Plant Communications
Research Article

CellPress
Partner Journal

Integrated network analyses identify *MYB4R1* neofunctionalization in the UV-B adaptation of Tartary buckwheat

Moyang Liu[1,2,3], Wenjun Sun[1], Zhaotang Ma[1,5], Chaocheng Guo[2,3], Jiahao Chen[2,3], Qi Wu[1], Xiyin Wang[4,*] and Hui Chen[1,*]

[1]College of Life Science, Sichuan Agricultural University, Ya'an 625014, China

[2]Joint Center for Single Cell Biology, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China

[3]Shanghai Collaborative Innovation Center of Agri-Seeds/School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China

[4]School of Life Science, North China University of Science and Technology, Tangshan 063210, China

[5]State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Key Laboratory of Major Crop Diseases and Rice Research Institute, Sichuan Agricultural University, Chengdu 611130, China

***Correspondence: Hui Chen (chenhui@sicau.edu.cn), Xiyin Wang (wangxiyin@vip.sina.com)**

https://doi.org/10.1016/j.xplc.2022.100414

ABSTRACT

A hallmark of adaptive evolution is innovation in gene function, which is associated with the development of distinct roles for genes during plant evolution; however, assessing functional innovation over long periods of time is not trivial. Tartary buckwheat (*Fagopyrum tataricum*) originated in the Himalayan region and has been exposed to intense UV-B radiation for a long time, making it an ideal species for studying novel UV-B response mechanisms in plants. Here, we developed a workflow to obtain a co-functional network of UV-B responses using data from more than 10,000 samples in more than 80 projects with multi-species and multi-omics data. Dissecting the entire network revealed that flavonoid biosynthesis was most significantly related to the UV-B response. Importantly, we found that the regulatory factor *MYB4R1*, which resides at the core of the network, has undergone neofunctionalization. *In vitro* and *in vivo* experiments demonstrated that *MYB4R1* regulates flavonoid and anthocyanin accumulation in response to UV-B in buckwheat by binding to L-box motifs in the *FtCHS*, *FtFLS*, and *FtUFGT* promoters. We used deep learning to develop a visual discrimination model of buckwheat flavonoid content based on natural populations exposed to global UV-B radiation. Our study highlights the critical role of gene neofunctionalization in UV-B adaptation.

Key words: neofunctionalization, deep learning, MYB, UV-B radiation, Tartary buckwheat

Liu M., Sun W., Ma Z., Guo C., Chen J., Wu Q., Wang X., and Chen H. (2022). Integrated network analyses identify *MYB4R1* neofunctionalization in the UV-B adaptation of Tartary buckwheat. Plant Comm. **3**, 100414.

INTRODUCTION

Ultraviolet-B (UV-B) is perceived by plants as an environmental signal and is a potential abiotic stressor that influences growth and adaptation (Shi and Liu, 2021). As an environmental signal, it can regulate aspects of photomorphogenesis, such as the inhibition of hypocotyl elongation, cotyledon expansion, and flavonoid accumulation (Kim et al., 1998; Kliebenstein et al., 2002; Favory et al., 2009; Wargent et al., 2009; Yadav et al., 2020); as an abiotic stress, sustained high-intensity radiation can cause direct DNA damage, oxidative stress, and lipid and protein oxidation, leading to abnormal growth and development (Britt, 1995; Shi and Liu, 2021). Plants use a variety of strategies, including morphological changes, endogenous macromolecules, and secondary metabolites, to repair DNA damage, detoxify reactive oxygen species, and reduce cellular exposure to UV-B to limit and repair UV-B-induced damage and promote adaptation to high-UV-B conditions (Jansen et al., 1998; Yin and Ulm, 2017; Vanhaelewyn et al., 2020). The deployment of antioxidant defenses and the accumulation of flavonoids, such as flavanols, anthocyanins, and proanthocyanidins, are among the most significant components of the hierarchical protection approach (Shi and Liu, 2021).

Although flavonoids have been demonstrated to play a crucial role in the stress response, flavonoid biosynthesis and regulatory mechanisms require further exploration. Flavonoid biosynthesis

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

includes eight branches (stilbene, aurone, flavone, isoflavone, flavonol, phlobaphene, proanthocyanidin, and anthocyanin biosynthesis) and four important intermediate metabolites (chalcone, flavanone, dihydroflavonol, and leucoanthocyanidin), whose structural diversity and antioxidant properties contribute to their physiological roles in plant interactions with the biotic and abiotic environment (Winkel-Shirley, 2001; Sasaki and Nakayama, 2015; Xu et al., 2015). MYB transcription factors play a key role in regulating flavonoid biosynthesis to enhance UV-B stress tolerance (Ambawat et al., 2013). MYB proteins are categorized into four types depending on the number and placement of their MYB structural domain repeats: 1R-MYB/MYB-related, R2R3-MYB, 3R-MYB, and 4R-MYB (Dubos et al., 2010). R2R3-MYB members are primarily engaged in the regulation of flavonoid biosynthesis, and most are positive regulators that increase the expression of genes in the structural flavonoid pathway (Jaakola, 2013). Overexpression of *AN4* (an R2R3-MYB-encoding gene) can, for example, boost anthocyanin biosynthesis by increasing the expression of anthocyanin biosynthetic genes, including *CHS*, *CHI*, *F3H*, and *DFR* (Zhang et al., 2021a). A few repressors have also been identified, such as *AtMYB3*, which influences malic acid, erucic acid, and anthocyanin biosynthesis (Zhou et al., 2017). Despite previous studies, flavonoid biosynthesis and regulation are not yet fully understood because flavonoids are part of a large, complex, and highly coordinated metabolic network rather than a collection of separate components (Winkel-Shirley, 2001; Zeng et al., 2020).

Tartary buckwheat (*Fagopyrum tataricum*) has been cultivated for 4000 years and is a nutritionally balanced and flavonoid-rich crop. It is cultivated on all continents because of its ecological adaptability, quick growing time, and tolerance to low-nutrient conditions (Zhang et al., 2021b). Tartary buckwheat originated in the Himalayan area, which has strong UV-B radiation levels as well as low temperatures and pressures (Ferl et al., 2002; Norsang et al., 2009; Zhang et al., 2015). Tartary buckwheat has undergone whole-genome duplication (WGD), inducing significant expansion of numerous gene families that influence abiotic stress tolerance, including MYBs, the calcium-dependent protein kinase (CDPK) family, bZIP transcription factors, and the mitogen-activated protein (MAP) kinase family. Expansion of the R2R3-MYB subfamily related to flavonoid biosynthesis, in particular, shows that the family has evolved more diversified functions in regulating flavonoid biosynthesis and abiotic stress responses (Zhang et al., 2017), making Tartary buckwheat an attractive species for studying novel UV-B response mechanisms in plants (Liu et al., 2018; Zhang et al., 2021b). The buckwheat genome and many transcriptome resources have been published (Zhang et al., 2017, 2021b; Liu et al., 2018; Song et al., 2020), but owing to the lack of methods for deeply mining the data and stable genetic transformation systems, significant challenges remain.

Recently, several tools, such as COXPRESdb (Obayashi et al., 2019), WeGET (Szklarczyk et al., 2016), GeneMANIA (Warde-Farley et al., 2010), WGCNA (Langfelder and Horvath, 2008), and CLIC (Li et al., 2017), have been used to assign putative new functions to genes through correlations or coexpression networks. At their core, these approaches rely on the concept of guilt by association: transcripts or proteins that exhibit similar expression patterns are often functionally related (Eisen et al., 1998). However, these approaches typically rely on discrete subsets of genes in which the expression correlation exceeds a hard or soft threshold, and this strongly influences the final result. The GeneBridge toolkit is a newly developed and effective tool for identifying new functions and new relationships between genes based on large-scale transcriptomes (Li et al., 2019). The toolkit considers the continuous expression level of the target gene in the whole population and establishes a relationship with the biological process based on the coexpression pattern of the target gene relative to all genes. Its prediction performance is better than that of the new function methods used for gene allocation, such as WeGET (Szklarczyk et al., 2016) and COXPRESdb (Obayashi et al., 2019), and the simpler method based on the average of the correlation coefficient between genes (average *r)*. Using this tool, a new function of *DDT* that affects mitochondrial respiration was successfully discovered in animal studies in combination with experiments (Li et al., 2019).

Here, we developed a workflow to obtain a co-functional network of UV-B responses from multi-omics data of nearly 10 000 samples. This led to the observation that flavonoid biosynthesis formed a top-level association with the UV-B response, and their conserved association in plants was characterized using the OneKP project. We found that *MYB4R1*, which resides at the core of the co-functional network of flavonoid biosynthesis and UV-B response in Tartary buckwheat, was neofunctionalized. Finally, we characterized the function of *FtMYB4R1* in regulating flavonoid and anthocyanin biosynthesis and developed a visual discriminatory model for buckwheat flavonoid content based on deep learning.

RESULTS

Obtaining a co-functional network of the UV-B response

We obtained plant transcriptomes from the GEO, ArrayExpress, and CNGBdb public repositories that contained whole-genome gene expression data previously generated with next-generation sequencing technologies. The GeneBridge toolkit is a newly developed and effective tool for identifying new functions of genes and new connections between functions based on large-scale transcriptome data (Li et al., 2019). Therefore, to fully characterize the co-functional network of the UV-B response, we first collected transcriptomes from more than 80 projects, including 10 000 samples of typical monocot and dicot species (*Arabidopsis thaliana* and *Oryza sativa*) and *F. tataricum* (Supplemental Table 1). Next, for these transcriptomes, using genes that were involved in the UV-B response (Gene ontology [GO]: GO0010224), we applied the Gene-Module Association Determination (G-MAD) tool in the GeneBridge toolkit to the different species to search for genes potentially related to the UV-B response (Figure 1A).

In *A. thaliana*, a total of 739 genes related to the UV-B response were obtained. This collection contains all 43 genes known to be annotated in the UV-B response, such as UVR8 (Tissot and Ulm, 2020; Dong et al., 2021; Serrano et al., 2021), providing evidence that the GeneBridge toolkit can recover the functions of known genes (Figure 1B and Supplemental Table 2). Similarly, 336 and 2097 genes that were related to the UV-B response were obtained in *O. sativa* and *F. tataricum*, respectively (Figure 1C and Supplemental Table 2). We examined an extant public dataset of

