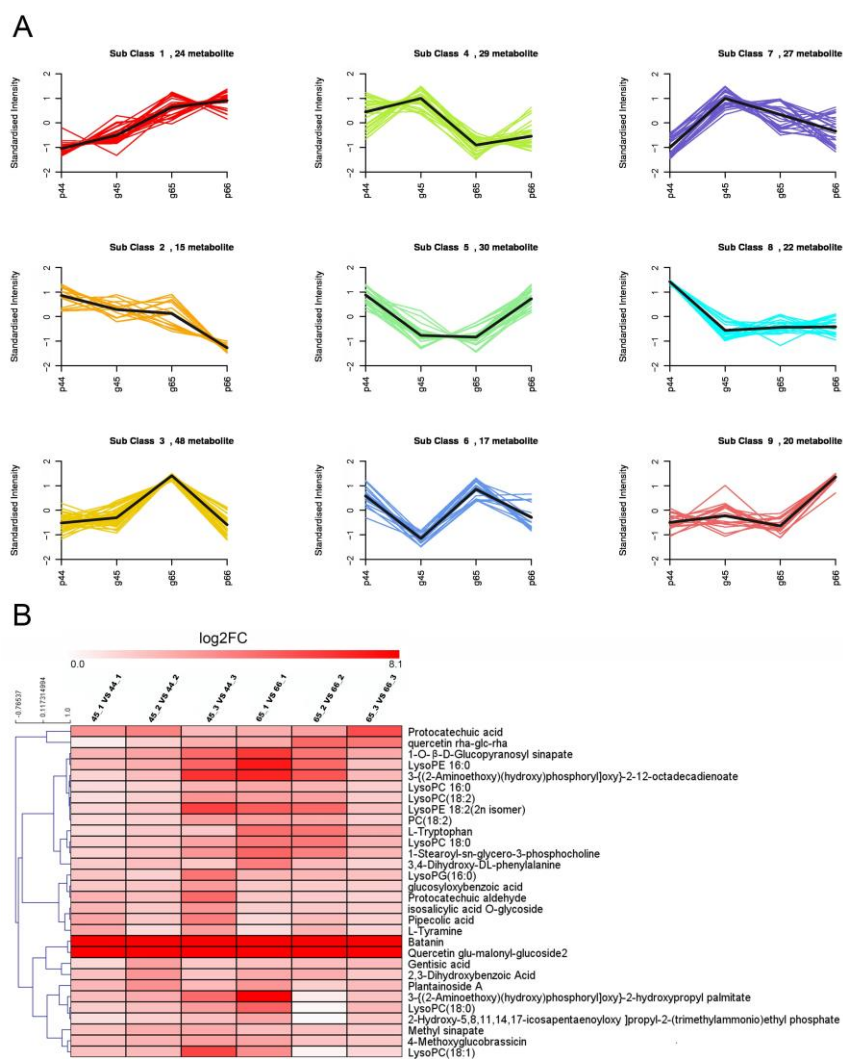




**Figure 5.** Pie chart of the classification of 454 metabolites.

Based on the OPLS-DA results and using the criteria  $VIP > 1$  and absolute  $\log_2FC > 1$ , 103 metabolites (49 downregulated and 54 upregulated) exhibited an altered abundance between the 45 and 44 samples (Figure S2A), 116 metabolites (66 downregulated and 50 upregulated) were significantly different between the 65 and 66 samples (Figure S2B), and 106 metabolites (47 downregulated and 59 upregulated) were significantly different between the 45 and 65 samples (Figure S2C). To further clarify the reasons for the formation of different leaf phenotypes, all the DAMs were analyzed using a K-means clustering algorithm. The results showed that a total of nine subclasses were obtained, among which the content of candidate metabolites in subclass 5 exhibited a higher accumulation in the 44 and 66 samples (purple leaf veins) compared with 45 and 65 (green veins) (Figure 6A). The metabolites in subclass 5 could be considered as potential metabolites of purple leaf veins. In subclass 5, the metabolites' accumulation of 30 DAMs was displayed in the form of a heat map. The more metabolites that accumulated, the more closely related they were to the formation of purple leaf veins. As Figure 6B shows, quercetin glu-malonyl-glucoside2 and betanin were the most accumulated metabolites in the 44 and 66 samples compared with the 45 and 65 samples, respectively. Betanin is classified as an anthocyanidins and quercetin glu-malonyl-glucoside2 is an important flavonoid, and both of them are major pigments in the formation of plant leaf color.

To explore the vital candidate metabolites causing purple leaf veins and leaf edge cracks in mustard, Venn diagrams were drawn based on the DAMs of 45 vs. 44 and 65 vs. 66 as well as 45 vs. 44 and 45 vs. 65. The results indicated that 16 metabolites were highly accumulated in purple leaf vein samples (44 and 66) compared with green vein samples (45 and 65), and 5 metabolites were downregulated in purple leaf vein samples (Figure S3A; Table 2). This was consistent with the above results showing that the metabolites of quercetin glu-malonyl-glucoside2 and betanin were highly accumulated in purple leaf veins, indicating once again that these two metabolites were important for the change in the leaf vein color. In addition, one metabolite belonging to anthocyanins was downregulated in the two purple leaf vein samples. Regarding the leaf shape diversity, 15 metabolites were upregulated in the notched leaf samples (44 and 65); among these, the highest content of metabolites was alkaloids, followed by amino acids and derivatives. In addition, 14 metabolites were downregulated, and the highest content of metabolites was phenolic acid (Figure S3B; Table 3).



**Figure 6.** K-means analysis of differential metabolites. **(A)** Dynamics of metabolites for the four mustard samples. The bold black line represents the average pattern of all candidate metabolites in each class, the different colored lines represent different classes, and each line represents the standardized value of the metabolite over three biological replicates. **(B)** Hierarchical clustering heat map of 30 differential accumulation metabolites in subclass 5.

