ADVANCED SCIENCE Open Access
www.advancedscience.com

Evolution and Domestication of a Novel Biosynthetic Gene Cluster Contributing to the Flavonoid Metabolism and High-Altitude Adaptability of Plants in the *Fagopyrum* Genus

*Xu Huang, Yuqi He, Kaixuan Zhang, Yaliang Shi, Hui Zhao, Dili Lai, Hao Lin, Xiangru Wang, Zhimin Yang, Yawen Xiao, Wei Li, Yinan Ouyang, Sun Hee Woo, Muriel Quinet, Milen I. Georgiev, Alisdair R. Fernie, Xu Liu, and Meiliang Zhou**

The diversity of secondary metabolites is an important means for plants to cope with the complex and ever-changing terrestrial environment. Plant biosynthetic gene clusters (BGCs) are crucial for the biosynthesis of secondary metabolites. The domestication and evolution of BGCs and how they affect plant secondary metabolites biosynthesis and environmental adaptation are still not fully understood. Buckwheat exhibits strong resistance and abundant secondary metabolites, especially flavonoids, allowing it to thrive in harsh environments. A non-canonical BGC named ***UFGT3*** cluster is identified, which comprises a phosphorylase kinase (***PAK***), two transcription factors (***MADS1/2***), and a glycosyltransferase (***UFGT3***), forming a complete molecular regulatory module involved in flavonoid biosynthesis. This cluster is selected during Tartary buckwheat domestication and is widely present in species of the ***Fagopyrum*** genus. In wild relatives of cultivated buckwheat, a gene encoding anthocyanin glycosyltransferase (AGT), which glycosylates pelargonidin into pelargonidin-3-***O***-glucoside, is found inserted into this cluster. The pelargonidin-3-***O***-glucoside can help plants resist UV stress, endowing wild relatives with stronger high-altitude adaptability. This study provides a new research paradigm for the evolutionary dynamics of plant BGCs, and offers new perspectives for exploring the mechanism of plant ecological adaptability driven by environmental stress through the synthesis of secondary metabolites.

1. Introduction

Since plants first colonized terrestrial environments about 500 million years ago,[1] they have had to face enormous challenges from the environment. In order to cope with these environmental conditions, plants have evolved multiple strategies for adaptation, such as morphological adaptation, establishing symbiotic relationships with microorganisms, sensitivity to temperature and water changes, complex transcriptional regulation, evoking signal transduction pathways, and the synthesis of secondary metabolites. Flavonoids, belonging to polyphenolic compounds, are important plant secondary metabolites participating in plant-environment interactions.[2-6] Therefore, understanding the biosynthetic mechanism of flavonoids is crucial for exploring evolutionary and domestication mechanisms of plants' adaptation to environmental stress.

So far, over 9000 flavonoids have been identified in plants.[7-9] Besides being diverse, flavonoid biosynthesis is extremely complex, with multiple intermediates and derivatives interconverted.[10]

X. Huang, Y. He, K. Zhang, Y. Shi, D. Lai, H. Lin, X. Wang, Z. Yang, Y. Xiao, W. Li, Y. Ouyang, X. Liu, M. Zhou
National Key Facility for Crop Gene Resources and Genetic Improvement
Institute of Crop Sciences
Chinese Academy of Agricultural Sciences
Beijing 100081, China
E-mail: zhoumeiliang@caas.cn

S. H. Woo
Department of Agronomy
Chungbuk National University
Cheongju 28644, South Korea

M. Quinet
Groupe de Recherche en Physiologie Végétale (GRPV)
Earth and Life Institute-Agronomy (ELI-A)
Université Catholique de Louvain
Croix du Sud 45, boîte L7.07.13, Louvain-la-Neuve B-1348, Belgium

M. I. Georgiev
Laboratory of Metabolomics
Institute of Microbiology
Bulgarian Academy of Sciences
Plovdiv 4000, Bulgaria

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/advs.202403603

DOI: 10.1002/advs.202403603

Additionally, it is highly susceptible to environmental influence, meaning that the levels of flavonoids are highly variable and relatively unstable.[11] Glycosyltransferases, such as UDP-glucose-flavonoid-glycosyltransferase (UFGT) and anthocyanin glycosyltransferase (AGT), can alter the solubility and bioavailability of flavonoids by glycosylation, thus imparting different functional properties to flavonoids in various cellular components.[12] In addition, the biosynthesis and metabolism of flavonoids were regulated by various transcription factors.[13-15] For example, in *Eriobotrya japonica*, MADS-box transcription factor EjCAL directly binds to the promoter of the glycosyltransferase gene *EjUF3GaT1* and activates gene expression, regulating the biosynthesis of hyperoside.[16] The MADS family transcription factor MdJa2 can directly bind to the promoter of downstream target genes, inhibiting the synthesis of anthocyanins and proanthocyanidins in red-fleshed apple.[17] In buckwheat, MYB and ERF transcription factors have been identified to directly bind to promoters of genes encoding flavonoid biosynthetic enzymes or transporters thus regulating gene expression.[18-20] These diverse enzymes and transcription factors lead to plants possessing different types and abundances of flavonoids.

The synthesis of metabolites typically involves the coordinated action of multiple genes, which may exist in the form of gene clusters, known asBGCs. Gene clusters (operons) were originally discovered in prokaryotes, where ≈50% of the genes were clustered into gene clusters.[21,22] Besides being well-characterized in bacteria, there has been extensive research on gene clusters in fungi.[23-25] In plants, BGCs usually consist of a set of genes encoding enzymes in the biosynthesis or modification pathways.[26-28] For instance, *UDP-glycosyltransferases UGT85B1* is one of the four core members of the cyanogenic glycoside dhurrin gene cluster in *Sorghum bicolor*.[29] UGT85B1 can interact with the proteins encoded by two other members of this gene cluster, CYP79A1, and CYP71E1, forming a channeling complex that facilitates the rapid flow of metabolic intermediates during dhurrin biosynthesis.[30] However, in plants, BGCs containing specific transcription factors within gene clusters have not been reported.[31] In conclusion, despite the fact that BGCs are unique and important for plant biosynthesis, their regulatory patterns, as well as their formation and evolution, remain incompletely understood.

Metabolomics is a powerful tool for studying the modification of flavonoid compounds.[32] In previous research, through GWAS analysis of metabolites contents in Tartary buckwheat germplasm resources, a gene encoding FtUFGT3 was identified as associated with flavonoid contents.[33] In this study, enzymatic studies confirmed that FtUFGT3 is a key enzyme in flavonoid biosynthesis, and could catalyze the glycosylation of kaempferol, quercetin, and myricetin. Further analysis revealed that *FtUFGT3*, along with three adjacent genes (two genes encoding MADS-box transcription factors FtMADS1 and FtMADS2, as well as one gene encoding a protein kinase FtPAK) formed a molecular module type BGC that co-regulates flavonoid biosynthesis in Tartary buckwheat. This BGC is located within the region that underwent selection during Tartary buckwheat domestication and is widely present in buckwheat plants. In wild relatives of buckwheat, a *UDP-glucose transferase* (*AGT*) was found inserted into this gene cluster. This AGT could glycosylate pelargonidin into pelargonidin-3-*O*-glucoside, which could help plants alleviate the UV damage thus enhancing the adaptability to high-altitude environments. This study provides new insights into the molecular mechanisms of flavonoid biosynthesis and ecological adaptation in buckwheat. It reveals the impact of genomic evolution on plant ecological adaptability, contributing to the development and utilization of plant resources.