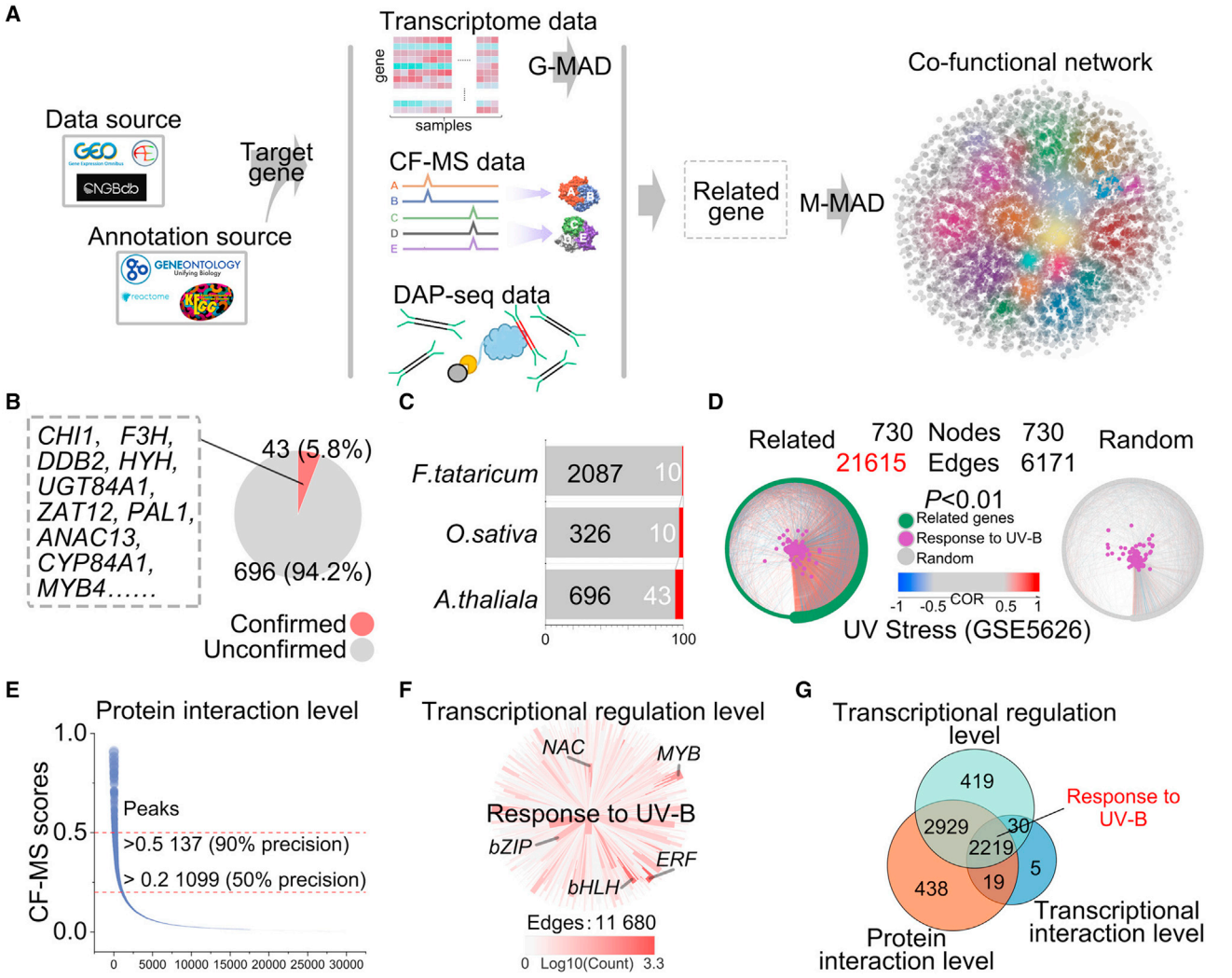

Figure 1. Co-functional network of the UV-B response.

(A) Architecture of the workflow. Multi-species and multi-omics data were obtained from public databases. The transcript-level-related genes of the target were obtained using G-MAD, the protein-level-related genes of the target were obtained using CF-MS, and the DNA–protein-level-related genes of the target were obtained using DAP-seq. The related genes at each level were combined to obtain the related functions of the target using M-MAD.

(B) UV-B-response-related genes were obtained by G-MAD.

(C) UV-B-response-related genes were obtained by G-MAD in *A. thaliana*, *O. sativa*, and *F. tataricum*. Red indicates genes known to be associated with the UV-B response.

(D) Correlation network of the UV-B response and the related genes or random genes with the same number of related genes based on the UV-B stress dataset (GSE5626). The complexity of the network (number of edges) and the degree of correlation (color of edges) were compared ($P < 0.01$). Nodes indicate the number of input genes.

(E) UV-B-response-related genes were obtained based on CF-MS data.

(F) UV-B-response-related TFs were obtained based on DAP-seq data.

(G) Venn diagram of functions at the transcript level, protein level, and DNA–protein level based on GO enrichment.

UV-B stress (GSE5626) and confirmed by correlation network analysis that the expression of these genes was closely correlated with the UV-B response network, in contrast to the absence of such a correlation in a random gene set (Figure 1D and Supplemental Table 3).

Because the related genes above were obtained based on the transcript level, we next screened for genes that may be related to the UV-B response at the level of protein interactions and transcriptional regulation. Public data from co-fractionation mass spectrometry (CF-MS), a high-throughput method for detecting interacting proteins (McWhite et al., 2020), enabled the identification of 3016 proteins that interacted with the UV-B response pathway proteins (Figure 1E and Supplemental Table 4). Public data from DNA affinity purification sequencing (DAP-seq), a method for high-throughput detection of transcriptional regulation *in vitro* (Bartlett et al., 2017), enabled the identification of 3120 transcription factors regulating UV-B response pathway genes (Figure 1F and Supplemental Table 5). We performed GO enrichment analysis of the related genes obtained at the transcript, protein interaction, and transcriptional regulation levels and found that 2219 GO terms shared by the three levels

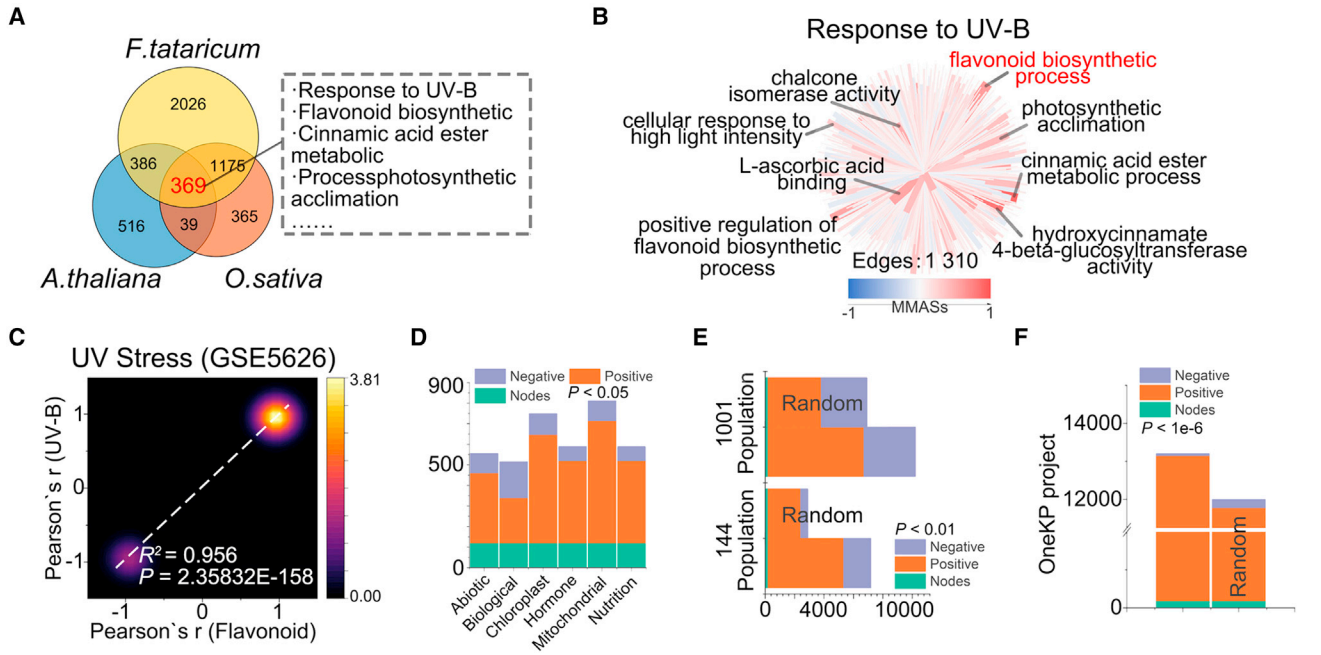

Figure 2. Comprehensive analysis of the co-functional network of the UV-B response.

(A) Venn diagram of the UV-B-response-related functions in *A. thaliana*, *O. sativa*, and *F. tataricum* based on M-MAD.

(B) Analysis of the UV-B-response-related top functions based on M-MAD. The color depth indicates the MMASs of the co-functional network of the UV-B response.

(C) The correlation density map between the UV-B response and flavonoid biosynthesis based on the UV-B stress dataset (GSE5626). The column represents the correlation between the co-functional network and the UV-B response, and the row represents the correlation between the co-functional network and flavonoid biosynthesis.

(D) Microenvironment analysis of the co-functional network of flavonoid biosynthesis and UV-B response ($P < 0.05$). Nodes indicate the number of input genes.

(E) Correlation network of the co-functional network of flavonoid biosynthesis and UV-B response based on 144 natural epigenetic variant populations and the 1001 Genomes Project ($P < 0.01$). Nodes indicate the number of input genes.

(F) Correlation network of the co-functional network of flavonoid biosynthesis and UV-B response based on the OneKP project ($P < 1e-6$). Nodes indicate the number of input genes.

included UV-B response, cellular response to high light intensity, cinnamic acid ester metabolic process, and photosynthetic adaptation, demonstrating that these genes are closely related to the UV-B response at different levels (Figure 1G and Supplemental Table 6). We then used the M-MAD tool in the GeneBridge toolkit to infer linkages between genes that were closely related to the UV-B response and other biological functions and to obtain module-module association scores (MMASs). Thus, co-functional networks of the UV-B response consisting of 1310, 1948, and 3956 GO terms were identified in *A. thaliana*, *O. sativa*, and *F. tataricum*, respectively, using functions as nodes and MMASs as edges (Supplemental Table 7).

Cross-species conserved functions, robustness, and microenvironment analysis of the co-functional network of the UV-B response

To explore the conserved functions of the UV-B response among species, we retrieved M-MAD results and performed an overlap analysis of the three species results. The results revealed that 389 functions, such as hydroxycinnamate 4-beta-glucosyltransferase activity, chalcone isomerase activity, and photosynthetic acclimation, were shared among species (Figure 2A and Supplemental Table 7). Flavonoid biosynthesis had the highest association with the UV-B response (Figure 2B and Supplemental Table 7). Flavonoids are known to play an important role in plant resistance to UV-B stress (Wan et al., 2018; Shi and Liu, 2021; Yoon et al., 2021). Correlation network analysis using the UV-B stress dataset (GSE5626) confirmed that the genes obtained by G-MAD correlate flavonoid biosynthesis with the UV-B response (Figure 2C and Supplemental Table 8).

It is known that environmental stresses cause species to form specific functional networks, which are significantly activated by the corresponding stresses (He et al., 2020; Trivedi et al., 2020). We attempted to understand the environmental stress context in which the co-functional network of flavonoid biosynthesis and UV-B response has formed. In stressful microenvironments (including nutrient, mitochondrial, hormonal, chloroplast, biotic, and abiotic stresses) caused by chemical or genetic strategies or by environmental stressors of a biotic and abiotic nature, correlation network analysis revealed that, in addition to biotic stress, the co-functional network of flavonoid biosynthesis and UV-B response can form a close association. This finding suggests that the network can respond synergistically to multiple types of stresses induced by UV-B, consistent with the multiple biological functions of flavonoids (Wan et al., 2018; Shi and Liu, 2021; Yoon et al., 2021) (Figure 2D and Supplemental Table 9).