**Table 2.** The consistently significantly regulated metabolites in the 45 vs. 44 and 65 vs. 66 groups.

Metabolome ID	Compounds	Class I	Class II	Log2FC (45 vs. 44)	Log2FC (65 vs. 66)
CWJP002833	Kaempferol-di-O-malonyl-glucoside-O-malonyl-diglucoside	Flavonoids	Flavonols	13.84	1.64
Hmgp006596	Cyclobassinin	Alkaloids	Plumerane	1.128	1.64
Hmyn007168	LysoPG(16:0)	Lipids	Glycerol ester	1.32	1.09
Lmmp002995	Quercetin glu-malonyl-glucoside2	Flavonoids	Flavonols	12.68	12.25
Lmyn001269	Kaempferol 3-O-β-D-sophoroside	Flavonoids	Flavonols	1.03	12.34

**Table 2.** *Cont.*

Metabolome ID	Compounds	Class I	Class II	Log2FC (45 vs. 44)	Log2FC (65 vs. 66)
mws0126	1-Stearoyl-sn-glycero-3-phosphocholine	Lipids	LPC	1.156	1.72
mws0183	Protocatechuic acid	Flavonoids	Flavanols	1.63	1.70
mws0639	2,3-Dihydroxybenzoic acid	Organic acids	Organic acids	1.18	1.10
mws1138	Betanin	Alkaloids	Alkaloids	14.19	13.17
pmb0876	LysoPE 16:0	Lipids	LPE	1.68	1.74
pmb2936	Disinapoyl hexoside	Phenolic acids	Phenolic acids	1.76	12.55
pmb2940	1-O-β-D-Glucopyranosyl sinapate	Phenolic acids	Phenolic acids	1.66	1.97
pmd0136	LysoPC 18:0	Lipids	LPC	1.05	1.73
pmn001409	Plantainoside A	Phenolic acids	Phenolic acids	1.32	1.06
pmp001277	3-{{(2-Aminoethoxy)(hydroxy) phosphoryl}oxy}-2-hydroxypropyl palmitate	Alkaloids	Alkaloids	1.69	1.68
pmp001286	LysoPC(18:0)	Lipids	LPC	1.15	1.67
Hmpp002612	Luteolin-7-O-β-D-gentiobioside	Flavonoids	Flavonoid	−2.19	−1.31
Lmbp002592	Kaempferol-3,7-di-O-β-D-glucopyranoside	Flavonoids	Flavonols	−2.34	−1.82
mws0170	Cyanidin chloride	Flavonoids	Anthocyanins	−2.26	−2.32
pme0256	Xanthine	Nucleotides and derivatives	Nucleotides and derivatives	−1.01	−1.09
pme0516	Inositol	Others	Saccharides and Alcohols	−1.66	−2.42

**Table 3.** The consistently significantly regulated metabolites in the 45 vs. 44 and 45 vs. 66 groups.

Metabolome ID	Compounds	Class I	Class II	Log2FC (45 vs. 44)	Log2FC (45 vs. 66)
Cmln000394	4-(Methylthio)-3-OH-butyl glucosinolate (glucoraphanin)	Others	Glucosinolates	1.41	1.37
CMLN000400	4-Methylsulfinyl-3-butenyl thioglucoside (glucoraphenin)	Others	Glucosinolates	1.83	1.33
GQ512003	Fer-agmatine	Alkaloids	Phenolamine	1.25	3.54
Hmgp006596	Cyclobrassinin	Alkaloids	Plumerane	1.13	1.86
mws0997	Petunidin 3-O-glucoside	Flavonoids	Anthocyanins	2.00	1.04
pmb0489	N-hexosyl-p-coumaroyl putrescine	Alkaloids	Phenolamine	3.60	2.75
pmb0490	N-p-coumaroyl putrescine	Alkaloids	Phenolamine	2.68	2.44
pmb0494	N-sinapoyl putrescine	Alkaloids	Phenolamine	2.17	1.98
pmb0496	N-feruloyl agmatine	Alkaloids	Phenolamine	1.15	3.72
pmb0608	Chrysoeriol O-malonylhexoside	Flavonoids	Flavonoid	1.21	1.10
pmb3081	Glucarate O-phosphoric acid	Others	Saccharides and alcohols	1.27	1.70

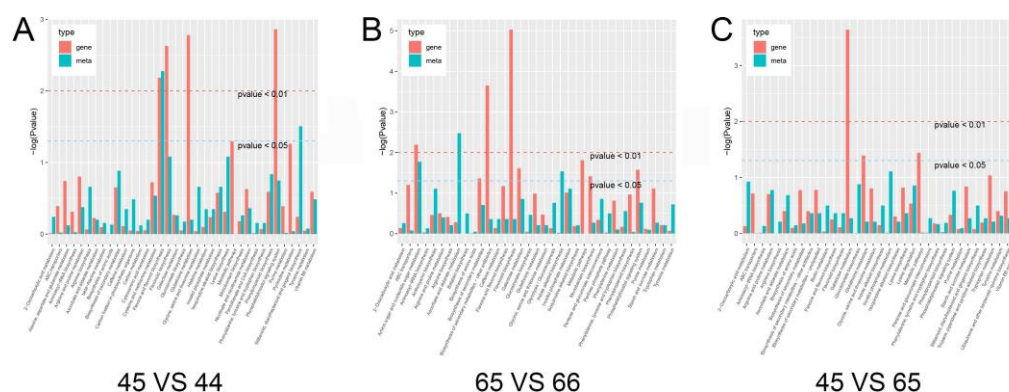
Table 3. Cont.

Metabolome ID	Compounds	Class I	Class II	Log2FC (45 vs. 44)	Log2FC (45 vs. 66)
pme0008	L-citrulline	Amino acids and derivatives	Amino acids and derivatives	1.18	1.27
pme2596	4-Pyridoxic acid	Others	Vitamin	1.26	1.30
pme3388	H-homoArg-OH	Amino acids and derivatives	Amino acids and derivatives	2.30	1.84
Rfmb319	Pipecolic acid	Amino acids and derivatives	Amino acids and derivatives	1.03	1.56
Cmyn001733	1-Methylpropyl glucosinolate	Others	Glucosinolates	−4.29406	−3.09
CWJP002007	Kaempferol-3-O-feruloyl-sophoroside-7-oglucoside	Flavonoids	Flavonols	−1.57628	−1.29
Hmbn002228	kaemperferol-3-O-[2-O-(6-O-E-feruloyl)]-β-D-glucopyranosyl]-β-galactopyranoside	Others	Others	−1.14448	−2.39
Lmhn002423	vnilloylmalic acid	Phenolic acids	Phenolic acids	−4.00812	−2.29
Lnrp102522	Kaempferol glc-glc-rha	Flavonoids	Flavonols	−3.30626	−1.23
mws1212	Methyl ferulate	Phenolic acids	Phenolic acids	−1.32806	−1.24
mws1550	S-allyl-L-cysteine	Amino acids and derivatives	Amino acids and derivatives	−1.65983	−1.36
pmb0382	O-feruloyl 4-hydroxycoumarin	Lignans and Coumarins	Coumarins	−2.10723	−3.09
pmb2620	3,4-Dimethoxycinnamic acid	Phenolic acids	Phenolic acids	−1.35293	−1.09
pmb3072	3-O-p-coumaroyl shikimic acid O-hexoside	Phenolic acids	Phenolic acids	−1.81168	−2.06
pme2914	3-Hydroxy-3-methylpentane-1,5-dioic acid	Amino acids and derivatives	Amino acids and derivatives	−1.70006	−1.99
pme3382	N-acetylthreonine	Amino acids and derivatives	Amino acids and derivatives	−1.98759	−1.27
Zmhn001375	Vanillic acid-C-glucoside	Phenolic acids	Phenolic acids	−1.91953	−1.30
Zmhn001883	Vanillic acid glycoside	Phenolic acids	Phenolic acids	−1.64013	−1.12