2. Results

2.1. FtUFGT3 Is a Crucial Enzyme Catalyzing the Glycosylation of Various Flavonoid Substrates

In our previous research, GWAS analysis identified a significant SNP (Ft1: 4617722), located in the promoter of *FtUFGT3*, associated with the content of quercetin-3-*O*-glucoside in Tartary buckwheat germplasm resources.[33,34] (**Figure 1A**; Figure S1, Supporting Information). Similarly, through GWAS, the contents of substances such as kaempferol-3-*O*-glucoside-7-*O*-rhamnoside and kaempferol-3-*O*-flavonoid glucoside, were found closely associated with this SNP (Figures S2–S4, Supporting Information), suggesting that FtUFGT3 may influence the synthesis of multiple flavonoid compounds. Further analysis revealed that this locus underwent selection during Tartary buckwheat domestication based on cross-population composite likelihood ratio (XP-CLR) and F_{st} analysis (Figure 1B). Previous studies have shown that FtUFGT3 could also glycosylate cyanidin and kaempferol.[33,35].To investigate whether FtUFGT3 could glycosylate other flavonoids, we next analyzed the metabolite content in accessions with different genotype based on this SNP.[31] The results showed that accessions harboring the G-genotype exhibited a higher ratio of quercetin-3-*O*-glucoside to quercetin compared to those harboring A-genotype (Figure 1C–E). A similar result was also found in the ratio of myricetin-3-*O*-glucoside to myricetin (Figure S5, Supporting Information). As previous research illustrated the expression of *FtUFGT3* was significantly higher in A-genotype than in the G-haplotype.[33] we speculated that FtUFGT3 might also participate in the glycosylation of quercetin and myricetin.

To verify the substrates of FtUFGT3, molecular docking simulations were performed with these candidate substrates. The results showed that quercetin (Figure 1F; Figure S6, Supporting Information), kaempferol (Figures S7 and S8, Supporting Information), myricetin (Figures S9 and S10, Supporting Information), and UDP-glucose (Figure S11, Supporting Information) exhibited relatively high binding efficiency. Therefore, we hypothesize that FtUFGT3 can catalyze the glycosylation of not only kaempferol but also quercetin and myricetin. To verify this hypothesis, FtUFGT3-MBP protein was obtained through prokaryotic expression (Figure S12, Supporting Information). Using quercetin, kaempferol, and myricetin as substrates, the in vitro enzymatic activity of FtUFGT3 was measured. The

M. I. Georgiev, A. R. Fernie
Center of Plant Systems Biology and Biotechnology
Plovdiv 4000, Bulgaria
A. R. Fernie
Department of Molecular Physiology
Max-Planck-Institute of Molecular Plant Physiology
Potsdam 14476, Germany

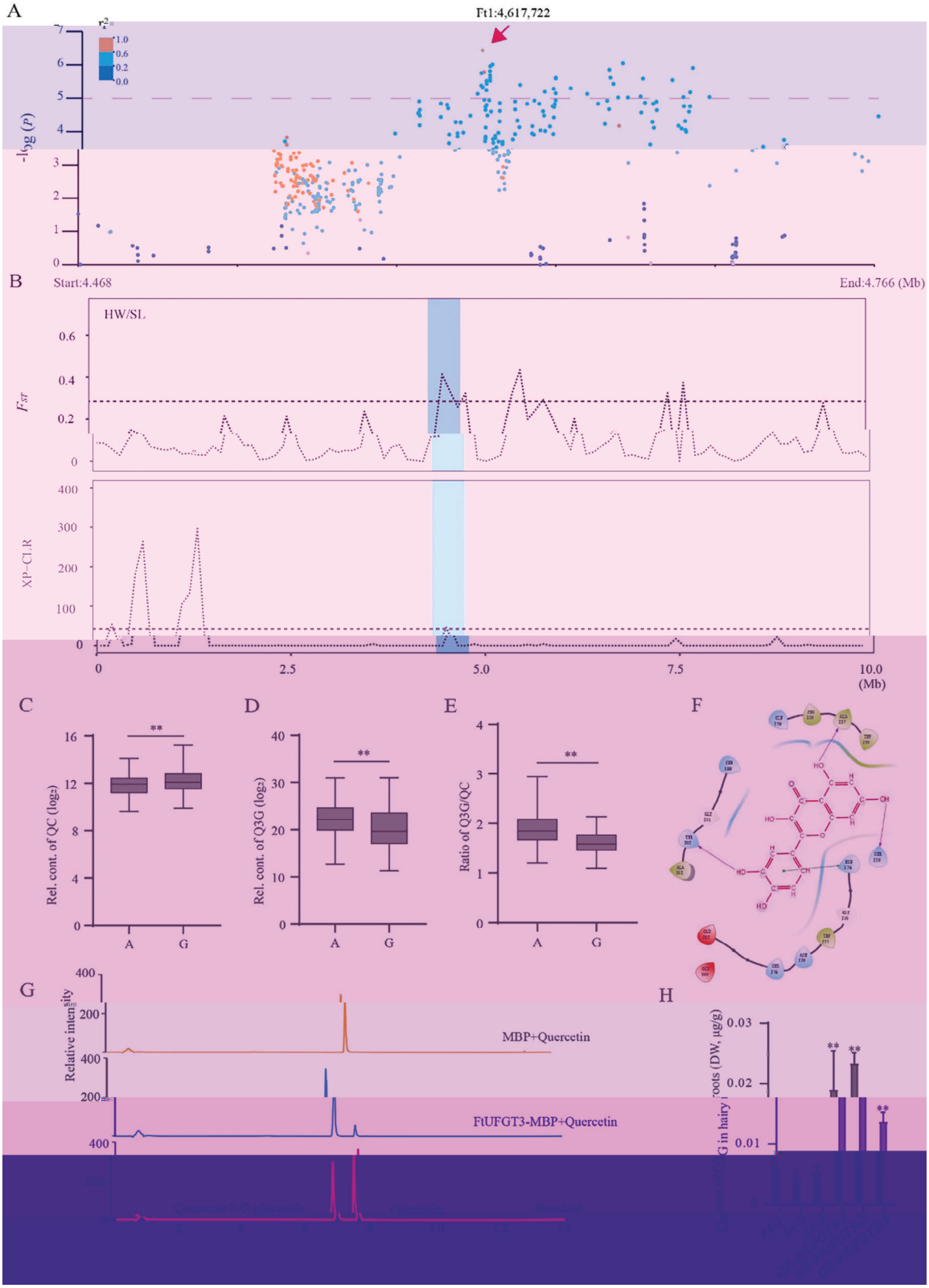

Figure 1. The multi-catalytic function of FtUFGT3.A) Ft1:4617722 (*FtUFGT3*) was identified through GWAS on quercetin-3-*O*-glucoside in Tartary buckwheat germplasm resources. The dashed line indicates the threshold $-\log P = 5$. The red arrow indicates the lead SNP. B) The upper part illustrates population differentiation (F_{st}) with selective sweeps in Tartary buckwheat. The lower part is the XP-CLR plot of FtUFGT3. F_{st} and XP-CLR are plotted using 10 Mb sliding window. The black horizontal dashed line represents the genome-wide cutoff with the highest being 5%. C) Box plots show the