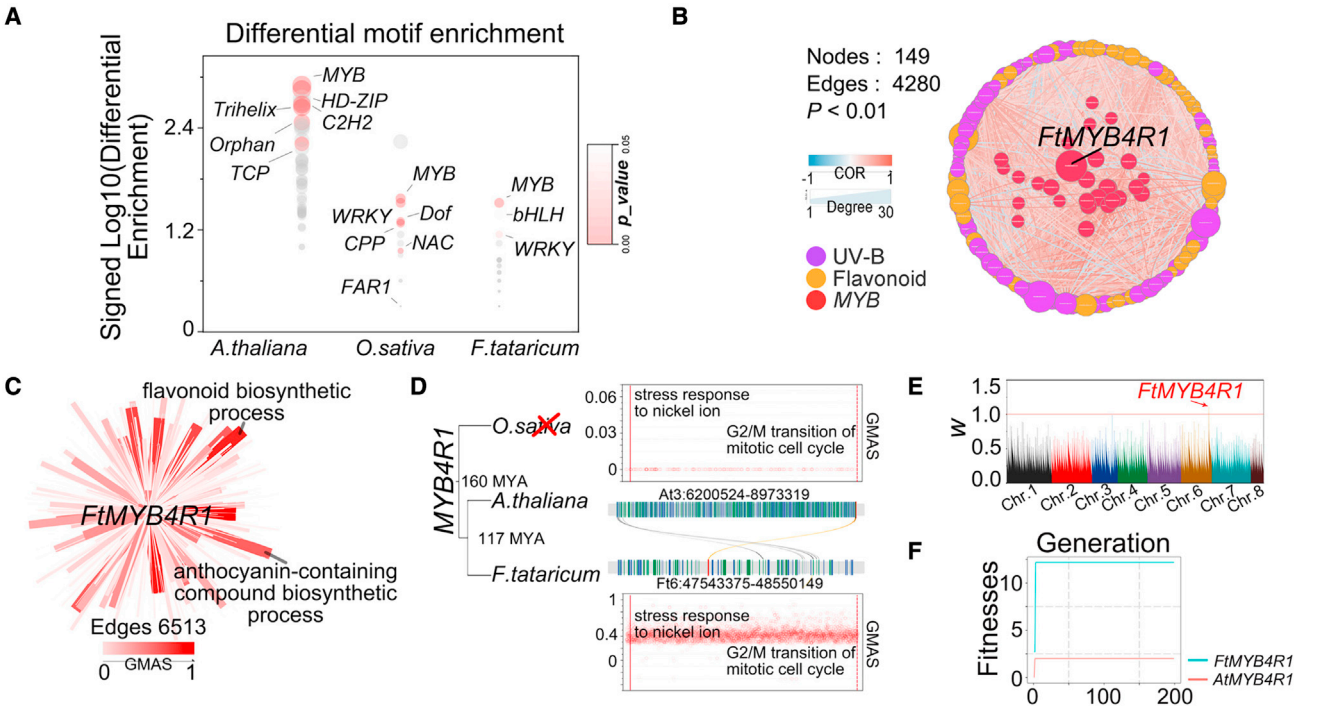

Figure 3. Neofunctionalization of the candidate regulators of the co-functional network of UV-B response.

(A) Enrichment of the TF motifs in each species (*A. thaliana*, *O. sativa*, and *F. tataricum*) identifies candidate regulators of the co-functional network of the UV-B response. The symbol size indicates the degree of enrichment, and the color depth indicates the *P* value.

(B) Identification of core regulators of the co-functional network of the UV-B response in *F. tataricum*. The color depth indicates the degree of correlation. The symbol size represents the degree in the network: the higher the degree, the better the connectivity in the network ($P < 0.01$).

(C) *FtMYB4R1*-related functions in *F. tataricum* based on G-MAD. The color depth indicates the degree of correlation.

(D) Schematic diagram of *MYB4R1* neofunctionalization.

(E) Selection pressure analysis of *MYB4R1* shows that *FtMYB4R1* has undergone positive selection.

(F) Fitness analysis based on a genetic algorithm (set to 1000 generations).

To demonstrate the conservation of the co-functional network of flavonoid biosynthesis and UV-B response in *A. thaliana* populations, we performed a correlation network analysis for the expression levels of UV-B response, flavonoid biosynthesis, and G-MAD-acquired genes in 144 natural epigenetic variant populations (Schmitz et al., 2013) and the 1001 Genomes Project (Kawakatsu et al., 2016). We found that they form a close association, in contrast to the absence of such a correlation in a random gene set (Figure 2E and Supplemental Table 10). Orthologs ($e = 1 \times 10^{-5}$) of the UV-B response, flavonoid biosynthesis, and G-MAD-acquired genes were also identified using the OneKP program (Matasci et al., 2014), which covered all green plants for more than 1 billion years, and correlation network analysis was performed using gene expression data to verify that this co-functional network was conserved in the plants (Figure 2F and Supplemental Table 11). In summary, the co-functional network of flavonoid biosynthesis and UV-B response in plants is capable of responding to a wide range of stresses.

Candidate regulators, neofunctionalization of core members, and natural selection fitness analysis of the co-functional network of the UV-B response

To identify regulators that may drive the co-functional network of the UV-B response, we performed TF-motif differential enrichment analysis of the UV-B response-related genes in each species (*A. thaliana*, *O. sativa*, and *F. tataricum*) using the Plant Transcriptional Regulatory Map database (Tian et al., 2020). The analysis revealed that the regulatory factors of the three species were differentiated, but the shared and top regulatory factors were MYBs (Figure 3A and Supplemental Table 12). This result is consistent with the important function of MYBs in the UV-B response (Wan et al., 2018; Shi and Liu, 2021; Yoon et al., 2021).

To identify the core members among the *F. tataricum* candidate regulators, we performed a correlation network analysis of MYB, flavonoid biosynthesis, and UV-B response genes and found that *FtMYB4R1* had the best connectivity (highest degree value) among the MYBs at the core of the network (Figure 3B and Supplemental Table 13).However, a function for *FtMYB4R1* has not yet been reported. Synteny analysis combined with PhyloGenes database results (Zhang et al., 2020) showed that *FtMYB4R1* has one ortholog in *A. thaliana* but none in *O. sativa* (Figure 3C and Supplemental Table 14). According to the G-MAD results, *MYB4R1* was found to have the top functions of response to UV-B, flavonoid biosynthetic process, and anthocyanin-containing compound biosynthetic process in *F. tataricum* (Figure 3C and Supplemental Table 14). By contrast, in *A. thaliana*, it has only a few functions in the G2/M transition of the mitotic cell cycle and stress response to nickel ions (Yanhui et al., 2006; Wang et al., 2008). We speculate that the high-UV-B habitat of *F. tataricum* prompted the neofunctionalization of its *MYB4R1* (Figure 3D and Supplemental Table 15).

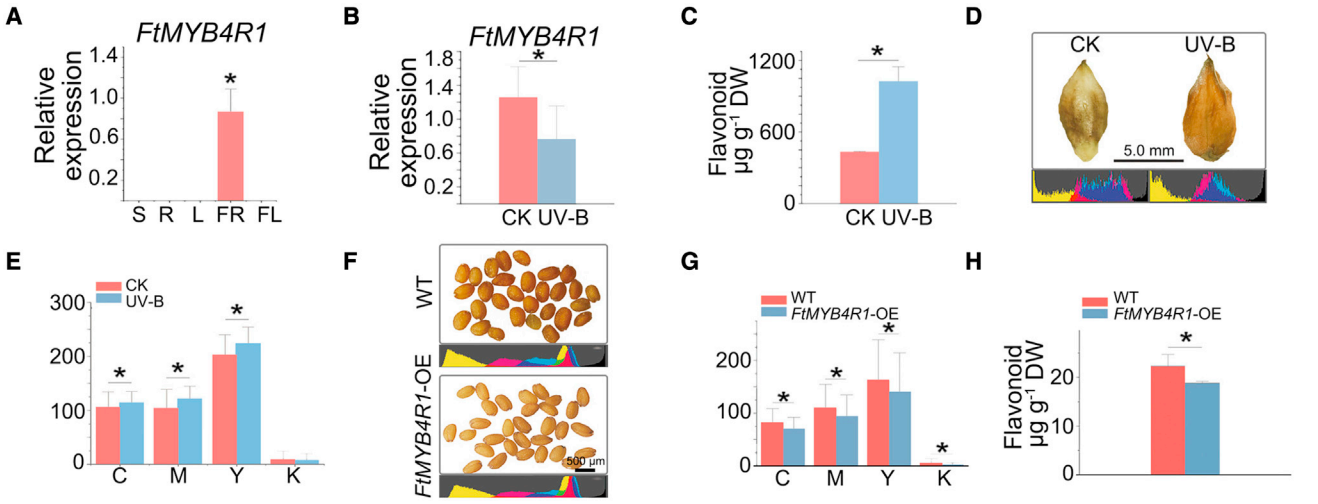

Figure 4. *FtMYB4R1* negatively regulates flavonoid biosynthesis and alters seed color under UV-B stress.

(A) Spatiotemporal analysis of *FtMYB4R1* expression by qRT–PCR. S, stem; R, root; L, leaf; FR, fruit; FL, flower. Data are shown as the mean ± standard deviation (SD), n = 3 (*$P < 0.05$, LSD).

(B) *FtMYB4R1* expression by qRT–PCR under UV-B stress. UV-B radiation (Philips TL-K 20 W/12 RS, 0.2 mW cm^{-2}) was applied for 2 h. Data are shown as the mean ± SD, n = 3 (*$P < 0.05$, LSD).

(C) Flavonoid content of buckwheat seeds under UV-B stress (μg g^{-1} dry weight [DW]). Data are shown as the mean ± SD, n = 10 (*$P < 0.05$, LSD).

(D) Phenotype of buckwheat seeds treated with UV-B. Below is the histogram (CMYK mode) corresponding to the buckwheat seed image above.

(E) The average values of C, M, Y, and K in the histogram of the buckwheat seed image after UV-B processing. C, cyan; M, magenta; Y, yellow; K, black. Data are shown as the mean ± SD, n = 3 (*$P < 0.05$, LSD).

(F) Phenotype of homozygous *A. thaliana* seeds obtained through stable overexpression of *FtMYB4R1*. Below is the histogram (CMYK mode) that corresponds to the *A. thaliana* seed image above.

(G) The average values of C, M, Y, and K in the histogram of the seed image from *A. thaliana* stably overexpressing *FtMYB4R1*. Data are shown as the mean ± SD, n = 3 (*$P < 0.05$, LSD).

(H) Flavonoid content (μg g^{-1} dry weight, DW) in the imaged seeds of *A. thaliana stably overexpressing FtMYB4R1*. Data are shown as the mean ± SD, n = 10 (*$P < 0.05$, LSD).

To assess this possibility, we analyzed the selection pressure on *MYB4R1* orthologs of different species and found that *MYB4R1* was positively selected and retained only in *F. tataricum* (Figure 3E and Supplemental Table 16). On the basis of Darwinian natural selection, a comparative analysis of the fitness of positively selected *FtMYB4R1* and *AtMYB4R1* gene sequences was performed using genetic algorithms (Mirjalili, 2019) (set to 1000 generations) (Supplemental Table 17). The results showed that *FtMYB4R1* reached the optimal solution with a higher fitness. This indicates that under natural selection conditions, *FtMYB4R1* can reach ecological adaptation in fewer generations, facilitating the long-term evolution of the species (Figure 3F). In summary, *FtMYB4R1*, which resides at the core of the co-functional network of UV-B responses, has undergone neofunctionalization in ecological adaptation.

FtMYB4R1 negatively regulates flavonoid content and seed color in response to UV-B

The G-MAD results indicated that the top function of *FtMYB4R1*, in addition to UV-B response, involves flavonoid biosynthesis and the biosynthetic process of anthocyanin-containing compounds (Figure 3C). Flavonoids are known to protect plants from UV-B damage (Wan et al., 2018; Shi and Liu, 2021; Yoon et al., 2021), and anthocyanins can produce different colors in plants (Silva et al., 2017; Alappat and Alappat, 2020). We first determined the spatiotemporal characteristics of *FtMYB4R1* expression and found that it was significantly highly expressed in the seeds (Figure 4A and Supplemental Table 20). We then exposed buckwheat to UV-B radiation (Philips TL-K 20 W/12 RS, 0.2 mW cm^{-2}) for 2 h. Analysis of gene expression levels revealed a significant decrease in *FtMYB4R1* under UV-B treatment (Figure 4B and Supplemental Table 20); a significant increase in seed flavonoid content (Figure 4C) was observed, together with a change in seed color (Figure 4D). To quantitatively analyze changes in seed color, the mean values of cyan (C), magenta (M), yellow (Y), and black (K) seed images were recorded separately using CMYK mode (Rose and Fulton, 2003; Mccue, 2009). CMY was found to increase significantly after UV-B treatment (Figure 4E). We next overexpressed *FtMYB4R1* driven by the CaMV35S promoter in *A. thaliana* and observed a change in seed color in T3-generation transgenic-positive lines (Figure 4F). CMYK values were all significantly reduced (Figure 4G), and there was a significant decrease in flavonoid content (Figure 4H). The above experiments demonstrate that the function of *FtMYB4R1* is consistent with that predicted by G-MAD, negatively regulating flavonoid content and seed color in response to UV-B.