### 2.5. Conjoint Analysis of Genes and Metabolites Related to Purple Leaf Veins and Notched Leaves

To reveal the relationship between key genes and metabolites associated with purple leaf veins and leaf edge cracks, the DEGs and DAMs in the 45 vs. 44, 65 vs. 66, and 45 vs. 65 groups were respectively mapped to the corresponding KEGG pathways. According to the results of the GO and KEGG enrichment analyses, the enrichment degrees of both the DEGs and DAMs in the pathways were shown with histograms (Figure 7). For instance, in the groups with purple leaf vein samples (45 vs. 44 and 65 vs. 66), the pathways for flavone, flavonol, and flavonoid biosynthesis were coenriched with higher credibility (Figure 7A,B). In addition, we analyzed the correlation between DEGs and DAMs in different groups and selected the results with a Pearson correlation coefficient greater than 0.8 (Table S6). The selected DEGs and DAMs were shown with correlation coefficients using a cluster heat map. The results showed that the pathway for flavonoid biosynthesis had a high enrichment degree for both differential genes and metabolites in the 45 vs. 44 and 65 vs. 66 groups

(Figure S4A,B). In the 45 vs. 44 and 45 vs. 65 groups, for notched leaves, the pathways for phenolic acids and flavonoid biosynthesis had a high enrichment degree (Figure S4A,C).

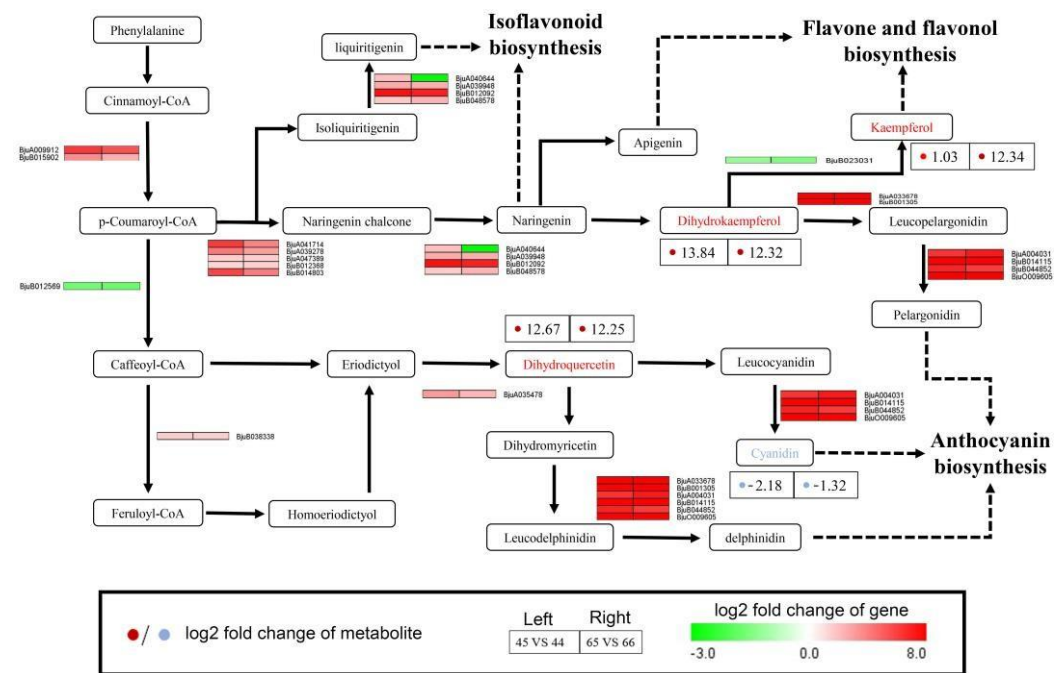


**Figure 7.** KEGG enrichment analysis of the DEGs and DAMs that were enriched in the same pathway. (A) KEGG enrichment analysis of the DEGs and DAMs in the 45 vs. 44 group. (B) KEGG enrichment analysis of the DEGs and DAMs in the 65 vs. 66 group. (C) KEGG enrichment analysis of the DEGs and DAMs in the 45 vs. 65 group. The abscissa represents the metabolic pathway, the red bars represent the enriched differential genes, and the blue bars represent the enriched differential metabolites (as expressed by  $\log(p\text{-value})$ ). The higher the ordinate, the stronger the enrichment degree.

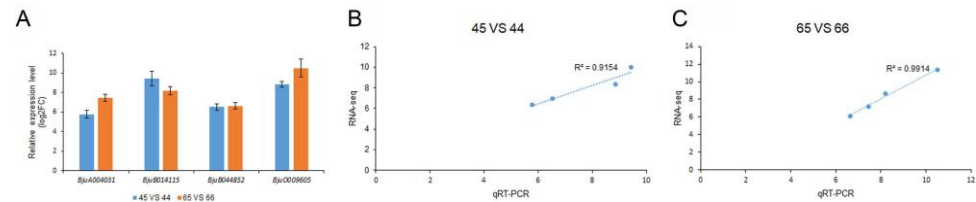
### 2.6. Integrating Related Genes and Metabolites in the Flavonoid Biosynthesis Pathway Provides Insights into Purple Vein Formation

Previous studies have shown that flavonoids and anthocyanins are important metabolites that lead to leaf color changes. To further clarify this point in this study, the DEGs and DAMs identified in both purple leaf vein samples (44 and 66) were selected to map to the flavonoid biosynthesis pathway. A total of 21 DEGs and four metabolites were screened (Figure 8). The results showed that the metabolites of dihydrokaempferol, dihydroquercetin, and kaempferol were highly accumulated in both 44 and 66. Dihydrokaempferol and kaempferol were the important metabolites in flavone and flavonol biosynthesis. In this pathway, the gene *BjuB023031* was downregulated, which regulated the synthesis of kaempferol. This was inconsistent with the accumulation of kaempferol. In addition, dihydrokaempferol also participated in anthocyanin biosynthesis. On this branch, the genes *BjuA033678*, *BjuB001305*, *BjuA004031*, *BjuB014115*, *BjuB044852*, and *BjuO009605*, which were annotated as anthocyanin synthase oxygenase (*ANS*), showed high expression levels. Dihydroquercetin was a vital intermediate metabolite involved in anthocyanin biosynthesis. Following this metabolite, there were two branches for anthocyanin biosynthesis. On one branch, although the expression levels of genes associated with cyanidin biosynthesis were high, cyanidin involved in the pathway of anthocyanin biosynthesis presented as less accumulated in 44 and 66 compared with 45 and 65 separately. However, on the other branch, genes associated with delphinidin biosynthesis showed higher expression levels, while the contents of delphinidin in 44 and 66 were similar to those in 45 and 65, respectively.