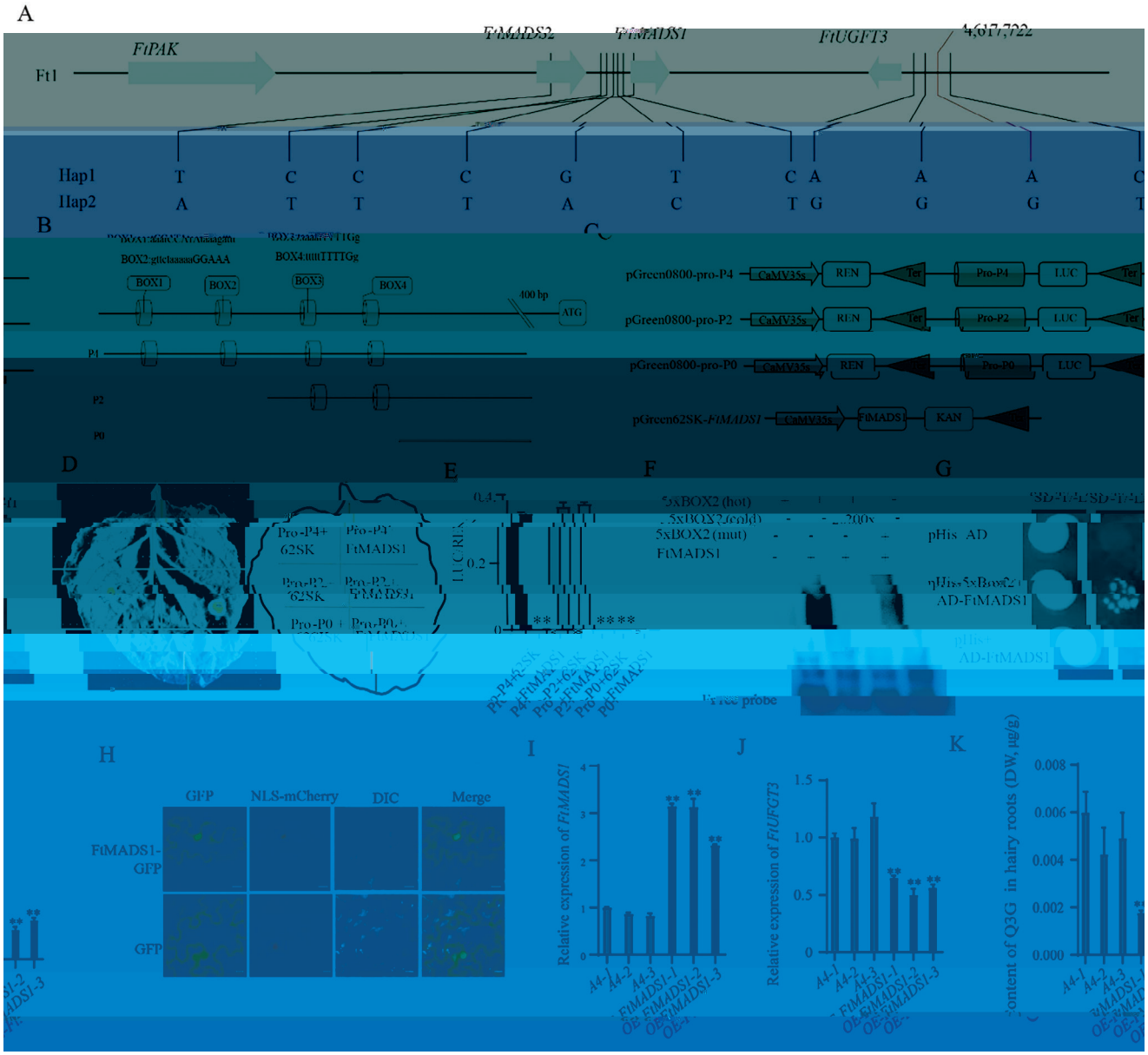

Figure 2. FtMADS1 inhibits *FtUFGT3* expression through directly binding to the promoter of *FtUFGT3*. A) The haplotype of gene cluster *FtPAK-FtMADS1/2-FtUFGT3*. B) A diagram showing the distribution of MADS-box elements on the upstream 2000 bp region of the *FtUFGT3* promoter and the promoter fragments used in this study. C) Schematic diagram of the recombinant vector in the transcription activation experiment. D) The results of dual bioluminescence assay. The fusion proteins were expressed in *N-benthamiana* using agroinfiltration. Chemiluminescence images were captured 36 h after infiltration using 3 mg mL^{-1} luciferin. Similar results were obtained in 3 biological replicates. E) The ratio of the luciferase (LUC) activity and the recombinant enzyme activity of the reference gene (renilla, REN) of D). Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P <$ 0.01, Student's *t*-test. F) Results of the EMSA for the interaction between FtMADS1 and 5×BOX2. G) Y2H results of FtMADS1 and FtMADS2. SD-L/-T, the SD basic culture medium lacked Leu and Trp. SD-L/-T/-H/-A, SD basal medium lacked Leu, Trp, His, and Ade, and added 10 mm3-AT. H)Subcellular localization of FtMADS1. FtMADS1-GFP, pCAMBIA1300-FtMADS1 recombination plasmid; NLS-mCherry, nuclear marker; DIC, differential interference contrast ; Merge, merge channel. Bar = 25 μm. I) The relative expression level of *FtMADS1* in three *OE-FtMADS1* hairy root lines. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's t-test. J) The relative expression level of *FtUFGT3* in three *OE-FtMADS1* hairy root lines. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's t-test. K) Content of quercetin-3-*O*-glucoside in three *OE-FtMADS1* hairy roots lines. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's *t*-test.