FtMYB4R1 regulates flavonoid and anthocyanin biosynthesis by binding the L-box motif in the promoters of *FtCHS*, *FtFLS*, and *FtUFGT*

To resolve the regulatory mechanism of *FtMYB4R1*, we transiently overexpressed *FtMYB4R1* in buckwheat seeds and observed that the seed color was changed and the CMYK values were reduced, especially the most significant C values (Figure 5A). Expression analysis of key pathway genes revealed that genes encoding the common upstream enzymes of flavonoid and anthocyanin

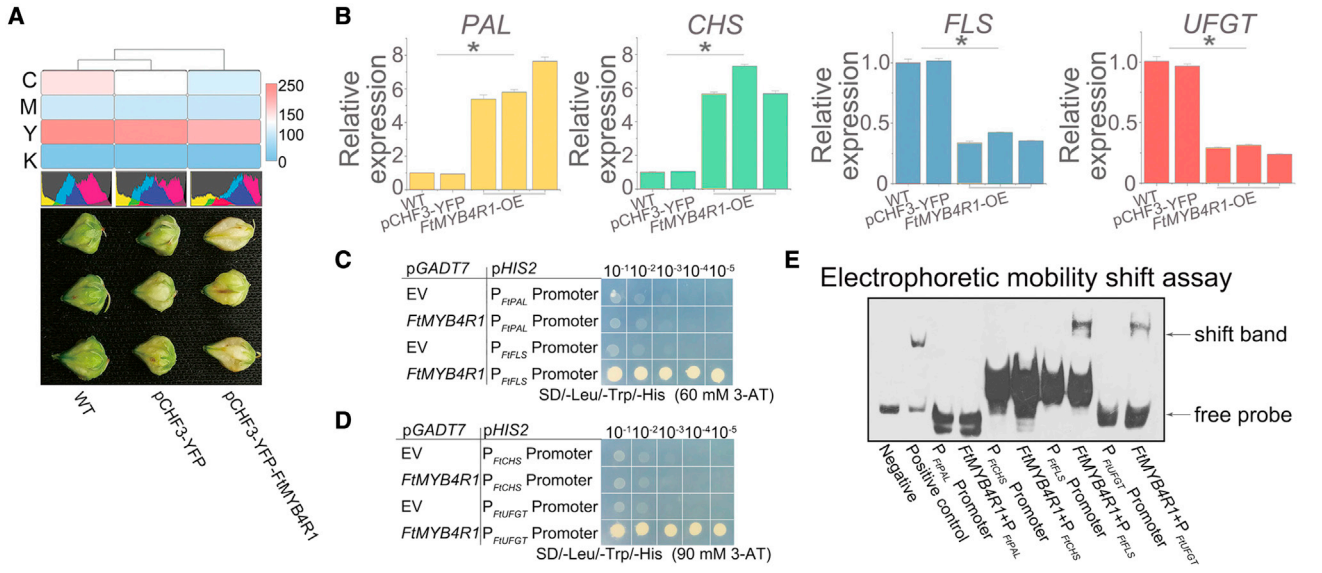

Figure 5. *FtMYB4R1* regulates flavonoid content and seed color by combining with the L-box motif in promoters of flavonoid and anthocyanin biosynthesis genes.

(A) Phenotype of *F. tataricum* seeds transiently overexpressing *FtMYB4R1*. Below is the histogram (CMYK mode) corresponding to the *F. tataricum* seed image above; the heatmap on the right shows the average values of C, M, Y, and K in the image histogram.

(B) The expression levels of key genes in flavonoid and anthocyanin biosynthesis during transient overexpression experiments with *FtMYB4R1*. Data are shown as the mean ± SD, n = 3 (*$P < 0.05$, LSD, compared with wild type [WT]).

(C) Y187 yeast cells transformed with different plasmid combinations (*FtMYB4R1* and P_{FtPAL} promoter; *FtMYB4R1* and P_{FtFLS} promoter) were diluted stepwise to 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5} and cultivated on SD/-Trp-Leu-His medium containing 60 mM 3-amino-1,2,4-triazole (3-AT).

(D) Y187 yeast cells transformed with different plasmid combinations (*FtMYB4R1* and P_{FtCHS} promoter; *FtMYB4R1* and P_{FtUFGT} promoter) were diluted stepwise to 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5} and cultivated on SD/-Trp-Leu-His medium containing 90 mM 3-AT.

(E) Binding reactions between *FtMYB4R1* and the *FtPAL*, *FtCHS*, *FtFLS*, and *FtUFGT* promoters were determined by EMSA.

biosynthesis, *FtPAL* and *FtCHS*, were upregulated, whereas genes encoding the respective downstream enzymes of these pathways, *FtFLS* and *FtUFGT*, were downregulated (Figure 5B and Supplemental Table 20).

On the basis of these results, the direct effect of *FtMYB4R1* on flavonoid and anthocyanin biosynthesis was investigated by yeast one-hybrid experiments. MYBs can regulate biosynthetic pathways by binding to GTTAG and L-box (ACTTTG) sequences in gene promoters (Espley et al., 2009; Yuan et al., 2013). We found that the *FtPAL*, *FtCHS*, *FtFLS*, and *FtUFGT* promoters all contained GTTAG sequences, and all promotors except that of *FtPAL* contained L-box sequences. *FtPAL* and *FtCHS* promoter fragments containing GTTAG sequences were cloned into the p*HIS2* vector; *FtFLS* and *FtUFGT* promoter fragments containing L-box sequences were also cloned into the p*HIS2* vector; and *FtMYB4R1* was cloned into the p*GADT7* vector. The results showed that *FtMYB4R1* could bind to the core sequences (L box) of the *FtFLS* and *FtUFGT* promoters but not to the core sequences (GTTAG) of other promoters (Figures 5C and 5D and Supplemental Figure 1). In yeast one-hybrid experiments, *FtMYB4R1* did not bind to *FtCHS*, probably because the fragment of the *FtCHS* promoter that contained the L-box motif was not used in these experiments.

To verify the direct interaction of *FtMYB4R1* with the gene promoters *in vitro*, we performed an electrophoretic mobility shift assay (EMSA). The results showed that *FtMYB4R1* promoted a shift in the band of biotin-labeled *FtFLS* and *FtUFGT* promoter DNA fragments that contained the L-box motif, and the intensity of the shifted band decreased with the addition of excess unlabeled wild-type fragments (Figure 5E). By contrast, *FtMYB4R1* did not promote a shift in the band of biotin-labeled *FtPAL* and *FtCHS* promoter DNA fragments containing the GTTAG motif (Figure 5E). In summary, *FtMYB4R1* regulates flavonoids and anthocyanins by binding to L-box motifs in the *FtCHS*, *FtFLS*, and *FtUFGT* promoters.

Visual recognition model of buckwheat flavonoid content developed based on population analysis under global UV-B radiation

To clarify that the neofunctionalization of *FtMYB4R1* is an evolutionarily driven result of UV-B adaptation in buckwheat, we assembled a natural buckwheat population, including 26 accessions from low-radiation regions (LR; 949–1241 kMh m^{-2}), 41 from medium-radiation regions (MR; 1387–1680 kMh m^{-2}), and 28 from high-radiation regions (HR; 1826–2118 kMh m^{-2}), which were identified based on the UV-B radiation budget and temporal and spatial variations across China from 2005 to 2012 (Liu et al., 2017) (Figure 6A and Supplemental Table 18). Measurements of seed flavonoid content and *FtMYB4R1* expression showed that buckwheat flavonoid content increased with increasing UV-B radiation level, whereas *FtMYB4R1* expression level decreased (Figure 6B and Supplemental Table 18).

In the above study, *FtMYB4R1* was shown to change seed color while regulating flavonoid accumulation. We hypothesized that seed color could be used to characterize the flavonoid content

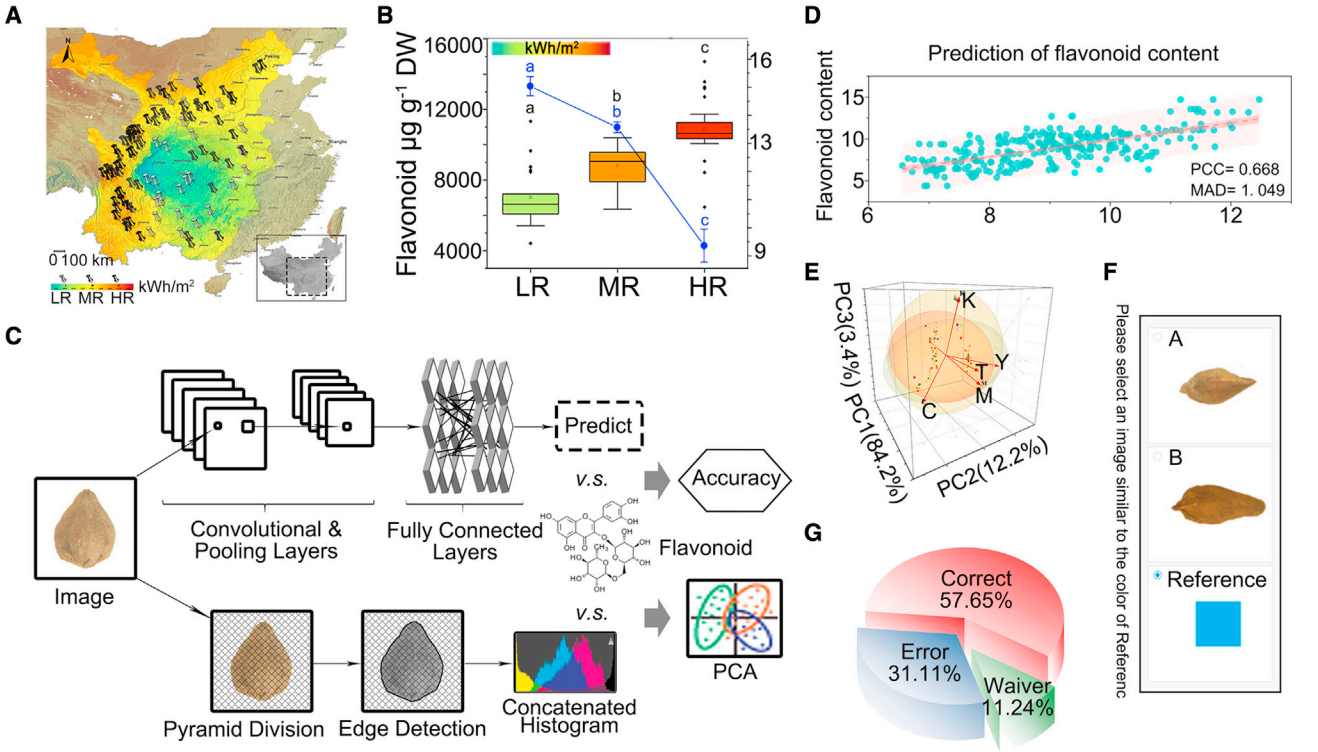

Figure 6. Visual recognition model of flavonoid content in buckwheat.

(A) Geographic distribution of 95 Tartary buckwheat accessions in regions of China with different radiation conditions, including 26 accessions from low-radiation regions (LR) (949–1241 kMh m^{-2}), 41 accessions from moderate-radiation regions (MR) (1387–1680 kMh m^{-2}), and 28 accessions from high-radiation regions (HR) (1826–2118 kMh m^{-2}). The color in the map indicates the radiation intensity (949–2118 kMh m^{-2}). Bars, 100 km.

(B) Boxplot representing the flavonoid content in mature seeds of 95 Tartary buckwheat accessions. Lowercase letters indicate significant differences between treatments ($\alpha = 0.05$, LSD). The color in the map indicates the radiation intensity (1095–1972 kMh m^{-2}). The blue line chart shows the expression level of *FtMYB4R1* in mature seeds of 95 Tartary buckwheat accessions measured by qRT–PCR.

(C) Schematic diagram of the flavonoid content estimation model that was trained based on convolutional neural networks containing five convolutional layers and three fully connected layers.

(D) Correlation between predicted flavonoid content and measured rutin content (MAD, 1.049; PCC, 0.668).