To further validate the RNA-Seq expression profile data, four genes (*BjuA004031*, *BjuB014115*, *BjuB044852*, and *BjuO009605*) with high expression in the flavonoid biosynthesis pathway were selected for the qRT-PCR assays. The results showed that all four genes were consistent with respect to the expression profiles between the qRT-PCR analysis and the RNA-Seq data (Figure 9).



**Figure 8.** Detailed information on the DAMs and DEGs in the pathways for phenylpropanoid and flavonoid biosynthesis. The rectangle represents the DEGs (the red color represents high expression and the green color represents low expression). The dot represents the DAMs (the red color represents high expression and the blue color represents low expression).



**Figure 9.** Gene expression and correlation analysis between qRT-PCR and RNA-Seq. (A) qRT-PCR analysis of 4 selected candidate genes. The blue color indicates the log<sub>2</sub>FC values in the comparison group of 45 vs. 44, the orange color indicates the log<sub>2</sub>FC values in the comparison group of 65 vs. 66. (B) Correlation analysis of gene expression between RNA-Seq and qRT-PCR in the comparison group of 45 vs. 44. (C) Correlation analysis of gene expression between RNA-Seq and qRT-PCR in the comparison group of 65 vs. 66.

### 3. Discussion

It is well known that genes determine the phenotypes of organisms and that metabolites are thought of as a “readout” of physiological states [33]. Therefore, exploring the association of gene-expression levels and metabolite abundances has attracted interest in recent years [34,35]. In the present study, we used four mustard samples that showed different phenotypes. Among these samples (44, 45, 65, and 66), 45 had green veins and an entire leaf, 44 showed purple veins and a serrated leaf, 65 showed green veins and a lobed leaf, and 66 exhibited purple veins and a lobed leaf (Figure 1). Using these samples, an integrated transcriptome and metabolome analysis was conducted to explain the variations in leaf color and leaf shape in mustard.

Flavonoids are the largest type of secondary metabolites and are the main compounds that determine the color of flowers, fruits, and leaves [36,37]. As important subclasses of flavonoids, flavonols and proanthocyanidins were proved to be the main pigment components in cucumber black thorn and orange peel [38], while the yellow peels of fruits such as lemons, oranges, limes, grapes, and tomatoes are often rich in flavanones [39,40].

Shi et al. showed that structurally modified anthocyanins and major potential copigmented flavonoids were the chemicals primarily responsible for the purple coloration of tea leaves [41]. In this study, dihydrokaempferol and kaempferol, which are important metabolites in the flavone synthesis pathway, were highly abundant in purple leaf veins compared with green veins (Figure 8). However, during the transformation from dihydrokaempferol to kaempferol, we identified only one differentially expressed regulatory gene (*BjuB023031*) that was downregulated in purple leaf veins. There may be some unknown regulatory mechanisms that lead to an increased accumulation of kaempferol. Excessive accumulation of kaempferol causes changes in the color of plant organs. Silva et al. (2020) found that high amounts of anthocyanins and other flavonoids such as rutin and kaempferol were responsible for the purple color of tomato [42]. In summary, our results showed that the excessive accumulation of dihydrokaempferol and kaempferol was involved in the formation of purple leaf veins in the 44 and 66 samples.

In addition, previous studies showed that the formation of purple leaves was mainly due to the accumulation of anthocyanins [43–45]; the pathway of anthocyanins in higher plants is well known, but the regulation of the biosynthetic mechanism in mustard remains unclear. Therefore, we focused on the regulatory pathway of anthocyanin synthesis. There were three anthocyanin synthesis pathways (Figure 8). It is well known that the metabolite of cyanidin is the most predominant anthocyanin found in nature [46]. In our study, there were four differentially expressed genes (*BjuA004031*, *BjuB014115*, *BjuB044852*, and *BjuO009605*) that regulated cyanidin synthesis. Although these four genes were consistently highly expressed in the purple leaf veins of the 44 and 66 samples, the accumulation of cyanidin was downregulated. Metabolites often present the final phenotype of plants, as the low accumulation of cyanidin results in low anthocyanin. Therefore, this branched pathway in anthocyanin biosynthesis was not the regulation pathway of purple leaf vein formation in mustard. Pelargonidin-based anthocyanins are another important branch for color characteristics [47]. Some scholars found that pelargonidin-based anthocyanins had a critical impact on the color and tinctorial strength of anthocyanin extracts [48]. Our results revealed that four genes were highly expressed; these were the vital genes in regulating the synthesis of pelargonidin. In this branch, from phenylalanine to anthocyanin, almost all differential genes were upregulated, indicating that the synthesis of pelargonidin-based anthocyanins was an important candidate regulatory pathway for purple leaf vein formation. The last pathway was the synthesis of delphinidin-based anthocyanins. Due to the large accumulation of delphinidin-related anthocyanins, the tea cultivar ‘Ziyan’ showed that a dark purple [49] and excessive accumulation of delphinidin-based anthocyanins resulted in novel bluer-colored *Chrysanthemums* [50]. In our results, genes and metabolites involved in the synthesis of delphinidin-based anthocyanins showed significant differences. Except for one gene (*BjuB012569*) regulating caffeoyl-CoA synthesis, which was downregulated, the other differentially expressed genes and proteins in this pathway were upregulated in purple-veined leaves. As a precursor metabolite of delphinidin synthesis, the dihydroquercetin metabolic level increased significantly in the two purple leaf vein samples. The above results indicated that the delphinidin-based anthocyanins also contributed to the purple leaf vein formation in the 44 and 66 samples.

The shape of a plant leaf is also an obvious phenotype that not only restricts the yield of crops, but also affects stress resistance in plants. In this study, four mustard leaf samples showed three different phenotypes: 45 had an entire leaf, 44 showed a serrated leaf, and 65 and 66 showed a lobed leaf (Figure 1). The K-means cluster analysis showed genes in subclass 7 were highly expressed in the serrated and lobed leaves but showed low expression in the entire leaf. Moreover, the deeper the degree of leaf notching, the higher the expression levels of these genes (Figure 3). A GO-enriched analysis indicated that “isomerase activity”, “extracellular region”, “cell plate”, “apoplast”, and “extracellular space” were the most significantly enriched cellular components. Previous studies showed that the environment also affected leaf shape, with different geographical locations resulting in different leaf morphologies for potato [29]. In a similar finding, genes that were highly

expressed in 44 and 65 compared with 45 were enriched in “systemic acquired resistance”, “alpha-farnesene synthase activity”, and “glucan endo-1,3-beta-D-glucosidase activity” (Figure 4D). All of them were related to adversity, indicating that the formation of serrated and lobed leaves was selected by environmental pressure [51].