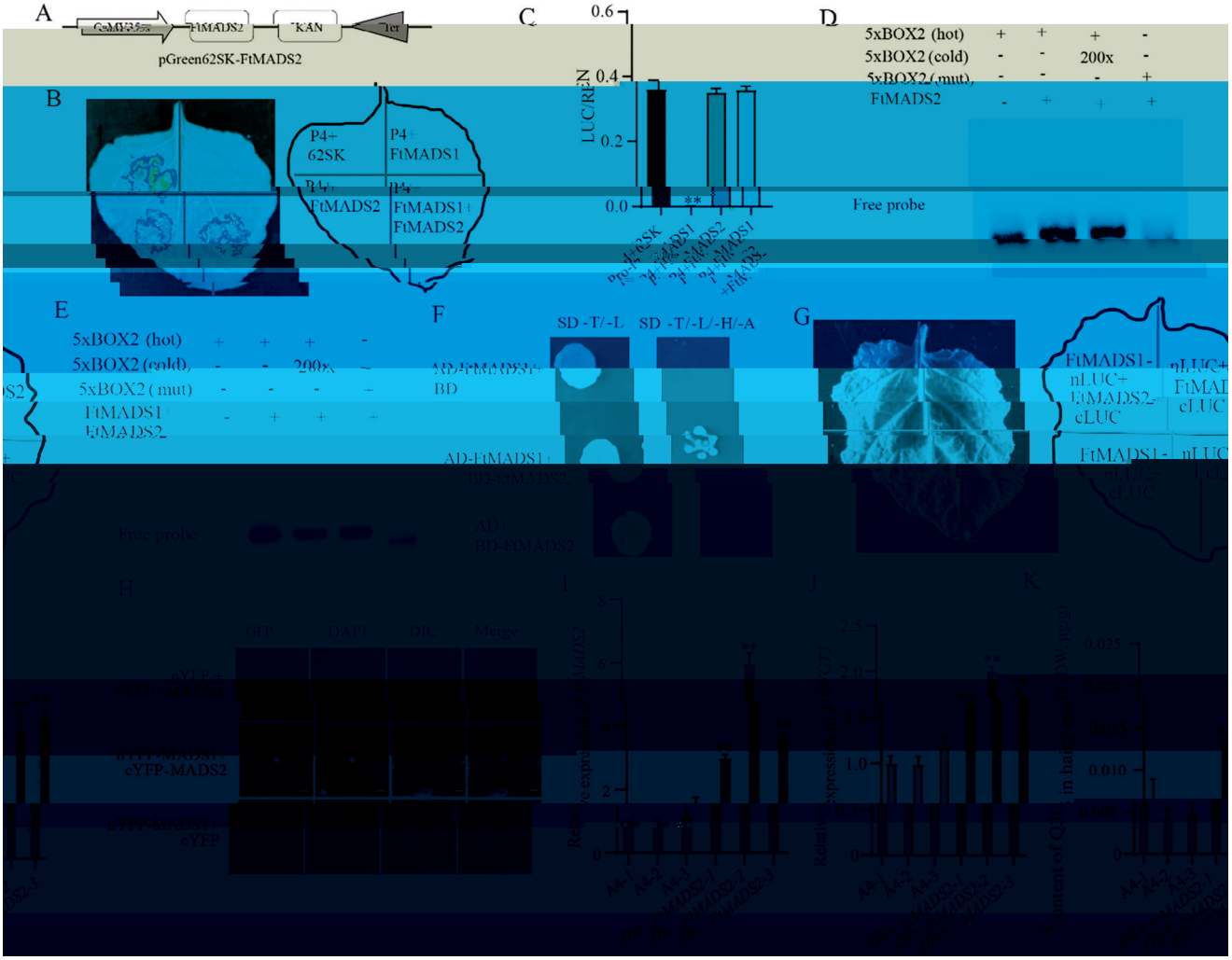

Figure 3. The interaction between FtMADS2 and FtMADS1 alleviates the inhibitory effect of FtMADS1 on *FtUFGT3* expression.A) Schematic diagram of the recombinant vector in the dual bioluminescence assay. B) The results of dual bioluminescence assay. The fusion proteins were expressed in *N-benthamiana* using agroinfiltration. Chemiluminescence images were captured 36 h after infiltration using 3 mg mL^{-1} luciferin. Similar results were obtained in 3 biological replicates. C) The ratio of the luciferase (LUC) activity and the recombinant enzyme activity of the reference gene renilla (REN) of B. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's *t*-test. D) Results of the EMSA for the interaction between FtMADS2 and 5×BOX2. E) Results of the EMSA for the interaction between FtMADS1+FtMADS2 and 5×BOX2. F) The yeast two-hybrid (Y2H) results of the interaction of FtMADS1 and FtMADS2. SD-L/-T, SD basic medium lacked the Leu and Trp; SD-L/-T/-H/-A, SD basal medium lacked Leu, Trp,His and Ade, and contained 50 mM 3-AT. G) The interaction between FtMADS1 and FtMADS2 was confirmed using LCA. The N- or C-terminal fragment of LUC (nLUC or cLUC) was fused with their respective proteins. The experiment was performed according to the grouping shown in the figure. The constructed fusion proteins were co-expressed in *N.benthamiana* using agroinfiltration. Images of chemiluminescence were recorded by applying 3 mg mL^{-1} luciferin 36 h after infiltration. Similar results were obtained in 3 biological replicates. H) The BiFC assay demonstrates interactions between FtMADS1 and FtMADS2 in the leaf epidermal cells of *N. benthamiana*. The N- or C-terminal fragment of YFP (nYFP or cYFP) was fused with their respective proteins. The experiment was performed according to the grouping shown in the figure. The constructed fusion proteins were co-expressed in *N.benthamiana* using agroinfiltration. 4′,6-diamidino-2-phenylindole (DAPI) serves as a reliable marker for cellular nuclei. The results were observed through confocal microscopy after 48 h. The scale bar represents 25 μm. I) The relative expression level of *FtMADS2* in three *OE-FtMADS2* hairy root lines. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's t test. J) The relative expression level of *FtUFGT3* in three *OE-FtMADS2* hairy root lines. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's t test. K) Content of quercetin-3-*O*-glucoside in three *OE-FtMADS2* hairy roots lines. Data show the arithmetic mean ± SD from 3 biological replicates. $^{**}P < 0.01$, Student's *t*-test.

in the presence of FtMADS2, FtMADS1 no longer binds to the *FtUFGT3* promoter (Figure 3E). These results collectively suggest that although FtMADS2 cannot directly regulate the activity of *FtUFGT3* promoter, it can relieve the inhibition effect of FtMADS1 on *FtUFGT3* expression.

To explore the mechanism by which FtMADS2 relieves the inhibitory effect of FtMADS1 on *FtUFGT3* expression, we next investigated whether FtMADS2 could interact with FtMADS1. Y2HGold revealed that the yeast strain co-transformed with FtMADS1 and FtMAD2 was able to grow normally on

SD-T/-L-/H/-A medium, while the control group could not (Figure 3F). Subsequently, luciferase complementation assay revealed strong luciferase activity on the co-transformation of FtMADS1-nLUC with FtMAD2-cLUC, but no bioluminescence was observed in the negative controls (Figure 3G,H). Bimolecular fluorescent complimentary (BiFC) assay also showed strong fluorescence in the nucleus but no fluorescence was observed in the negative controls (Figure 3I). These results collectively indicated that FtMADS2 alleviated the transcription suppression effect of FtMADS1 via its interaction with FtMADS1. In order to investigate the function of FtMADS2 in vivo, *FtMADS2* overexpressing Tartary buckwheat hairy roots were constructed (Figure 3J). The expression of *FtUFGT3* was significantly upregulated (Figure 3K), which was in contrast with the function of *FtMADS1*, indicating that *FtMADS2* can positively regulate the expression of *FtUFGT3* in Tartary buckwheat. Accordingly, the content of quercetin-3-*O*-glucoside was significantly upregulated compared to the empty vector controls (Figure 3L). In summary, these results demonstrated that the interaction between FtMADS2 and FtMADS1 alleviates the transcription inhibition effect of FtMADS1 on *FtUFGT3* expression.