(E) Principal component analysis was performed using the image CMYK values as input values and the flavonoid contents as observation values, with regions of different radiation intensity as groups. Principal components 1, 2, and 3 are shown.

(F) Schematic diagram of volunteer participation in the survey. The volunteers selected the image close to the reference color in A and B. After choosing, the right or wrong answer was judged based on the flavonoid content of the seeds. In this diagram, A has a higher flavonoid content.

(G) Statistics of the results from the 150 volunteers who participated in the survey.

in natural buckwheat populations. For this purpose, we imported 285 seed images of the buckwheat natural population and the corresponding flavonoid content data into a neural network (AlexNet) (Wang et al., 2019) for training (Figure 6C) and finally obtained a prediction model with a mean deviation of ±1.049 mg g^{-1} (Median Absolute Deviation (MAD), 1.049; Person's Correlation Coefficient (PCC), 0.668) (Figure 6D). Principal component analysis was performed using the CMYK images as input values and flavonoid contents as observation values, with regions of different radiation intensities as groups. The results showed that PC1 (84.2%) exceeded 80%, and C correlated most strongly with PC1 (factor loading value, 0.773); therefore, C could characterize the flavonoid content of the buckwheat seeds (Figure 6E and Supplemental Table 19).

To verify the usefulness of this result, 150 visually healthy volunteers were invited to participate in a study in which buckwheat seed images that were similar to a color block were selected (C, 255) (10 parallel experiments per volunteer) (Figure 6F). We then judged whether their selections were correct based on the measured flavonoid content. We found that 57.65% of the buckwheat varieties with high flavonoid contents could be selected directly by visual inspection (Figure 6G). In summary, *FtMYB4R1* regulates flavonoid and anthocyanin accumulation as a result of UV-B-driven adaptive evolution, and, based on this finding, a visual recognition model of buckwheat flavonoid content was constructed.

DISCUSSION

The damage caused by UV-B radiation to plants is mainly due to denaturation of nucleic acids, proteins, and other macromolecules (Balskus and Walsh, 2010; Casati et al., 2011); destruction of membrane structure, and changes in membrane permeability, resulting in cell autolysis; and reduction in chlorophyll content (Booij-James et al., 2000). Thus, the efficiency of photosynthesis is affected, which further affects respiration, secondary metabolism, growth and development,

and other biological processes (Wan et al., 2018; Shi and Liu, 2021; Yoon et al., 2021). An important question is whether plants can evolve rapidly to adapt to environmental change; if not, migration or extinction may be the only options during times of environmental change (Davis et al., 2005). Plants have established multiple strategies to cope with the adverse effects of UV-B radiation; these include increasing the number of branches and tillers (Hamid et al., 2020), synthesizing UV-absorbing substances (Tohge et al., 2011; Kaling et al., 2015; Tohge and Fernie, 2017), and altering hormone homeostasis (Vanhaelewyn et al., 2016; Yang et al., 2020; Miao et al., 2021). Our studies show that flavonoid biosynthesis is most highly associated with UV-B responses and can synergistically respond to multiple types of UV-B-induced stress. Flavonoids are a class of secondary metabolites that are widely present in plants and play an important role in flower color, resistance to UV radiation, and resistance to pathogenic bacterial infection (Saslowsky et al., 2005). We observed that the co-functional network of flavonoid biosynthesis and UV-B response is conserved across a wide range of plant populations. However, how plants synergistically respond to other stresses requires further study.

Previous studies have shown that functional changes in well-connected genes that are centrally located in genetic networks tend to produce strains with high phenotypic diversity during evolution and can even enhance the evolutionary and adaptive capacity of organisms (Li et al., 1993). Our analysis of the regulators of the co-functional UV-B response network revealed that MYB transcription factors are significantly conserved regulators across species. MYBs play a key role in plant metabolism, development, and stress response and are an important target for the evolution of crop adaptation and breeding (Dubos et al., 2010). The buckwheat coregulator *MYB4R1* is under positive selection and can rapidly and efficiently promote ecological adaptation. Unlike its orthologs in *A. thaliana*, the neofunctionalized *FtMYB4R1* can regulate flavonoid and anthocyanin biosynthesis in response to UV-B by binding to L-box motifs in the *FtCHS*, *FtFLS*, and *FtUFGT* promoters. The first MYB repressor was discovered 20 years ago (Tamagnone et al., 1998), and research on MYB repressors has increased greatly in recent years (Chen et al., 2019). In buckwheat, *FtMYB16* is a flavonoid repressor that depends on the EAR motif for specific expression in roots (Zhang et al., 2018). We found that *FtMYB4R1* is a flavonoid and anthocyanin repressor that depends on the L-box motif for specific expression in seeds. However, numerous studies are needed to fully characterize this response in buckwheat and to assess whether it involves the typical *HY5* and *UVR8* pathways.

Deep learning methods have been widely used in recent years, especially in the field of intelligent agriculture for tasks such as pest detection, crop and weed detection and classification, and plant identification (Singh et al., 2016). The prominent advantage of deep learning is feature learning; that is, the combination of lower-level features to form higher-level features (LeCun et al., 2015). This technique can improve the accuracy of detection, classification, and recognition; for example, after several iterative trainings in disease recognition, the recognition accuracy can reach more than 90% (Fuentes et al., 2021). Second, deep learning involves versatility and generalization. For example, deep learning exhibits good anti-interference and robustness in pathological image recognition for fruit and vegetables (Yan et al., 2021). In addition, deep learning recognition is efficient; e.g., for the early detection of grape diseases (Nguyen et al., 2021). Based on the core regulator *FtMYB4R1*, which regulates both flavonoid and anthocyanin biosynthesis, we developed a visual identification model of buckwheat flavonoid content using deep learning to help consumers simply and quickly select the right buckwheat variety. However, the model still requires further optimization to improve its recognition accuracy.

Concluding remarks

We comprehensively applied several technical approaches, including computational biology, molecular biology, and artificial intelligence (AI), to dissect the co-functional network of the UV-B response and reveal the key role of neofunctionalization in adaptive evolution. Thus, we envision that integrated AI-based strategies can assess functional innovations over long timescales and help us decipher complex evolutionary phenomena at an unprecedented depth.

METHODS

Plant materials, growth conditions, and UV-B treatments

A total of 95 Tartary buckwheat accessions were collected in areas of China with different UV-B radiation levels, and the flavonoid content was determined in mature seeds from each accession. These accessions included 16 wild, 13 breeding, and 66 endemic accessions (Supplemental Table 18). Xiqiao no. 2 Tartary buckwheat was grown in the experimental field. We collected samples from three replicates of Xiqiao no. 2 and determined the transcript levels of *FtMYB4R1* in different tissues. Entire plants, roots, flowers, stems, leaves, and seeds were collected separately and stored at −80°C until further use. Xiqiao no. 2 plants were treated with UV-B light for 2 h in the presence of adequate moisture using a UV-B lamp (Philips TL-K 20 W/12 RS, 0.2 mW cm^{-2}).

Data sources and sequence retrieval

The public transcriptome data originated from research on the responses of *A. thaliana* to stimuli including nutrition stress (Van Aken and Whelan, 2012), mitochondrial stresses (Van Aken and Whelan, 2012), hormone stress (auxin, gibberellin, abscisic acid, ethylene, methyl jasmonate, brassinosteroid, zeatin, and salicylic acid; Goda et al., 2008; Clifton et al., 2005), chloroplast stresses (Van Aken and Whelan, 2012), biological stresses (*Botrytis cinerea*, Goda et al., 2008; elicitor Flg22, Goda et al., 2008; *Erysiphe orontii*, Goda et al., 2008; *Phytophthora infestans*, Goda et al., 2008; *Blumeria patens*, Jensen et al., 2008; elicitor EF-Tu, Zipfel et al., 2006; and *Erysiphe cichoracearum*, Nishimura et al., 2003), and the abiotic stresses ozone (Short et al., 2012), high light (Kleine et al., 2007), H_2O_2 (Clifton et al., 2005), heat (Goda et al., 2008), salt (Goda et al., 2008), UV (Goda et al., 2008), oxidative stress (Goda et al., 2008), and osmotic stress (Goda et al., 2008). DAP-seq datasets and CF-MS of *A. thaliana* were collected. Transcriptome datasets of 144 natural *A. thaliana* accessions (GEO: GSE43858), the *A. thaliana* 1001 genomes project (GEO: GSE80744), and the OneKP project datasets were also collected (Supplemental Table 1).

GeneBridge analysis

The data were preprocessed with the PEER program to remove potential covariables (Stegle et al., 2012) that may affect the analysis and then imported into the GeneBridge toolkit. The term "module" is used in the GeneBridge toolkit to refer to knowledge-based gene sets, ontology terms, and biological pathways from different resources; only ontological terms were used in this study. Gene-Module Association Determination (G-MAD) in the GeneBridge toolkit uses expression data from large-scale cohorts to propose potential gene functions. G-MAD uses a

competitive gene set testing method, in which the correlation adjusted mean rank gene set test (CAMERA) adjusts the correlations between genes (Wu and Smyth, 2012), to calculate the connections between genes of interest and biological modules. Gene-module connections that survived multiple test corrections for gene or module numbering with enrichment *P* values were assigned a connection score of 1 or −1 depending on the direction of the enrichment and were assigned 0 otherwise. The results were then subjected to meta-analysis across datasets, and gene-module association scores (GMASs) were calculated as the mean of the linkage scores weighted by sample size and intergenic correlation coefficients within modules ($\overline{p}$). Another tool, Module-Module Association Determination (M-MAD), uses the compilation of expression data to study the connections between modules. The results obtained from G-MAD for individual modules for all genes were used to calculate their associations with all modules. Using CAMERA (Wu and Smyth, 2012), the enrichment scores of all genes in the target modules were used as gene-level statistics to calculate the enrichment against all modules. The obtained cross-module condensed *P* values were converted to 1, 0, or −1 based on Bonferroni thresholds, and a meta-analysis was then performed on all datasets to obtain MMASs. Whereas the module similarity network was based only on existing gene annotations, the module association network relied on the analysis of complete expression datasets. Thus, it could reveal novel biological connections between modules that were not included in the literature-based annotations.

GO enrichment analysis

GO enrichment was performed using the GOseq R package (release 3.15). GO terms with a corrected *P* value <0.05 were considered to be significantly enriched. The GO annotations were functionally classified using WEGO software for gene function distributions.

Identification of *MYB4R1* synteny

MYB4R1 synteny analysis was performed using the default parameters of BLAST and MCScanX as described previously (Wang et al., 2012). The potential anchoring points (E-value < 1×10^{-5}; top five matches) between each possible pair of chromosomes in multiple genomes were identified using BLASTP. Protein sequences were searched against the given genome or genomes of other species. The homologous blocks in each genome or between different genomes were determined with ColinearScan (maximal gap ≤50 genes; *P* value <0.05) (Wang et al., 2006). The maximum gap between neighboring genes along a chromosome that exhibited collinearity with genes along the corresponding chromosome sequence was set to 50 intervening genes (Wang et al., 2017).

Measuring gene expression by qRT–PCR

The expression pattern of *FtMYB4R1* in four tissues was verified by qRT–PCR. The transcript levels of *FtMYB4R1* in mature seeds from 95 Tartary buckwheat accessions were also verified by qRT–PCR. The qRT–PCR conditions consisted of pre-denaturation at 95°C for 3 min, followed by 40 cycles of denaturation at 95°C for 5 s and annealing and extension at 60°C for 30 s. The primers used in the experiment were designed with Primer 3 (http://frodo.wi.mit.edu/) (Supplemental Table 20). Histone H3 was chosen as the internal reference gene. The $2^{-\Delta\Delta Ct}$ method was used to calculate the expression data. For each biological replicate, three technical replicates were performed for a total of three biological replicates.

UV radiation data

UV radiation data were obtained from the national-scale network for the measurement of solar radiation levels and temporal and spatial variations over China. These data included UV radiation data recorded hourly by 40 stations of the China Ecosystem Research Network from 2005 to 2012 (Supplemental Table 18). UV radiation was measured using a CUV3 broadband radiometer with an accuracy of 5% in accordance with the measurement standards of the World Meteorological Organization (WMO) (Liu et al., 2017).