## 4. Materials and Methods

### 4.1. Plant Materials

The purple leaf vein lines (44 and 66) and green leaf vein lines (45 and 65), which are self-bred lines bred by Anhui Science and Technology University, were planted in the experimental field at Anhui Science and Technology University (Fengyang, China). Among the four samples, 45 showed an entire leaf, 44 showed a serrated leaf, and 65 and 66 showed a lobed leaf. All the samples were collected from mature functional leaves, frozen in liquid nitrogen, and stored at  $-80\text{ }^{\circ}\text{C}$  until use in RNA and metabolite extraction. All experiments were analyzed with three biological repeats.

### 4.2. RNA Sequencing and Data Analysis

The total RNA was extracted using a TaKaRa MiniBEST Plant RNA Extraction Kit (Code No.: 9769) following the manufacturer’s instructions. An Agilent 2100 RNA Nano 6000 Assay Kit (Agilent Technologies, Santa Clara, CA, USA) was used to detect the integrity and concentration of the RNA samples. Sequencing libraries were established using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (New England BioLabs, Inc., Ipswich, MA, USA) and sequenced on an Illumina Navoseq 6000 platform. Paired-end reads of these samples were generated.

By removing adapter sequences, ploy-N, and low-quality reads with Trimmomatic software (version = 0.39) [52], clean reads were obtained and checked with FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>, accessed on 5 January 2022) for the subsequent steps. Clean reads were mapped to the *B. juncea* reference genome [53] using TopHat v2.0.12, and the expression level of each gene was measured in fragments per kilobase of transcript per million fragments mapped (FPKM). The DESeq R package (1.18.0) was used to identify the differentially expressed genes (DEGs) with an adjusted false discovery rate (FDR)  $< 0.05$  and  $|\log_2 \text{FC (fold change)}| \geq 1$ .

Functional categorization of DEGs was analyzed using Blast2GO (version 2.6.4) (<http://www.blast2go.org>, accessed on 9 January 2022) [54]. Moreover, the DEGs were blasted against the online Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<http://geneontology.org>, accessed on 9 January 2022) to retrieve their KOs and were subsequently mapped to pathways in KEGG [55]. GO enrichment of three ontologies (biological process, molecular function, and cellular component) and KEGG pathway enrichment analyses were applied based on Fisher’s exact test; the functional categories and pathways with a  $p$ -value  $< 0.05$  were considered to be significant.

### 4.3. Metabolic Analysis

A total of 100 mg of powder ground from each freeze-dried leaf was weighted and extracted overnight at  $4\text{ }^{\circ}\text{C}$  with 0.6 mL of 70% aqueous methanol. Following centrifugation at  $10,000 \times g$  for 10 min, the extracts were absorbed and filtrated. The purified samples were loaded into the UPLC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system; MS, Applied Biosystems 4500 Q TRAP). The parameters were set as follows: UPLC: column, Waters ACQUITY UPLC HSS T3 C18 ( $1.8\text{ }\mu\text{m}$ ,  $2.1\text{ mm} \times 100\text{ mm}$ ); the mobile phase consisted of solvent A (pure water with 0.04% acetic acid) and solvent B (acetonitrile with 0.04% acetic acid). Sample measurements were performed with a gradient program. The column oven was set to  $40\text{ }^{\circ}\text{C}$ . The injection volume was  $4\text{ }\mu\text{L}$ . The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-MS. The ESI source operation parameters were set as follows: ion source, turbo spray; source temperature,  $550\text{ }^{\circ}\text{C}$ ; ion spray voltage (IS),  $5500\text{ V}$  (positive ion mode)/ $-4500\text{ V}$  (negative ion mode); ion sources gas I (GSI), gas II (GSII), and curtain gas (CUR) were set at 50, 60,

and 30.0 psi, respectively; the collision gas (CAD) was high [56]. The raw files produced by the UPLC-ESI-MS/MS assays were analyzed with Analyst 1.6.3 software. Q3 was used for metabolite quantification, while Q1, Q3, RT (retention time), DP (declustering potential), and CE (collision energy) were used for metabolite identification.

A combination of fold change and the variable importance in projection (VIP) value of the OPLS-DA model was used to screen the differentially accumulated metabolites (DAMs) between groups. Significantly regulated metabolites between groups were determined using  $VIP > 1$  and absolute  $\log_2FC$  (fold change)  $> 1$ . The identified metabolites were annotated using the KEGG compound database (<http://www.kegg.jp/kegg/compound>, accessed on 23 January 2022); annotated metabolites were then mapped to KEGG pathway database (<http://www.kegg.jp/kegg/pathway.html>, accessed on 23 January 2022). The pathways mapped with significantly regulated metabolites were then fed into a metabolite sets enrichment analysis (MSEA); their significances were determined using the hypergeometric test's  $p$ -values.

#### 4.4. Combined Analysis of DEGs and DAMs

The obtained DEGs and DAMs were mapped to the KEGG pathway for integration. The principal component analysis (PCA) diagram, hierarchical clustering analysis diagram, and Venn diagrams were drawn using the R software package. The Pearson correlation coefficients (PCCs) of the DEGs and the significantly regulated metabolites were calculated using the COR function in R. The genes and metabolites with a  $PCC > 0.8$  in the KEGG pathway were used to establish the related network, which was visualized using Cytoscape software (version 3.7.1) [57].

#### 4.5. qRT-PCR Validation of RNA-Seq Results

In total, 4 genes were selected for qRT-PCR analysis to examine the expression profiles of the RNA-Seq. The gene of BjActin was used as the internal control. The primers of the 4 genes for qRT-PCR were designed using Primer 6 software (Table S7). The cDNA was synthesized from the total RNA used for the RNA-Seq, and then the qRT-PCR analysis was performed on a Bio-Rad CFX96 using  $2 \times$  SYBR green PCR master mix (Applied Biosystems) with three biological replicates. The relative expression of each gene was calculated using the  $2^{-\Delta\Delta CT}$  method.

#### 4.6. Statistical Analysis

Statistical analysis was performed with SPSS software. All the data are shown with the average values of three repeated samples. Statistical significance was determined using the least significant difference test ( $p < 0.05$ ).