2.4. The Interaction Between FtPAK and FtMADS2 Enhances the Inhibitory E□ect of FtMADS1 on *UFGT3* Expression

Previous research illustrated that phosphorylation of MADS transcription factors is crucial for their function, and this phosphorylation process is dependent on the interaction with protein kinases.[36,37] We found that within the 100 kb up- and downstream regions of the lead SNP (Ft1: 4617722), a gene encoding protein kinase (FtPAK) was also highly expressed during Tartary buckwheat seed development.[33] which was in accoradance with *FtUFGT3*, *FtMADS1* and *FtMADS2*. To examine whether this protein kinase can interact with FtMADSs, a luciferase complementation assay was further conducted. Live imaging showed a distinct luciferase activity when FtPAK-nLUC and FtMAD2-cLUC were co-present (**Figure** 4A,B). Subsequently, recombinant proteins of FtPAK-His, FtMADS1-GST, and FtMADS2-GST were obtained in order to verify their interactions via pull-down analyses. The results indicated that FtPAK could interact with FtMADS2 but not FtMADS1, and this interaction requires the presence of ATP (Figure 4C). Results of Y2H analyses indicated that the yeast strain co-transformed with FtPAK and FtMADS2 was able to grow on SD-T/-L-/H/-A medium, while the negative controls could not (Figure 4D). Subcellular localization experiments demonstrated that FtPAK is localized in the nucleus and cytoplasm, while FtMADS2 is solely localized in the nucleus (Figure 4E). However, when tobacco cells co-transformed with FtMADS2 and FtPAK, the subcellular location of FtMADS2 changed. The number of cells with FtMADS2 cytoplasm-located was increased in FtMADS2 and FtPAK co-transformed, compared to that transformed with FtMADS2 alone (Figure 4E,F). To investigate whether this interaction could result in FtMADS2 phosphorylation, Phostag experiments were further performed and showed that FtMADS2 could be phosphorylated by FtPAK (Figure 4G), indicating thacription iTJ3□ irFtMADS2

Figure 4. The interaction between FtPAK and FtMADS2 enhances the inhibitory effect of FtMADS1 on UFGT3. A) The interaction between FtPAK and FtMADS2 was confirmed using LCA. The N- or C-terminal fragment of LUC (nLUC or cLUC) was fused with their respective proteins. The experiment was performed according to the grouping shown in the figure. The constructed fusion proteins were co-expressed in *N-benthamiana* using agroinfiltration. Images of chemiluminescence were recorded by applying 3 mg mL^{-1} luciferin 36 h after infiltration. Similar results were obtained in 3 biological replicates.

the UV-B and salt resistance tests also support this point. As shown in **Figure 5**A,B, after UV-B and NaCl treatment, compared to *F. longistylum*, the leaves of Tartary buckwheat displayed curling and wilting, indicating that wild relatives have stronger resistance to UV and NaCl. Therefore, we hypothesize that the *AGT* gene inserted in the UFGT3 BGC gives the wild relatives stronger high-altitude adaptability. We further investigate whether this gene was also regulated by UFGT3 cluster. The dual-luciferase reporter assay further showed that MADS1 can inhibit *AGT* expression (Figure 5C). Further research revealed that in the presence of both MADS2 and MADS1, the expression of *AGT* is no longer inhibited (Figure 5C), indicating that the interaction between MADS2 and MADS1 can relieve the inhibition effect of MADS1 on *AGT*. Furthermore, when PAK and MADS1/2 are simultaneously present, the inhibition effect of MADS1 on *AGT* expression is reinstated (Figure 5C). This result indicates that PAK phosphorylated MADS2 disrupts the interaction between MADS1 and MADS2, allowing MADS1 to inhibit *AGT* expression.

The function of AGT was further analyzed. Subcellular localization showed that FlAGT and FpAGT are located in both the cytoplasm and nucleus (Figure 5D). As AGT is an anthocyanin glucosyltransferase, we tested the molecular docking of FlAGT with several major anthocyanins (cyanidin, petunidol, pelargonidin, and delphinidin). The results showed that FlAGT exhibited varying degrees of catalytic activity toward them (Figures S42 and S45, Supporting Information). In vitro enzyme activity experiments revealed that FlAGT could only glycosylate pelargonidin into pelargonidin-3-*O*-glucoside (Figure 5E), but could not glycosylate the other three substances. Overexpression of *FlAGT* in hairy roots significantly increased the content of pelargonidin 3-*O*-glucoside (Figure S46, Supporting Information; Figure 5F), further confirming the function of *FlAGT*. It is worth noting that after UV-B, there was a significant upregulation of pelargonidin-3-*O*-glucoside content in the *F. longistylum*, while there was no significant difference compared to untreated buckwheat (Figure S47, Supporting Information).

We investigated the protective effects of pelargonidin-3-*O*-glucoside on buckwheat and other crops against UV-B damage. Measurement of pelargonidin-3-*O*-glucoside levels revealed a significant increase in treated buckwheat, indicating its effective entry into the plant through the surface (Figure S48, Supporting Information). UV-B treatment combined with diaminobenzidine (DAB) staining visualized damage in plants treated with pelargonidin-3-*O*-glucoside, showing a notable reduction in rust disease on the young leaves of Tartary buckwheat compared to controls (Figure 5H; Figures S52B, Supporting Information). Additionally, symptoms of rust were diminished on young wheat leaves, and root growth was healthy (Figure 5I; Figure S51A, Supporting Information). In rice, surface damage on seedlings was also reduced compared to controls (Figure 5J). Similar protective effects were observed in barley and *F. homotropicum*, as shown in Figures S49 and S50 (Supporting Information), with visual results in Figure S51B,C (Supporting Information). In the field experiment conducted at an altitude of 3500 m, three widely cultivated varieties of Tartary buckwheat (ZK3, CQ1, CQ2) were subjected to foliar spraying with pelargonidin 3-*O*-glucoside. The results revealed that the treated varieties exhibited an increase in both plant height and grain number (Figures S53 and S54, Supporting Information). These findings suggest that pelargonidin 3-*O*-glucoside universally enhances plant resistance to ultraviolet stress and holds significant developmental and utilitarian value.

In conclusion, a model can be formulated to describe the regulatory patterns of the *UFGT3* BGC in cultivated buckwheat and wild relatives (**Figure 6**). In cultivated buckwheat, when the expression level of *PAK* is low, MADS2 can interact with MADS1, alleviating the transcriptional inhibitory effect of FtMADS1 on the expression of *FtUFGT3*, thus promoting the glycosylation of its substrates. However, when the expression level of *PAK* is high, PAK can interact with phosphate MADS2, leading to the translocation of MADS2 from the nucleus to the cytoplasm. This means that MADS2 can no longer interact with MADS1, causing the MADS1 transcription factor to bind to the MADS-BOX2 element on the *UFGT3* promoter, thereby inhibiting

B) The results of the pull-down assay to detect the interaction between FtPAK and FtMADS1 and FtMADS2. In the experiment, recombinant proteins FtPAK-His, FtMADS1-GST, and FtMADS2-GST were obtained through prokaryotic expression. As shown in Figure, FtPAK-His was co-incubated with FtMADS1-GST or FtMADS2-GST. GST-tagged beads were used for the pull-down assay and anti-His was used for western blot detection. P represents Beads protein elution solution, W1 represents the first beads wash solution, W2 represents the final beads wash solution, and Anti-GST and Anti-His represent the antibodies used for western blot. C) Y2H results of FtMADS1 and FtMADS2. SD-L/-T, the SD basic culture medium lacked Leu and Trp. SD-L/-T/-H/-A, SD basal medium lacked Leu, Trp, His and Ade, and added 10 mm3-AT. D) Subcellular localization of FtMADS2 and FtAPK. FtMADS2-GFP, pCAMBIA1300-FtMADS2 recombination plasmid. FtPAK-GFP, pCAMBIA1300-FtPAK recombination plasmid; FtPAK-Myc, pCAMBIA1307-FtPAK recombination plasmid. NLS-mCherry, nuclear marker; Bar = 25 μm. E) The percentage of cells with cytoplasmic localization out of the total number of cells. 20 cells were counted in each measurement, and the experiment was independently repeated three times. Data are mean ± SD. $^{**}P < 0.01$, Student's t test. F) The Phostag experiment validates the phosphorylation of FtMADS1/2 by FtPAK. G) Schematic diagram of the recombinant vector in the dual bioluminescence assay. H) The results of dual bioluminescence assay. The fusion proteins were expressed in *N. benthamiana* using agroinfiltration.