Image analysis

Under stable environmental light conditions, images of seeds were obtained with an Olympus SZ61 stereomicroscope (Tokyo, Japan) in manual mode (ISO200, 1/196 s, *f*/2) using three biological replicates. Adobe Photoshop CC 2018 (California, USA) was used to select the seeds from the images. The histogram was read in CMYK mode, and the average C, M, Y, K, bright (B), and T values were recorded for each image.

Flavonoid extraction and high-performance liquid chromatography

The flavonoids in Tartary buckwheat consist mainly of rutin (Zhang et al., 2017; Liu et al., 2018). To extract rutin, 0.01 g of seed powder was mixed with 1 ml of methanol and incubated at 60°C for 2 h in an ultrasonic bath, followed by centrifugation at 4000 rpm for 10 min and collection of the supernatant. The solution was then filtered using a 0.22-μm membrane filter. An Agilent C18 column on an Agilent 1260 system was used for high-performance liquid chromatography (HPLC). The HPLC conditions were as follows: column temperature, 30°C; mobile phase, methanol with 0.5% phosphoric acid at a 40:60 volume ratio; flow rate, 1 ml min^{-1}; and injection volume, 10 μl. The rutin content was estimated based on the calibration curve of standard rutin hydrate (98%) at a detection wavelength of 257 nm (Supplemental Table 18). Three biological replicates were performed for each sample.

Principal component analysis

We performed principal component analysis (PCA) on 95 accessions. The input data included the C, M, K, Y, B, and T values of each accession; the observations were the rutin contents of each accession; the groups were based on the level of radiation in the region (HR, MR, or LR); the method was correlation matrix analysis; and the number of components to be extracted was three (Supplemental Table 19).

Recognition of flavonoid content in Tartary buckwheat seeds based on the AlexNet model

The system was based on the cloud computing platform and the AI platform of Shanghai Jiao Tong University. The cloud computing platform was configured with 12 320 core central processing unit (CPU)s, 118 TB of memory, and 12 Pb of storage. The AI platform was configured with eight DGX-2 servers, 128 v100 CPUs, and 16 peta floating point operations per second (PFLOPS) of computing power. An AlexNet that contained five convolutional layers (CLs) and three fully connected layers (FCLs) was used as the basic transfer learning model (Supplemental Figure 2). To improve the model performance, data augmentation (DA) was first applied to the input images as follows: image rotation (90°, 180°, 270°). Because the original FCLs were developed to classify 1000 categories, AlexNet cannot be applied directly as a feature extractor. Therefore, in this study, the loss function and classification layer (Supplemental Table 21) were modified using a new randomly initialized fully connected layer with only one class. The learnable weights and biases of the improved AlexNet are shown in Supplemental Table 22, and the total weights and biases are 56 862 752 + 9569 = 56 872 321. The number of training epochs in this study was set to 15, the global learning rate was set to 10^{-4}, the batch size was 64, and the activation function was rectified linear units (ReLU). The learning rate of the new layers was 10 times that of the transferred layer because the transferred layers with pretrained weights/biases and new layers had random-initialized weights/biases. The training-validation-test division was set at random again, and the training procedure stopped when either the algorithm reached the maximum epoch or the validation performance decreased over a preset training epoch. The hyperparameters were iteratively tuned, and those optimal hyperparameters were found based on the validation set. After the hyperparameters were fixed, we ran the final model on the test set for 10 runs.

Generation of *FtMYB4R1* transgenic *A. thaliana*

The open reading frame (ORF) of *FtMYB4R1* was cloned into the *Kpn*I and *Bam*HI restriction sites of the binary vector pCAMBIA1301, which contained the cauliflower mosaic virus (CaMV) 35S promoter, the nopaline synthase (NOS) terminator, and the *hygromycin* gene. *A. thaliana* ecotype Columbia-0 plants transformed through the floral dip method with *Agrobacterium* strain GV3101 harboring the recombinant pCAMBIA1301-*FtMYB4R1* plasmids were grown in soil. The transformed plants were germinated on 1/2 Murashige and Skoog medium containing 25 mg l^{-1} *hygromycin* under the same growth conditions. Positive homozygotes were screened by PCR and GUS staining.

Yeast one-hybrid assays

The full-length ORF of *FtMYB4R1* was cloned into the pGADT7 vector using the *Eco*RI and *Xba*I restriction enzymes, and the core segments of the *FtPAL*, *FtCHS*, *FtFLS*, and *FtUFGT* promoters were cloned into the p*HIS2* vector using the *Eco*RI and *Bam*HI restriction enzymes. Plasmid pGADT7-*FtMYB4R1* was cotransformed with p*HIS2-FtPAL*, *-FtCHS*, *-FtFLS*, or *-FtUFGT* into yeast strain *Y187*, and the corresponding negative controls were constructed. The p*HIS2-FtPAL*-, *-FtCHS*-, *-FtFLS*-, and *-FtUFGT*-transformed yeast colonies were cultured on SD plates lacking Leu, Trp, and His and containing different concentrations (10, 20, 30, 60, and 90 mM) of 3-AT. The transformed yeast colonies (p*HIS2-FtPAL* and p*HIS2-FtFLS*) were inhibited on SD/-Leu-Trp-His plates containing 60 mM 3-AT, whereas the transformed yeast colonies (p*HIS2-FtCHS* and p*HIS2-FtUFGT*) were inhibited on SD/-Leu-Trp-His plates containing 90 mM 3-AT. The transformed yeast colonies were cultured in yeast peptone dextrose adenine (YPDA) medium to a cell optical density at 600 nm (OD_{600}) of approximately 0.6 and then diluted to 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5} with sterilized double-distilled water. Then, 2 μl of diluted yeast cell solution was transferred to SD/-Trp-Leu plates. In addition, 2 μl of diluted yeast cell solution (a combination of plasmids pGADT7 and p*HIS2-FtPAL*, p*GADT7* and p*HIS2-FtFLS*, p*GADT7-FtMYB4R1* and p*HIS2-FtPAL*, and p*GADT7-FtMYB4R1* and p*HIS2-FtFLS*) was transferred to SD/-Leu-Trp-His plates containing 60 mM 3-AT. Finally, 2 μl of diluted yeast cell solution (a combination of plasmids p*GADT7* and p*HIS2-FtCHS*, p*GADT7* and p*HIS2-FtUFGT*, p*GADT7-FtMYB4R1* and p*HIS2-FtCHS*, and p*GADT7-FtMYB4R1* and p*HIS2-FtUFGT*) was transferred to SD/-Leu-Trp-His plates containing 90 mM 3-AT. The cells were cultured at 30°C for 2–4 days, and colony growth was observed.

EMSA

The pET-28a-*FtMYB4R1* recombinant protein was expressed and purified from *Escherichia coli*. For EMSA, 5′-biotinylated oligonucleotides (5′-TTTAAAAGATGTTAGTTGCTCAAAC-3′; 5′-AAATATCACTGTTAGCGTAAGATTAAAAAAAAGGAAACTTGTTGGAACGTTTTGTTCTTATATGAACCATTATATCGAAAGTTAGAACTAATAAA-3′; 5′-TGTTTAAACAGACTTTGAAGACCTCTTTTTTCTATGTGACCAAGCCTCTTGTGTAGAAGCTTAGTTTGTTCTAGGGCATGGCGTTAGAGGGCTTTGTT-3′; and 5′-TGCGTTCTCCACTTTGCAGCTGTTGT-3′) were used as probes. After incubation with probes and nuclear extracts for 30 min at room temperature, the entire reaction mixture was placed on a nondenatured 0.5× Tris-borate-EDTA (TBE) 6% polyacrylamide gel, run for 1 h at 60 V and 4°C, and transferred to Biodyne B nylon membranes. Kits and the ChemiDoc XRS system were used to visualize the signals.

Transient transformation of Tartary buckwheat seeds with *FtMYB4R1*

The ORF of *FtMYB4R1* was cloned into the pCHF3-YFP vector. The recombinant plasmid was transformed into *Agrobacterium* strain GV3101 and incubated at 28°C for 2 days. *Agrobacterium* cultivation and infiltration preparation were performed according to previous studies. *Agrobacteria* containing the pCHF3-*FtMYB4R1*-YFP recombinant plasmid were resuspended in 10 ml of infiltration buffer containing 10 mM $MgCl_2$, 10 mM MES-KOH, and 100 μM acetosyringone and incubated at 21°C for 2 h. The mixed suspension was injected into Xiqiao no. 2 seeds. Two days after infiltration, transiently transformed, empty vector-transformed, and wild-type Tartary buckwheat seeds were photographed and observed. The expression levels of genes related to the rutin synthesis pathway were measured by qRT–PCR in the three groups of seeds.

Differential enrichment of TF binding sites

The TF enrichment tool from the Plant Transcriptional Regulatory Map database (http://plantregmap.gao-lab.org/tf_enrichment.php) was used to detect differential enrichment of TF binding sites (method: FunTFBS, threshold P value ≤ 0.05). The transcriptional regulation information used by the TF enrichment tool was obtained from the literature and from chromatin immunoprecipitation sequencing (ChIP-seq) data or inferred by combining TF binding motifs and regulatory element data. Fisher's exact test was performed to check whether a greater proportion of genes were targets of a TF.

Genetic algorithm analysis

Genetic algorithms (GAs) simulate the survival of fitter creatures and their genes. The GA uses the fitness (objective) function to assess the fitness of each individual in the population, and its main operators are selection, crossover, and mutation. These operators are applied to each generation to improve the quality of the next generation of genes. The best solution in the last population is returned as the best approximation to the global optimum of the given problem. The rates of selection, crossover, and mutation can be changed or set to fix the numbers during optimization. In this study, the mutation value of the GA was set to 0.2, the crossover value was set to 0.5, the selection value was set to 0.2, and the number of generations was 1000 (Supplemental Table 17).

Correlation network construction and network visualization

The correlation network was constructed using the R package imsbInfer available at GitHub (https://github.com/wolski/imsbInfer). Correlation was estimated using Pearson's rank correlation measure; red edges represent positive correlations, and blue edges represent negative correlations.

Statistical analysis

The experimental data were analyzed using Origin Pro 2021 software (OriginLab, Northampton, MA, USA). The least significant difference test (LSD) was used for mean separation at significance levels of 0.05 and 0.01.

Accession numbers

The data that support the findings of this study are available in the supplemental information.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

FUNDING

This research was sponsored by the National Key R&D Program of China (2021YFD1200105), the National Natural Science Foundation of China (3210150112), the Sichuan Province Science and Technology Support Program (2021YFH0086), the Shanghai Sailing Program (20YF1422000), and the Startup Fund for Youngman Research at Shanghai Jiao Tong University (20X100040052).

AUTHOR CONTRIBUTIONS

M.-Y.L. planned and designed the research. M.-Y.L. analyzed the data. W.-J.S. and M.-Y.L. wrote the original manuscript. W.-J.S. and Z.-T.M. determined the expression of genes by qRT–PCR. W.-J.S. performed a transgenic verification experiment. W.-J.S. and Z.-T.M. determined the rutin contents and seed color changes of the Tartary buckwheat population after exposure to UV light. C.-C.G. used the AI method to integrate the plant data. J.-H.C. developed the AlexNet model. M.-Y.L. reviewed and

edited the manuscript. M.-Y.L. and Q.W. revised the manuscript. H.C. and X.-Y.W. supervised the research. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We are grateful for the support of the cloud computing and AI platform of Shanghai Jiao Tong University.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. No conflict of interest is declared.

Received: April 22, 2022
Revised: June 20, 2022
Accepted: July 20, 2022
Published: August 2, 2022

REFERENCES

Alappat, B., and Alappat, J. (2020). Anthocyanin pigments: beyond aesthetics. Molecules **25**:5500. https://doi.org/10.3390/molecules25235500.

Ambawat, S., Sharma, P., Yadav, N.R., and Yadav, R.C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. Physiol. Mol. Biol. Plants **19**:307–321. https://doi.org/10.1007/s12298-013-0179-1.