### 5. Conclusions

In the present study, the main regulatory pathways of purple leaf veins and leaf edge cracks in mustard were revealed using combined transcriptomic and metabolomics analyses. Based on the differences in leaf color among the different mustard samples, the DEGs and DAMs with similar expression trends were identified using a K-means analysis. The highly expressed genes and metabolites in the purple leaf vein samples were mainly enriched in the flavonoid biosynthesis pathway. Integrating related genes and metabolites in the flavonoid biosynthesis pathway revealed that synthetic pathways of pelargonidin-based anthocyanins and delphinidin-based anthocyanins were the important pathways in regulating the formation of purple leaf veins in the 44 and 66 samples. For the formation of leaf edge cracks, "alpha-farnesene synthase activity" and "glucan endo-1,3-beta-D-glucosidase activity" related to adversity were mainly enriched in the serrated and lobed leaf samples, indicating that environmental pressure was the dominant factor in the formation of leaf edge cracks. In conclusion, our results provided new insights into the formation of purple leaf veins and leaf edge cracks in mustard, as well as a theoretical reference for further studies of leaf phenotypes.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/plants11172229/s1>, Figure S1: The analysis results of the metabolic profiles obtained by UPLC-MS/MS. (A) Principal component analysis (PCA) score plot of all metabolites in 12 samples. The samples with the same color represent three biological replications of the same sample. (B) Correlation analysis of metabolome data of four samples. (C) The hierarchical clustering analysis and heatmaps of differential accumulation metabolites; Figure S2: Volcano map of differential metabolite. Each point in the volcanic map represents a metabolite, the x-coordinate represents the logarithm of the fold change of a metabolite in the two samples, and the y-coordinate represents the VIP value; Figure S3: The Venn diagram depicts the common and unique number of differential accumulation metabolites among the four groups of samples; Figure S4: Correlation coefficient cluster heat map of differentially accumulated metabolites and differentially expressed genes; Table S1: The information of DEGs in different subclasses; Table S2: GO analysis of DEGs in subclass 3; Table S3: GO analysis of DEGs in subclass 7; Table S4: KEGG analysis of DEGs in subclass 3; Table S5: KEGG analysis of DEGs in subclass 7; Table S6: Genes and metabolites with the Pearson correlation coefficient greater than 0.8 in the 45 vs. 44 group; Table S7: List of primers used in quantitative RT-PCR. All primers are designed by Primer 6.

**Author Contributions:** X.W. and L.W. conceived the research and designed the experiments; K.Z. performed the research, analyzed the data, and wrote the manuscript; D.Y. and Y.H. participated in preparing the samples; J.Q. and M.K.N. helped with the bioinformatics analysis; L.J., C.Y. and Z.L. analyzed and interpreted the data. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (32002061), the Natural Science Foundation of Hebei Science and Technology Program (C2021402014 and C2019402330), the Talent Foundation of Anhui Science and Technology University (NXYJ202103), and the College Students' Innovative Entrepreneurial Training Plan Program (202110879081 and S202110879222).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw data from the comparative transcriptome analysis were deposited into the Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) under accession number PRJNA800112.

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

## References

1. Fadhil, A.B.; Saleh, L.A.; Altamer, D.H. Production of biodiesel from non-edible oil, wild mustard (*Brassica Juncea* L.) seed oil through cleaner routes. *Energy Sources Part A Recovery Util. Environ. Eff.* **2020**, *42*, 1831–1843. [\[CrossRef\]](#)
2. Voss-Fels, K.; Frisch, M.; Qian, L.; Kontowski, S.; Friedt, W.; Gottwald, S.; Snowdon, R.J. Subgenomic diversity patterns caused by directional selection in bread wheat gene pools. *Plant Genome* **2015**, *8*, 0013. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Jin, F.; Li, S.; Dang, L.; Chai, W.T.; Li, P.L.; Wang, N.N. *PL1* fusion gene: A novel visual selectable marker gene that confers tolerance to multiple abiotic stresses in transgenic tomato. *Transgenic Res.* **2012**, *21*, 1057–1070. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Heng, S.; Cheng, Q.; Zhang, T.; Liu, X.J.; Huang, H.; Yao, P.J.; Liu, Z.X.; Wan, Z.G.; Fu, T.D. Fine-mapping of the *BjPur* gene for purple leaf color in *Brassica juncea*. *Theor. Appl. Genet.* **2020**, *133*, 2989–3000. [\[CrossRef\]](#)
5. Neill, S.O.; Gould, K.S.; Kilmartin, P.A.; Mitchell, K.A.; Markham, K.R. Antioxidant activities of red versus green leaves in *Elatostema rugosum*. *Plant Cell Environ.* **2002**, *25*, 539–547. [\[CrossRef\]](#)
6. Oh, J.E.; Kim, Y.H.; Kim, J.H.; Kwon, Y.R.; Lee, H. Enhanced level of anthocyanin leads to increased salt tolerance in *Arabidopsis* PAP1-D plants upon sucrose treatment. *J. Korean Soc. Appl. Biol. Chem.* **2011**, *54*, 79–88. [\[CrossRef\]](#)
7. Gonzalez, A.; Zhao, M.; Leavitt, J.M.; Lloyd, A.M. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.* **2008**, *53*, 814–827. [\[CrossRef\]](#)
8. Qi, T.; Song, S.; Ren, Q.; Ren, G.C.; Wu, D.W.; Huang, H.; Chen, Y.; Fan, M.; Peng, W.; Ren, C.M.; et al. The Jasmonate-ZIM-domain proteins interact with the WD-Repeat/bHLH/MYB complexes to regulate Jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. *Plant Cell* **2011**, *23*, 1795–1814. [\[CrossRef\]](#)
9. Carey, C.C.; Strahle, J.T.; Selinger, D.A.; Chandler, V.L. Mutations in the *pale aleurone color1* regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar *TRANSPARENT TESTA GLABRA1* gene in *Arabidopsis thaliana*. *Plant Cell* **2004**, *16*, 450–464. [\[CrossRef\]](#)