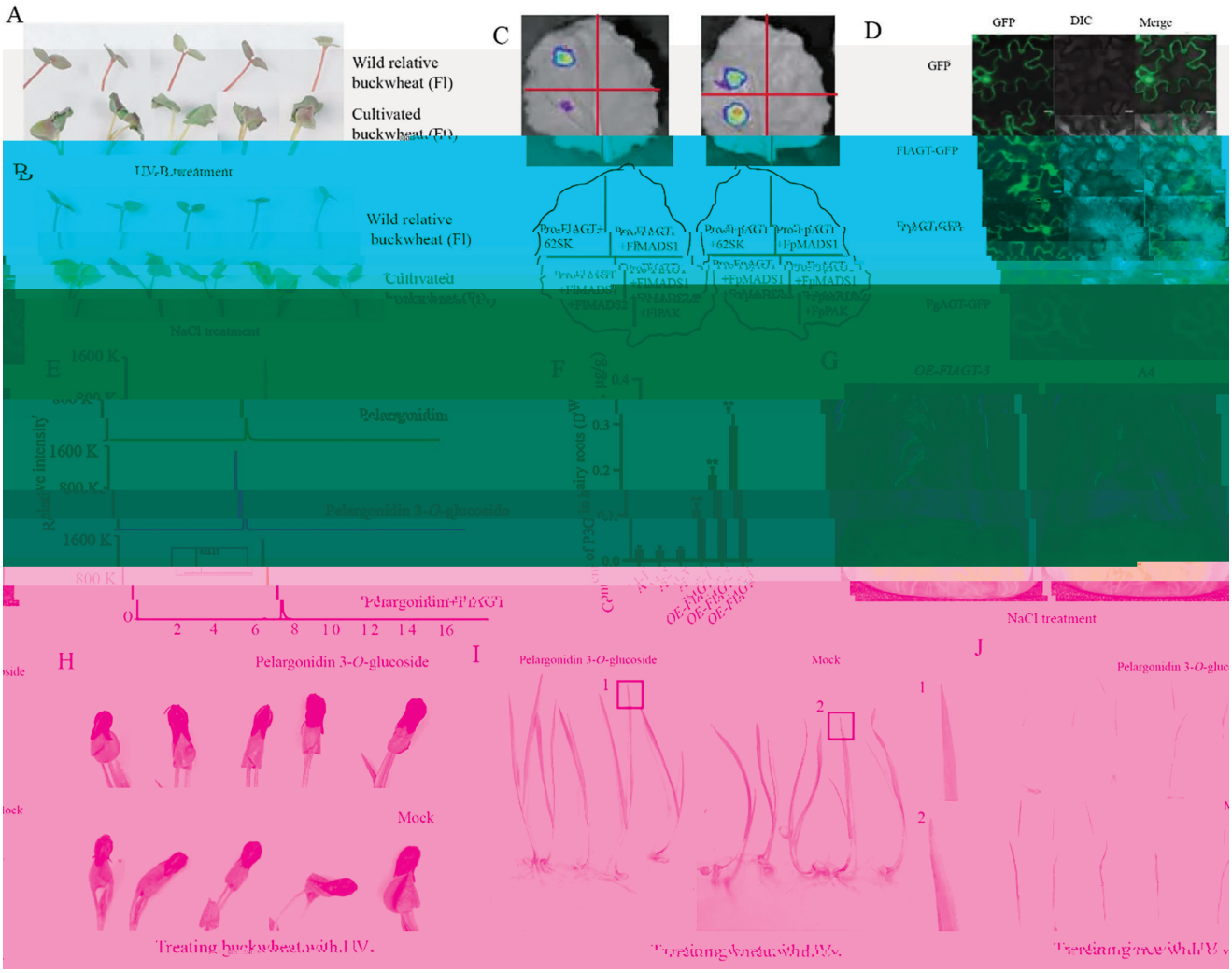

Figure 5. AGT contributes to the high-altitude adaptability of wild buckwheat. A) Wild buckwheat is more resistant to UV stress than cultivated buckwheat. Fl represents *F. longistylum*, Ft represents *F. tataticum*. B) Wild buckwheat is more resistant to NaCl stress than cultivated buckwheat. C) The results of dual-luciferase assay. The fusion proteins were expressed in *N. benthamiana* using agroinfiltration. Chemiluminescence images were captured 36 h after infiltration using 3 mg mL^{-1} luciferin. Similar results were obtained in three biological replicates. D) Subcellular localization of FlAGT, FpAGT, and FgAGT. FlAGT-GFP, pCAMBIA1300-*FlAGT* recombinant plasmid. FpAGT-GFP, pCAMBIA1300-*FpAGT* recombinant plasmid; FgAGT-GFP, pCAMBIA1300-*FgAGT* recombinant plasmid. Bar = 10 μm. E) Enzymatic assay for pelargonidin of FtAGT in vitro. The red boxes represent the enlarged areas. F) Detection of pelargonidin 3-*O*-glucoside expression in the overexpressed *FlAGT* hairy roots; P3G represents pelargonidin 3-*O*-glucoside. G) Phenotype of the overexpressed *FlAGT* hairy roots treated with 100 mM NaCl. H) Spraying pelargonidin 3-*O*-glucoside on the Tartary buckwheat seedlings, water was sprayed in the mock group. UV-B irradiation caused rust spots on the leaves. I) Spraying pelargonidin 3-*O*-glucoside on the leaves of wheat, water was sprayed in the mock group. UV-B irradiation caused rust spots on the leaves, and the red boxes with numbers represent the enlarged areas corresponding to the numbers on the right. J) Spraying pelargonidin 3-*O*-glucoside on the rice seedlings, water was sprayed in the mock group. UV-B irradiation caused rust spots on the seedlings.

UFGT3 expression and weakening its glycosylation activity on its substrates. Although the complex molecular regulatory mechanism of the *UFGT3* BGC in cultivated buckwheat is mainly conserved in wild relatives, an *AGT* gene is inserted between *PAK* and MADS2. AGT can glycosylate pelargonidin and is regulated in the same way as PAK, MADS1/2, similar to the regulation pattern of *UFGT3*. The addition of *AGT* enables the *UFGT3* gene cluster to have a more diversified synthesis and regulatory capability of flavonoids, endowing wild buckwheat relatives with a stronger adaptability to high-altitude climates.