Balskus, E.P., and Walsh, C.T. (2010). The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. Science **329**:1653–1656. https://doi.org/10.1126/science.1193637.

Bartlett, A., O'Malley, R.C., Huang, S.S.C., Galli, M., Nery, J.R., Gallavotti, A., and Ecker, J.R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. Nat. Protoc. **12**:1659–1672. https://doi.org/10.1038/nprot.2017.055.

Booij-James, I.S., Dube, S.K., Jansen, M.A., Edelman, M., and Mattoo, A.K. (2000). Ultraviolet-B radiation impacts light-mediated turnover of the photosystem II reaction center heterodimer in Arabidopsis mutants altered in phenolic metabolism. Plant Physiol. **124**:1275–1284. https://doi.org/10.1104/pp.124.3.1275.

Britt, A.B. (1995). Repair of DNA damage induced by ultraviolet radiation. Plant Physiol. **108**:891–896. https://doi.org/10.1104/pp.108.3.891.

Casati, P., Campi, M., Morrow, D.J., Fernandes, J.F., and Walbot, V. (2011). Transcriptomic, proteomic and metabolomic analysis of UV-B signaling in maize. BMC Genom. **12**:321. https://doi.org/10.1186/1471-2164-12-321.

Chen, L., Hu, B., Qin, Y., Hu, G., and Zhao, J. (2019). Advance of the negative regulation of anthocyanin biosynthesis by MYB transcription factors. Plant Physiol. Biochem. **136**:178–187. https://doi.org/10.1016/j.plaphy.2019.01.024.

Clifton, R., Lister, R., Parker, K.L., Sappl, P.G., Elhafez, D., Millar, A.H., Day, D.A., and Whelan, J. (2005). Stress-induced co-expression of alternative respiratory chain components in Arabidopsis thaliana. Plant Mol. Biol. **58**:193–212. https://doi.org/10.1007/s11103-005-5514-7.

Davis, M.B., Shaw, R.G., and Etterson, J.R. (2005). Evolutionary responses to climate change. Ecology **86**:1704–1714. https://doi.org/10.1890/03-0788.

Dong, H., Liu, X., Zhang, C., Guo, H., Liu, Y., Chen, H., Yin, R., and Lin, L. (2021). Expression of Tomato UVR8 in Arabidopsis reveals conserved photoreceptor function. Plant Sci. **303**, 110766. https://doi.org/10.1016/j.plantsci.2020.110766.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. Trends Plant Sci. **15**:573–581.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA **95**:14863–14868. https://doi.org/10.1073/pnas.95.25.14863.

Espley, R.V., Brendolise, C., Chagné, D., Kutty-Amma, S., Green, S., Volz, R., Putterill, J., Schouten, H.J., Gardiner, S.E., Hellens, R.P., et al. (2009). Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. Plant Cell **21**:168–183. https://doi.org/10.1105/tpc.108.059329.

Favory, J.J., Stec, A., Gruber, H., Rizzini, L., Oravecz, A., Funk, M., Albert, A., Cloix, C., Jenkins, G.I., Oakeley, E.J., et al. (2009). Interaction of COP1 and UVR8 regulates UV-B-induced photomorphogenesis and stress acclimation in Arabidopsis. EMBO J. **28**:591–601. https://doi.org/10.1038/emboj.2009.4.

Ferl, R.J., Schuerger, A.C., Paul, A.L., Gurley, W.B., Corey, K., and Bucklin, R. (2002). Plant adaptation to low atmospheric pressures: potential molecular responses. Life Support Biosph. Sci. **8**:93–101.

Fuentes, A., Yoon, S., Lee, M.H., and Park, D.S. (2021). Improving accuracy of tomato plant disease diagnosis based on deep learning with explicit control of hidden classes. Front. Plant Sci. **12**, 682230. https://doi.org/10.3389/fpls.2021.682230.

Goda, H., Sasaki, E., Akiyama, K., Maruyama-Nakashita, A., Nakabayashi, K., Li, W., Ogawa, M., Yamauchi, Y., Preston, J., Aoki, K., et al. (2008). The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. Plant J. **55**:526–542. https://doi.org/10.1111/j.0960-7412.2008.03510.x.

Hamid, A., Singh, S., Agrawal, M., and Agrawal, S.B. (2020). Effects of plant age on performance of the tropical perennial fodder grass, Cenchrus ciliaris L. subjected to elevated ultraviolet-B radiation. Plant Biol. **22**:805–812. https://doi.org/10.1111/plb.13116.

He, N., Li, Y., Liu, C., Xu, L., Li, M., Zhang, J., He, J., Tang, Z., Han, X., Ye, Q., et al. (2020). Plant trait networks: improved resolution of the dimensionality of adaptation. Trends Ecol. Evol. **35**:908–918. https://doi.org/10.1016/j.tree.2020.06.003.

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. Trends Plant Sci. **18**:477–483. https://doi.org/10.1016/j.tplants.2013.06.003.

Jansen, M.A., Gaba, V., and Greenberg, B.M. (1998). Higher plants and UV-B radiation: balancing damage, repair and acclimation. Trends Plant Sci. **3**:131–135. https://doi.org/10.1016/S1360-1385(98)01215-1.

Jensen, M.K., Hagedorn, P.H., de Torres-Zabala, M., Grant, M.R., Rung, J.H., Collinge, D.B., and Lyngkjaer, M.F. (2008). Transcriptional regulation by an NAC (NAM-ATAF1, 2-CUC2) transcription factor attenuates ABA signalling for efficient basal defence towards Blumeria graminis f. sp. hordei in Arabidopsis. Plant J. **56**:867–880. https://doi.org/10.1111/j.1365-313X.2008.03646.x.

Kaling, M., Kanawati, B., Ghirardo, A., Albert, A., Winkler, J.B., Heller, W., Barta, C., Loreto, F., Schmitt-Kopplin, P., and Schnitzler, J.P. (2015). UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics. Plant Cell Environ. **38**:892–904. https://doi.org/10.1111/pce.12348.

Kawakatsu, T., Huang, S.-S.C., Jupe, F., Sasaki, E., Schmitz, R.J., Urich, M.A., Castanon, R., Nery, J.R., Barragan, C., He, Y., et al. (2016). Epigenomic diversity in a global collection of Arabidopsis thaliana accessions. Cell **166**:492–505. https://doi.org/10.1016/j.cell.2016.06.044.

Kim, B.C., Tennessen, D.J., and Last, R.L. (1998). UV-B-induced photomorphogenesis in Arabidopsis thaliana. Plant J. **15**:667–674. https://doi.org/10.1046/j.1365-313x.1998.00246.x.

Kleine, T., Kindgren, P., Benedict, C., Hendrickson, L., and Strand, A. (2007). Genome-wide gene expression analysis reveals a critical role for CRYPTOCHROME1 in the response of Arabidopsis to high irradiance. Plant Physiol. **144**:1391–1406. https://doi.org/10.1104/pp.107.098293.

Kliebenstein, D.J., Lim, J.E., Landry, L.G., and Last, R.L. (2002). Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. Plant Physiol. **130**:234–243. https://doi.org/10.1104/pp.005041.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. **9**:559. https://doi.org/10.1186/1471-2105-9-559.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature **521**:436–444. https://doi.org/10.1038/nature14539.

Li, H., Rukina, D., David, F.P.A., Li, T.Y., Oh, C.M., Gao, A.W., Katsyuba, E., Bou Sleiman, M., Komljenovic, A., Huang, Q., et al. (2019). Identifying gene function and module connections by the integration of multispecies expression compendia. Genome Res. **29**:2034–2045. https://doi.org/10.1101/gr.251983.119.

Li, J., Ou-Lee, T.M., Raba, R., Amundson, R.G., and Last, R.L. (1993). Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell **5**:171–179. https://doi.org/10.1105/tpc.5.2.171.

Li, Y., Jourdain, A.A., Calvo, S.E., Liu, J.S., and Mootha, V.K. (2017). CLIC, a tool for expanding biological pathways based on co-expression across thousands of datasets. PLoS Comput. Biol. **13**, e1005653. https://doi.org/10.1371/journal.pcbi.1005653.

Liu, H., Hu, B., Wang, Y., Liu, G., Tang, L., Ji, D., Bai, Y., Bao, W., Chen, X., Chen, Y., et al. (2017). Two ultraviolet radiation datasets that cover China. Adv. Atmos. Sci. **34**:805–815. https://doi.org/10.1007/s00376-017-6293-1.

Liu, M., Ma, Z., Zheng, T., Sun, W., Zhang, Y., Jin, W., Zhan, J., Cai, Y., Tang, Y., Wu, Q., et al. (2018). Insights into the correlation between Physiological changes in and seed development of tartary buckwheat (Fagopyrum tataricum Gaertn.). BMC Genom. **19**:648. https://doi.org/10.1186/s12864-018-5036-8.

Matasci, N., Hung, L.-H., Yan, Z., Carpenter, E.J., Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Ayyampalayam, S., Barker, M., et al. (2014). Data access for the 1, 000 Plants (1KP) project. GigaScience **3**:17. https://doi.org/10.1186/2047-217X-3-17.

Mccue, C. (2009). Real World Print Production with Adobe Creative Suite Applications (O'Reilly).

McWhite, C.D., Papoulas, O., Drew, K., Cox, R.M., June, V., Dong, O.X., Kwon, T., Wan, C., Salmi, M.L., Roux, S.J., et al. (2020). A pan-plant protein complex map reveals deep conservation and novel assemblies. Cell **181**:460–474.e14. https://doi.org/10.1016/j.cell.2020.02.049.

Miao, T., Li, D., Huang, Z., Huang, Y., Li, S., and Wang, Y. (2021). Gibberellin regulates UV-B-induced hypocotyl growth inhibition in Arabidopsis thaliana. Plant Signal. Behav. **16**, 1966587. https://doi.org/10.1080/15592324.2021.1966587.

Mirjalili, S. (2019). Genetic algorithm. In Evolutionary Algorithms and Neural Networks: Theory and Applications (Springer International Publishing), pp. 43–55. https://doi.org/10.1007/978-3-319-93025-1_4.

Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S., and Kwasniewski, M.T. (2021). Early detection of plant viral disease using hyperspectral imaging and deep learning. Sensors **21**:742. https://doi.org/10.3390/s21030742.

Nishimura, M.T., Stein, M., Hou, B.H., Vogel, J.P., Edwards, H., and Somerville, S.C. (2003). Loss of a callose synthase results in salicylic acid-dependent disease resistance. Science **301**:969–972. https://doi.org/10.1126/science.1086716.

Norsang, G., Kocbach, L., Tsoja, W., Stamnes, J.J., Dahlback, A., and Nema, P. (2009). Ground-based measurements and modeling of solar UV-B radiation in Lhasa, Tibet. Atmos. Environ. **43**:1498–1502. https://doi.org/10.1016/j.atmosenv.2008.11.048.

Obayashi, T., Kagaya, Y., Aoki, Y., Tadaka, S., and Kinoshita, K. (2019). COXPRESdb v7: a gene coexpression database for 11 animal species supported by 23 coexpression platforms for technical evaluation and evolutionary inference. Nucleic Acids Res. **47**:D55–D62. https://doi.org/10.1093/nar/gky1155.

Rose, C., and Fulton, J. (2003). Sams Teach Yourself Adobe Photoshop Elements 2 in 24 Hours (O'Reilly).

Sasaki, N., and Nakayama, T. (2015). Achievements and perspectives in biochemistry concerning anthocyanin modification for blue flower coloration. Plant Cell Physiol. **56**:28–40. https://doi.org/10.1093/pcp/pcu097.

Saslowsky, D.E., Warek, U., and Winkel, B.S.J. (2005). Nuclear localization of flavonoid enzymes in Arabidopsis. J. Biol. Chem. **280**:23735–23740. https://doi.org/10.1074/jbc.M413506200.