10. Hu, W.; Zhou, T.; Han, Z.; Tan, C.; Xing, Y.Z. Dominant complementary interaction between *OsC1* and two tightly linked genes, *Rb1* and *Rb2*, controls the purple leaf sheath in rice. *Theor. Appl. Genet.* **2020**, *133*, 2555–2566. [[CrossRef](#)]
11. Chiu, L.W.; Zhou, X.; Burke, S.; Wu, X.L.; Prior, R.L.; Li, L. The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol.* **2010**, *154*, 1470–1480. [[CrossRef](#)] [[PubMed](#)]
12. Song, H.Y.; Lee, M.; Han, C.T.; Lee, J.; Kim, H.; Park, J.I.; Nou, I.S.; Kim, S.J.; Hur, Y. Purple *Brassica oleracea* var. *capitata* F. *rubra* is due to the loss of *BoMYBL2-1* expression. *BMC Plant Biol.* **2018**, *18*, 82. [[CrossRef](#)] [[PubMed](#)]
13. Guo, N.; Wu, J.; Zheng, S.; Cheng, F.; Liu, B.; Liang, J.L.; Cui, Y.; Wang, X.W. Anthocyanin profile characterization and quantitative trait locus mapping in zicaikai (*Brassica rapa* L. ssp. *chinensis* var. *purpurea*). *Mol. Breed.* **2015**, *35*, 113. [[CrossRef](#)]
14. Tang, Q.; Tian, M.; An, G.; Zhang, W.Y.; Chen, J.J.; Yan, C.H. Rapid identification of the purple stem (Ps) gene of Chinese kale (*Brassica oleracea* var. *alboglabra*) in a segregation distortion population by bulked segregant analysis and RNA sequencing. *Mol. Breed.* **2017**, *37*, 153. [[CrossRef](#)]
15. Runions, A.; Tsiantis, M.; Prusinkiewicz, P. A common developmental program can produce diverse leaf shapes. *New Phytol.* **2017**, *216*, 401–418. [[CrossRef](#)]
16. Vogel, S. Leaves in the lowest and highest winds: Temperature, force and shape. *New Phytol.* **2009**, *183*, 13–26. [[CrossRef](#)]
17. Sicard, A.; Thamm, A.; Marona, C.; Lee, Y.W.; Wahl, V.; Stinchcombe, J.R.; Wright, S.I.; Kappel, C.; Lenhard, M. Repeated evolutionary changes of leaf morphology caused by mutations to a homeobox gene. *Curr. Biol.* **2014**, *24*, 1880–1886. [[CrossRef](#)]
18. Rowland, S.D.; Zumstein, K.; Nakayama, H.; Cheng, Z.Z.; Flores, A.M.; Chitwood, D.H.; Maloof, J.N.; Sinha, N.R. Leaf shape is a predictor of fruit quality and cultivar performance in tomato. *New Phytol.* **2020**, *226*, 851–865. [[CrossRef](#)]
19. Li, Z.; Li, B.; Shen, W.H.; Huang, H.; Dong, A.W. TCP transcription factors interact with *AS2* in the repression of class-I KNOX genes in *Arabidopsis thaliana*. *Plant J.* **2012**, *71*, 99–107. [[CrossRef](#)]
20. Bar, M.; Ori, N. Compound leaf development in model plant species. *Curr. Opin. Plant Biol.* **2015**, *23*, 61–69. [[CrossRef](#)]
21. Sarvepalli, K.; Gupta, M.D.; Challa, K.R.; Nath, U. Molecular cartography of leaf development—role of transcription factors. *Curr. Opin. Plant Biol.* **2019**, *47*, 22–31. [[CrossRef](#)] [[PubMed](#)]
22. Weiss, D.; Ori, N. Mechanisms of cross talk between gibberellin and other hormones. *Plant Physiol.* **2007**, *144*, 1240–1246. [[CrossRef](#)] [[PubMed](#)]
23. Sarvepalli, K.; Nath, U. Interaction of TCP4-mediated growth module with phytohormones. *Plant Signal. Behav.* **2011**, *6*, 1440–1443. [[CrossRef](#)]
24. Hibara, K.; Karim, M.R.; Takada, S.; Taoka, K.I.; Furutani, M.; Aida, M.; Tasaka, M. Arabidopsis CUP-SHAPED COTYLEDON3 regulates postembryonic shoot meristem and organ boundary formation. *Plant Cell* **2006**, *18*, 2946–2957. [[CrossRef](#)] [[PubMed](#)]
25. Nikovics, K.; Blein, T.; Peaucelle, A.; Ishida, T.; Morin, H.; Aida, M.; Laufs, P. The balance between the *MIR164A* and *CUC2* genes controls leaf margin serration in *Arabidopsis*. *Plant Cell* **2006**, *18*, 2929–2945. [[CrossRef](#)]
26. Hasson, A.; Plessis, A.; Blein, T.; Adroher, B.; Grigg, S.; Tsiantis, M.; Boudaoud, A.; Damerval, C.; Laufs, P. Evolution and diverse roles of the CUP-SHAPED COTYLEDON genes in *Arabidopsis* leaf development. *Plant Cell* **2011**, *23*, 54–68. [[CrossRef](#)]
27. Saddic, L.A.; Huvermann, B.; Bezhani, S.; Su, Y.H.; Winter, C.M.; Kwon, C.S.; Collum, R.P.; Wagner, D. The *LEAFY* target *LMI1* is a meristem identity regulator and acts together with *LEAFY* to regulate expression of *CAULIFLOWER*. *Development* **2006**, *133*, 1673–1682. [[CrossRef](#)] [[PubMed](#)]
28. Hajheidari, M.; Wang, Y.; Bhatia, N.; Vuolo, F.; Franco-Zorrilla, J.M.; Karady, M.; Mentink, R.A.; Wu, A.H.; Oluwatobi, B.R.; Muller, b.; et al. Autoregulation of RCO by low-affinity binding modulates cytokinin action and shapes leaf diversity. *Curr. Biol.* **2019**, *29*, 4183–4192. [[CrossRef](#)]
29. Gupta, S.; Rosenthal, D.M.; Stinchcombe, J.R.; Baucom, R.S. The remarkable morphological diversity of leaf shape in sweet potato (*Ipomoea batatas*): The influence of genetics, environment, and G × E. *New Phytol.* **2020**, *225*, 2183–2195. [[CrossRef](#)]
30. Ni, X.; Huang, J.; Ali, B.; Zhou, W.J.; Zhao, J.Y. Genetic analysis and fine mapping of the LOBED-LEAF 1 (*BnLL1*) gene in rapeseed (*Brassica napus* L.). *Euphytica* **2015**, *204*, 29–38. [[CrossRef](#)]
31. Ni, X.; Liu, H.; Huang, J.; Zhao, J.Y. *LMI1*-like genes involved in leaf margin development of *Brassica napus*. *Genetica* **2017**, *145*, 269–274. [[CrossRef](#)] [[PubMed](#)]
32. Hu, L.; Zhang, H.; Yang, Q.; Meng, Q.W.; Han, S.Q.; Nwafor, C.C.; Khan, M.H.U.; Fan, C.C.; Zhou, Y.M. Promoter variations in a homeobox gene, *BnA10.LMI1*, determine lobed leaves in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* **2018**, *131*, 2699–2708. [[CrossRef](#)] [[PubMed](#)]
33. Fang, C.; Fernie, A.R.; Luo, J. Exploring the diversity of plant metabolism. *Trends Plant Sci.* **2019**, *24*, 83–98. [[CrossRef](#)] [[PubMed](#)]
34. Yamashita, H.; Kambe, Y.; Ohshio, M.; Kunihiro, A.; Tanaka, Y.; Suzuki, T.; Nakamura, Y.; Morita, A.; Ikka, T. Integrated metabolome and transcriptome analyses reveal etiolation-induced metabolic changes leading to high amino acid contents in a light-sensitive Japanese albino tea cultivar. *Front. Plant Sci.* **2021**, *11*, 2194. [[CrossRef](#)] [[PubMed](#)]
35. Szymański, J.; Bocobza, S.; Panda, S.; Sonawane, P.; Cárdenas, P.D.; Lashbrooke, J.; Kamble, A.; Shahaf, N.; Meir, S.; Bovy, A. Analysis of wild tomato introgression lines elucidates the genetic basis of transcriptome and metabolome variation underlying fruit traits and pathogen response. *Nat. Genet.* **2020**, *52*, 1111–1121. [[CrossRef](#)]
36. Deng, Y.; Lu, S. Biosynthesis and regulation of phenylpropanoids in plants. *Crit. Rev. Plant Sci.* **2017**, *36*, 257–290. [[CrossRef](#)]
37. Deng, Y.; Li, C.; Li, H.; Lu, S. Identification and characterization of flavonoid biosynthetic enzyme genes in *Salvia miltiorrhiza* (*Lamiaceae*). *Molecules* **2018**, *23*, 1467. [[CrossRef](#)]