3. Discussion

Genes relevant to plant secondary metabolism often exist in the form of BGCs in the genome. Studying BGCs can deepen our understanding of the biosynthetic mechanisms of metabolites. Their formation, domestication, and differentiation can significantly impact the stress resistance and ecological adaptability of plants. However, despite the increasing revelation of plant BGCs, the lack of genomic data has resulted in insufficient evolutionary dynamics analysis of BGCs within and between

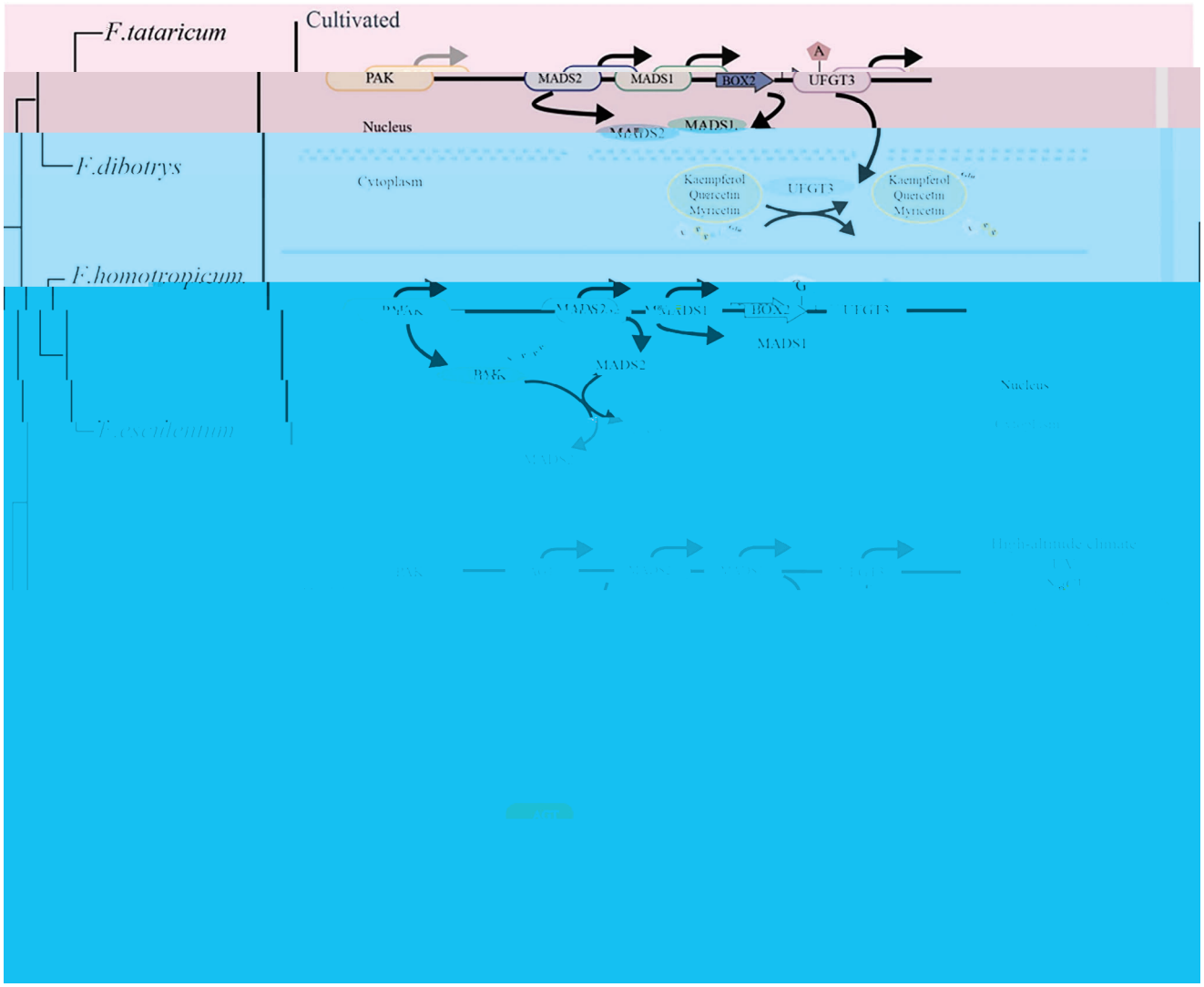

Figure 6. Molecular pattern schematics of *UFGT3* BGC in cultivated and wild relatives of buckwheat. In cultivated buckwheat, when the expression level of *PAK* is low, MADS2 can interact with MADS1, alleviating the transcriptional inhibitory effect of MADS1 on the expression of *UFGT3*, thus promoting the glycosylation activity of *UFGT3* on its substrates. However, when the expression level of *PAK* is high, PAK can phosphorylate MADS2, leading to the translocation of MADS2 from the nucleus to the cytoplasm. This means that MADS2 can no longer interact with MADS1, causing the MADS1 transcription factor to bind to the MADS-BOX2 element on the *UFGT3* promoter, thereby inhibiting *UFGT3* expression and weakening its glycosylation activity on its substrates. Similarly, the complex molecular regulatory mechanism of the *UFGT3* gene cluster in cultivated buckwheat also presents the same conservation in wild relatives. The difference is that in the wild relatives, an *AGT* gene is inserted between PAK and MADS2. AGT can glycosylate pelargonidin and is regulated in the same way as PAK, MADS1/2, similar to the regulation pattern of *UFGT3*. The additional *AGT* enables the *UFGT3* gene cluster to have a more diversified synthesis and regulatory capability of flavonoids, endowing wild buckwheat relatives with a stronger adaptability to high-altitude climates.

species.[31] In our study, we identified a glycosyltransferase gene, UFGT3, through GWAS and metabolomics data. By examining SNPs linked to nearby genes, we discovered and characterized a molecular module-type biosynthetic gene cluster (BGC), the UFGT3 gene cluster, in buckwheat. Comparative genomic analysis of seven Fagopyrum species (including both cultivated and wild relatives) revealed that this gene cluster is not only widely present within the genus but also influenced by domestication (Figures 1B and 2A) and evolved under environmental pressures (Figure 4N).

Intraspecifically, SNPs within BGCs often have significant impacts on the biosynthesis of metabolites. SNPs and genome-associated insertion/deletion events have been linked to the maintenance and diversification of BGCs in fungi.[42] In *Arabidopsis thaliana*, SNPs and small insertion/deletions are the most common sequence polymorphisms in the tryptophanol BGC.[43] A survey of haplotypes in the rice japonica and indica subtypes, along with *Oryza rufipogon Griff.*, showed that the intact rice diterpenoid gene cluster on chromosome 7 (*DGC7*), which encodes the entire biosynthetic pathway to 5,10-diketo-casbene,

For salt tolerance testing, *OE-FlAGT*-1, 2, and 3 lines with good growth on MS solid medium were transferred to liquid MS medium containing 100 mM NaCl. They were shaken as per the above method, and the phenotype was observed after two weeks. Hairy roots induced by A4 without the introduction of any vector were used as a control.

UV-B Treatment of Plants: Seed germination and seedling growth of all crops were conducted hydroponically. *Fagopyrum longistylum*, Tartary buckwheat (ZK3), wheat (ZM578), barley (HTX), and *F. homotropicum* were grown under a 16 h light/8 h dark cycle with 75% humidity. Rice (Japonica) seeds were germinated at 37 °C until sprouting and then transferred to the same cultivation conditions as the other crops. For UV light intervention experiments crops were exposed to 311 nm UV light (10 cm from the plants, 1.76 mw cm^{-2}) until UV stress symptoms appeared in seedlings. The UV-B irradiation experiments were conducted in a dark artificial climate chamber.