Schmitz, R.J., Schultz, M.D., Urich, M.A., Nery, J.R., Pelizzola, M., Libiger, O., Alix, A., McCosh, R.B., Chen, H., Schork, N.J., et al. (2013). Patterns of population epigenomic diversity. Nature **495**:193–198. https://doi.org/10.1038/nature11968.

Serrano, A.M., Vanhaelewyn, L., Vandenbussche, F., Boccalandro, H.E., Maldonado, B., Van Der Straeten, D., Ballaré, C.L., and Arana, M.V. (2021). Cryptochromes are the dominant photoreceptors mediating heliotropic responses of Arabidopsis inflorescences. Plant Cell Environ. **44**:3246–3256. https://doi.org/10.1111/pce.14139.

Shi, C., and Liu, H. (2021). How plants protect themselves from ultraviolet-B radiation stress. Plant Physiol. **187**:1096–1103. https://doi.org/10.1093/plphys/kiab245.

Short, E.F., North, K.A., Roberts, M.R., Hetherington, A.M., Shirras, A.D., and McAinsh, M.R. (2012). A stress-specific calcium signature regulating an ozone-responsive gene expression network in Arabidopsis. Plant J. **71**:948–961. https://doi.org/10.1111/j.1365-313X.2012.05043.x.

Silva, S., Costa, E.M., Calhau, C., Morais, R.M., and Pintado, M.E. (2017). Anthocyanin extraction from plant tissues: a review. Crit. Rev. Food Sci. Nutr. **57**:3072–3083. https://doi.org/10.1080/10408398.2015.1087963.

Singh, A., Ganapathysubramanian, B., Singh, A.K., and Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. **21**:110–124. https://doi.org/10.1016/j.tplants.2015.10.015.

Song, Y., Jia, Z., Hou, Y., Ma, X., Li, L., Jin, X., and An, L. (2020). Roles of DNA methylation in cold priming in tartary buckwheat. Front. Plant Sci. **11**, 608540. https://doi.org/10.3389/fpls.2020.608540.

Stegle, O., Parts, L., Piipari, M., Winn, J., and Durbin, R. (2012). Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat. Protoc. **7**:500–507. https://doi.org/10.1038/nprot.2011.457.

Szklarczyk, R., Megchelenbrink, W., Cizek, P., Ledent, M., Velemans, G., Szklarczyk, D., and Huynen, M.A. (2016). WeGET: predicting new genes for molecular systems by weighted co-expression. Nucleic Acids Res. **44**:D567–D573. https://doi.org/10.1093/nar/gkv1228.

Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F.A., Roberts, K., and Martin, C. (1998). The AmMYB308 and AmMYB330 transcription factors from Antirrhinum regulated phenylpropanoid and lignin biosynthesis in transgenic tobacco. Plant Cell **10**:135–154.

Tian, F., Yang, D.C., Meng, Y.Q., Jin, J., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. Nucleic Acids Res. **48**:D1104–D1113. https://doi.org/10.1093/nar/gkz1020.

Tissot, N., and Ulm, R. (2020). Cryptochrome-mediated blue-light signalling modulates UVR8 photoreceptor activity and contributes to UV-B tolerance in Arabidopsis. Nat. Commun. **11**:1323. https://doi.org/10.1038/s41467-020-15133-y.

Tohge, T., and Fernie, A.R. (2017). Leveraging natural variance towards enhanced understanding of phytochemical sunscreens. Trends Plant Sci. **22**:308–315. https://doi.org/10.1016/j.tplants.2017.01.003.

Tohge, T., Kusano, M., Fukushima, A., Saito, K., and Fernie, A.R. (2011). Transcriptional and metabolic programs following exposure of plants to UV-B irradiation. Plant Signal. Behav. **6**:1987–1992. https://doi.org/10.4161/psb.6.12.18240.

Trivedi, P., Leach, J.E., Tringe, S.G., Sa, T., and Singh, B.K. (2020). Plant–microbiome interactions: from community assembly to plant health. Nat. Rev. Microbiol. **18**:607–621. https://doi.org/10.1038/s41579-020-0412-1.

Van Aken, O., and Whelan, J. (2012). Comparison of transcriptional changes to chloroplast and mitochondrial perturbations reveals common and specific responses in Arabidopsis. Front. Plant Sci. **3**:281. https://doi.org/10.3389/fpls.2012.00281.

Vanhaelewyn, L., Prinsen, E., Van Der Straeten, D., and Vandenbussche, F. (2016). Hormone-controlled UV-B responses in plants. J. Exp. Bot. **67**:4469–4482. https://doi.org/10.1093/jxb/erw261.

Vanhaelewyn, L., Van Der Straeten, D., De Coninck, B., and Vandenbussche, F. (2020). Ultraviolet radiation from a plant perspective: the plant-microorganism context. Front. Plant Sci. **11**, 597642. https://doi.org/10.3389/fpls.2020.597642.

Wan, J., Zhang, P., Wang, R., Sun, L., Wang, W., Zhou, H., and Xu, J. (2018). UV-B radiation induces root bending through the flavonoid-mediated auxin pathway in arabidopsis. Front. Plant Sci. **9**:618. https://doi.org/10.3389/fpls.2018.00618.

Wang, J., Sun, P., Li, Y., Liu, Y., Yu, J., Ma, X., Sun, S., Yang, N., Xia, R., Lei, T., et al. (2017). Hierarchically aligning 10 legume genomes establishes a family-level genomics platform. Plant Physiol. **174**:284–300. https://doi.org/10.1104/pp.16.01981.

Wang, S.H., Xie, S., Chen, X., Guttery, D.S., Tang, C., Sun, J., and Zhang, Y.-D. (2019). Alcoholism identification based on an AlexNet transfer learning model. Front. Psychiatry **10**. https://doi.org/10.3389/fpsyt.2019.00205.

Wang, Y., Zhang, W.Z., Song, L.F., Zou, J.J., Su, Z., and Wu, W.H. (2008). Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in Arabidopsis. Plant Physiol. **148**:1201–1211. https://doi.org/10.1104/pp.108.126375.

Wang, Y., Seburn, K., Bechtel, L., Lee, B.Y., Szatkiewicz, J.P., Nishina, P.M., and Naggert, J.K. (2006). Defective carbohydrate metabolism in mice homozygous for the tubby mutation. Physiol. Genomics **27**:131–140. https://doi.org/10.1152/physiolgenomics.00239.2005.

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. **40**:e49. https://doi.org/10.1093/nar/gkr1293.

Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. **38**:W214–W220. https://doi.org/10.1093/nar/gkq537.

Wargent, J.J., Gegas, V.C., Jenkins, G.I., Doonan, J.H., and Paul, N.D. (2009). UVR8 in Arabidopsis thaliana regulates multiple aspects of cellular differentiation during leaf development in response to ultraviolet B radiation. New Phytol. **183**:315–326. https://doi.org/10.1111/j.1469-8137.2009.02855.x.

Winkel-Shirley, B. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. **126**:485–493. https://doi.org/10.1104/pp.126.2.485.

Wu, D., and Smyth, G.K. (2012). Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. **40**:e133. https://doi.org/10.1093/nar/gks461.

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci. **20**:176–185. https://doi.org/10.1016/j.tplants.2014.12.001.

Yadav, A., Singh, D., Lingwan, M., Yadukrishnan, P., Masakapalli, S.K., and Datta, S. (2020). Light signaling and UV-B-mediated plant growth regulation. J. Integr. Plant Biol. **62**:1270–1292. https://doi.org/10.1111/jipb.12932.

Yan, K., Guo, X., Ji, Z., and Zhou, X. (2021). Deep transfer learning for cross-species plant disease diagnosis adapting mixed subdomains. IEEE/ACM Trans Comput Biol Bioinform. PP. https://doi.org/10.1109/TCBB.2021.3135882.

Yang, Y., Zhang, L., Chen, P., Liang, T., Li, X., and Liu, H. (2020). UV-B photoreceptor UVR8 interacts with MYB73/MYB77 to regulate auxin responses and lateral root development. EMBO J. **39**, e101928. https://doi.org/10.15252/embj.2019101928.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., et al. (2006). The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. **60**:107–124. https://doi.org/10.1007/s11103-005-2910-y.

Yin, R., and Ulm, R. (2017). How plants cope with UV-B: from perception to response. Curr. Opin. Plant Biol. **37**:42–48. https://doi.org/10.1016/j.pbi.2017.03.013.

Yoon, H.I., Kim, H.Y., Kim, J., Oh, M.M., and Son, J.E. (2021). Quantitative analysis of UV-B radiation interception in 3D plant structures and intraindividual distribution of phenolic contents. Int. J. Mol. Sci. **22**:2701. https://doi.org/10.3390/ijms22052701.

Yuan, Y., Wu, C., Liu, Y., Yang, J., and Huang, L. (2013). The Scutellaria baicalensis R2R3-MYB transcription factors modulates flavonoid biosynthesis by regulating GA metabolism in transgenic tobacco plants. PLoS One **8**:e77275–e77289. https://doi.org/10.1371/journal.pone.0077275.

Zeng, X., Yuan, H., Dong, X., Peng, M., Jing, X., Xu, Q., Tang, T., Wang, Y., Zha, S., Gao, M., et al. (2020). Genome-wide dissection of Co-selected UV-B responsive pathways in the UV-B adaptation of Qingke. Mol. Plant **13**:112–127. https://doi.org/10.1016/j.molp.2019.10.009.

Zhang, B., Xu, X., Huang, R., Yang, S., Li, M., and Guo, Y. (2021a). CRISPR/Cas9-mediated targeted mutation reveals a role for AN4 rather than DPL in regulating venation formation in the corolla tube of Petunia hybrida. Hortic. Res. **8**:116. https://doi.org/10.1038/s41438-021-00555-6.

Zhang, K., Logacheva, M.D., Meng, Y., Hu, J., Wan, D., Li, L., Janovská, D., Wang, Z., Georgiev, M.I., Yu, Z., et al. (2018). Jasmonate-Responsive MYB factors spatially repress rutin biosynthesis in fagopyrum tataricum. J. Exp. Bot. **69**:1955–1966.

Zhang, K., He, M., Fan, Y., Zhao, H., Gao, B., Yang, K., Li, F., Tang, Y., Gao, Q., Lin, T., et al. (2021b). Resequencing of global Tartary buckwheat accessions reveals multiple domestication events and key loci associated with agronomic traits. Genome Biol. **22**:23. https://doi.org/10.1186/s13059-020-02217-7.

Zhang, L., Li, X., Ma, B., Gao, Q., Du, H., Han, Y., Li, Y., Cao, Y., Qi, M., Zhu, Y., et al. (2017). The tartary buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance. Mol. Plant **10**:1224–1237. https://doi.org/10.1016/j.molp.2017.08.013.

Zhang, P., Berardini, T.Z., Ebert, D., Li, Q., Mi, H., Muruganujan, A., Prithvi, T., Reiser, L., Sawant, S., Thomas, P.D., et al. (2020).

PhyloGenes: an online phylogenetics and functional genomics resource for plant gene function inference. Plant Direct **4**, e00293. https://doi.org/10.1002/pld3.293.

Zhang, R.-b., Yuan, Y.-j., Wei, W.-s., Gou, X.-h., Yu, S.-l., Shang, H.-m., Chen, F., Zhang, T.-w., and Qin, L. (2015). Dendroclimatic reconstruction of autumn–winter mean minimum temperature in the eastern Tibetan Plateau since 1600 AD. Dendrochronologia **33**:1–7. https://doi.org/10.1016/j.dendro.2014.09.001.

Zhou, M., Zhang, K., Sun, Z., Yan, M., Chen, C., Zhang, X., Tang, Y., and Wu, Y. (2017). LNK1 and LNK2 corepressors interact with the MYB3 transcription factor in phenylpropanoid biosynthesis. Plant Physiol. **174**:1348–1358. https://doi.org/10.1104/pp.17.00160.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T., and Felix, G. (2006). Perception of the bacterial PAMP EF-tu by the receptor EFR restricts agrobacterium-mediated transformation. Cell **125**:749–760. https://doi.org/10.1016/j.cell.2006.03.037.