38. Liu, M.; Zhang, C.; Duan, L.; Luan, Q.Q.; Li, J.L.; Yang, A.G.; Qi, X.Q.; Ren, Z.H. *CsMYB60* is a key regulator of flavonols and proanthocyanidins that determine the colour of fruit spines in cucumber. *J. Exp. Bot.* **2019**, *70*, 69–84. [[CrossRef](#)]
39. Peterson, J.J.; Dwyer, J.T.; Beecher, G.R.; Bhagwat, S.A.; Gebhardt, S.E.; Haytowitz, D.B.; Holden, J.M. Flavanones in oranges, tangerines (mandarins), tangors, and tangelos: A compilation and review of the data from the analytical literature. *J. Food Compos. Anal.* **2006**, *19*, S66–S73. [[CrossRef](#)]
40. Goulas, V.; Manganaris, G.A. Exploring the phytochemical content and the antioxidant potential of Citrus fruits grown in Cyprus. *Food Chem.* **2012**, *131*, 39–47. [[CrossRef](#)]
41. Shi, J.; Simal-Gandara, J.; Mei, J.; Ma, W.J.; Peng, Q.H.; Shi, Y.L.; Xu, Q.; Lin, Z.; Lv, H.P. Insight into the pigmented anthocyanins and the major potential co-pigmented flavonoids in purple-coloured leaf teas. *Food Chem.* **2021**, *363*, 130278. [[CrossRef](#)] [[PubMed](#)]
42. Da Silva Souza, M.A.; Peres, L.E.P.; Freschi, J.R.; Purgatto, E.; Lajolo, F.M.; Hassimotto, N.M. Changes in flavonoid and carotenoid profiles alter volatile organic compounds in purple and orange cherry tomatoes obtained by allele introgression. *J. Sci. Food. Agric.* **2020**, *100*, 1662–1670. [[CrossRef](#)] [[PubMed](#)]
43. Wang, L.; Pan, D.; Liang, M.; Abubakar, Y.S.; Li, J.; Lin, J.; Chen, S.; Chen, W. Regulation of anthocyanin biosynthesis in purple leaves of zijuan tea (*Camellia sinensis* var. *kitamura*). *Int. J. Mol. Sci.* **2017**, *18*, 833. [[CrossRef](#)] [[PubMed](#)]
44. Chaves-Silva, S.; Dos Santos, A.L.; Chalfun-Júnior, A.; Zhao, J.; Peres, L.E.P.; Benedito, V.A. Understanding the genetic regulation of anthocyanin biosynthesis in plants—tools for breeding purple varieties of fruits and vegetables. *Phytochemistry* **2018**, *153*, 11–27. [[CrossRef](#)] [[PubMed](#)]
45. Chen, X.; Wang, P.; Zheng, Y.; Gu, M.Y.; Lin, X.Y.; Wang, S.Y.; Jin, S.; Ye, N.X. Comparison of metabolome and transcriptome of flavonoid biosynthesis pathway in a purple-leaf tea germplasm Jinmingzao and a green-leaf tea germplasm Huangdan reveals their relationship with genetic mechanisms of color formation. *Int. J. Mol. Sci.* **2020**, *21*, 4167. [[CrossRef](#)] [[PubMed](#)]
46. Wallace, T.C.; Giusti, M.M. Anthocyanins. *Adv. Nutr.* **2015**, *6*, 620–622. [[CrossRef](#)]
47. Giusti, M.M.; Rodríguez-Saona, L.E.; Wrolstad, R.E. Molar absorptivity and color characteristics of acylated and non-acylated pelargonidin-based anthocyanins. *J. Agric. Food Chem.* **1999**, *47*, 4631–4637. [[CrossRef](#)]
48. Giusti, M.M.; Wrolstad, R.E. Acylated anthocyanins from edible sources and their applications in food systems. *Biochem. Eng. J.* **2003**, *14*, 217–225. [[CrossRef](#)]
49. Lai, Y.S.; Li, S.; Tang, Q.; Li, H.S.; Chen, S.X.; Li, P.W.; Xu, J.Y.; Xu, Y.; Guo, X. The dark-purple tea cultivar ‘Ziyan’ accumulates a large amount of delphinidin-related anthocyanins. *J. Agric. Food Chem.* **2016**, *64*, 2719–2726. [[CrossRef](#)]
50. Noda, N.; Aida, R.; Kishimoto, S.; Ishiguro, K.; Fukuchi-Mizutani, M.; Tanaka, Y.; Ohmiya, A. Genetic engineering of novel bluer-colored chrysanthemums produced by accumulation of delphinidin-based anthocyanins. *Plant Cell Physiol.* **2013**, *54*, 1684–1695. [[CrossRef](#)]
51. Chitwood, D.H.; Sinha, N.R. Evolutionary and environmental forces sculpting leaf development. *Curr. Biol.* **2016**, *26*, R297–R306. [[CrossRef](#)] [[PubMed](#)]
52. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
53. Yang, J.; Liu, D.; Wang, X.; Ji, C.; Cheng, F.; Liu, B.; Hu, Z.; Chen, S.; Pental, D.; Ju, Y.; et al. The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. *Nat. Genet.* **2016**, *48*, 1225–1232. [[CrossRef](#)] [[PubMed](#)]
54. Conesa, A.; Götz, S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genom.* **2008**, *2008*, 619832. [[CrossRef](#)]
55. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirekawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2007**, *36*, D480–D484. [[CrossRef](#)]
56. Chen, W.; Gong, L.; Guo, Z.; Wang, W.; Zhang, H.; Liu, X.; Yu, S.; Xiong, L.; Luo, J. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: Application in the study of rice metabolomics. *Mol. Plant* **2013**, *6*, 1769–1780. [[CrossRef](#)]
57. Kohl, M.; Wiese, S.; Warscheid, B. Cytoscape: Software for visualization and analysis of biological networks. *Humana Press* **2011**, *696*, 291–303.