Diaminobenzidine (DAB) Staining: DAB (1 mg mL^{-1}) was dissolved in PBS buffer (pH 3.8) and mixed thoroughly. After the samples were rinsed with phosphate buffer solution (PBS), they were subsequently immersed in the DAB staining solution prior to being shaken horizontally at 40 rpm for 6 h. The degree of staining was observed following decolorization with 95% alcohol.

Treatment of Plants with Flavonoids—Laboratory Experiment: All plant seeds were germinated and grown to the appropriate stages as described in the UV-B treatment of plants. One hour prior □□ □ □ □.□□□□ Tm7rimenttreatmentoprior□□□□□6 w□□□.□(,.□□□□-□(C)-□□□.Q□(2)-□□(C)-□□□.Q□(2)-□.□□□r

Vazyme, Nanjing, China). The quantitative reverse transcription polymerase chain reaction (qRT-PCR) was performed using the Taq Pro Universal SYBR qPCR Master Mix (Q712, v20.1, Vazyme, China) in accordance with the manufacturer's instructions. Primers are depicted in Data Set 2 (Supporting Information).

Luciferase Complementation Imaging Assay (LCA): For LCA assays, the full length of the target protein was amplified using specific primers (Data Set 2, Supporting Information) and introduced into pCAMBIA1300-cLUC or pCAMBIA1300-nLUC. The recombinant vectors were transformed into GV3101 and then co-transformed into the *N. benthamiana leaves*. Luminescence signals from pavement cells were detected after applying 3 mg mL^{-1} luciferin by a charge-coupled device (CCD) system (plant in vivo imaging system). The luminescence from LUC was detected using a Mithras LB940 microplate reader52..

Statistical Analysis: GraphPad Prism 8.0 and SPSS22 were employed for conducting the statistical analysis. The statistical significance of the observed differences was assessed using Student's t-test (*$P < 0.05$; **$P < 0.01$). Run molecular docking calculations using AMDock, and visualize and analyze the results using PyMOL.

, 122□□□□.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

X.H., Y.H., K.Z., Y.S., H.Z., D.L., H.L., and X.W. contributed equally to this work This research was supported by the National Key R&D Program of China (2023YFF1002500), National Natural Science Foundation of China (32161143005, 32241042, 32301794), the Youth Innovation Program of Chinese Academy of Agricultural Sciences (Y2022QC02), the European Union's Horizon 2020 research and innovation programme, project PlantaSYST (SGA-CSA No. 739582 under FPA No. 664620), and the BG05M2OP001-1.003-001-C01 project, financed by the European Regional Development Fund through the Bulgarian "Science and Education for Smart Growth" Operational Programme to M.I.G. and A.R.F.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

M.Z. designed and managed the project. M.Z., Y.H., and K.Z. organized the funding for this research. D.L. and W.L. provided the genetic materials. X.H., Y.H., K.Z., Y.S., H.L., and X.W. performed data analysis and figure design. X.H., H.Z., D.L., Z.Y., Y.X., and Y.O. performed most of the experiments. X.H., Y.H., K.Z., S.H.W., M.Q., M.I.G., and A.R.F. wrote the manuscript. All authors read and approved the manuscript.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords

buckwheat, evolution, flavonoid metabolism, novel gene cluster, UV resistance

Received: April 6, 2024
Revised: August 15, 2024
Published online:

[1] S. Cheng, W. Xian, Y. Fu, B. Marin, J. Keller, T. Wu, W. Sun, X. Li, Y. Xu, Y. u. Zhang, S. Wittek, T. Reder, G. Günther, A. Gontcharov, S. Wang, L. Li, X. Liu, J. Wang, H. Yang, X. Xu, P. M. Delaux, B. Melkonian, G. K. a.-S. Wong, M. Melkonian, *Cell* 2019, *179*, 1057.
[2] G. Agati, E. Azzarello, S. Pollastri, M. Tattini, *Plant Sci.* 2012, *196*, 67.
[3] M. L. Falcone Ferreyra, S. P. Rius, P. Casati, *Front Plant Sci.* 2012, *3*, 222.
[4] D. Zhao, Y. Zhang, Y. Lu, L. Fan, Z. Zhang, J. Zheng, M. Chai, *J. Genet Genomics 49*, 547.
[5] S. Hussain, M. J. Rao, M. A. Anjum, S. Ejaz, U.-U. d.-D. Umar, M. A. Ali, M. F. Khalid, M. Sohail, S. Ercisli, M. Zia-Ul-Haq, S. Ahmad, S. A. H. Naqvi, *AMB Express* 2019, *9*, 147.
[6] D. Šamec, E. Karalija, I. Šola, V. Vujčić Bok, B. Salopek-Sondi, *Plants* (Basel) 2021, *10*, 118.
[7] S. Martens, A. Mithöfer, *Phytochemistry* 2005, *66*, 2399.

M. S. Motawia, B. Hamberger, B. L. Møller, J. E. Bassard, *Sci.* 2020, *354*, 890.
[31] H. W. Nützmann, A. Huang, A. Osbourn, *New Phytol.* 2016, *211*, 771.
[32] Z. Tian, J. Jia, B. o. Yin, W. Chen, *J Genet Genomics* 2024, *51*, 714.
[33] K. Zhang, M. He, Y. u. Fan, H. Zhao, B. Gao, K. Yang, F. Li, Y. u. Tang, Q. Gao, T. Lin, M. Quinet, D. Janovská, V. Meglic, J. Kwiatkowski, O. Romanova, N. Chrungoo, T. Suzuki, Z. Luthar, M. Germ, S. H. Woo, M. I. Georgiev, M. Zhou, *Genome Biol.* 2021, *22*, 1.
[34] H. Zhao, Y. He, K. Zhang, S. Li, Y. Chen, M. He, F. He, B. Gao, D. i. Yang, Y. u. Fan, X. Zhu, M. Yan, N. Giglioli-Guivarc'h, C. Hano, A. R. Fernie, M. I. Georgiev, D. Janovská, V. Meglic, M. Zhou, *Plant Biotechnol J* 2023, *21*, 150.
[35] J. Zhou, C. L. Li, F. Gao, X. P. Luo, Q. Q. Li, H. X. Zhao, H. P. Yao, H. Chen, A. n.-H. u. Wang, Q. i. Wu, *J. Agric. Food Chem.* 2016, *64*, 6930.
[36] H. Fujita, M. Takemura, E. Tani, K. Nemoto, A. Yokota, T. Kohchi, *Plant Cell Physiol.* 2003, *44*, 735.
[37] W. Yu, F. Peng, Y. Xiao, G. Wang, J. Luo, *Front Plant Sci* 2018, *9*, 1856.
[38] M. He, Y. He, K. Zhang, X. Lu, X. Zhang, B. Gao, Y. u. Fan, H. Zhao, R. Jha, M. d. N. Huda, Y. u. Tang, J. Wang, W. Yang, M. Yan, J. Cheng, J. Ruan, E. Dulloo, Z. Zhang, M. I. Georgiev, M. A. Chapman, M. Zhou, *New Phytol.* 2022, *235*, 1927.
[39] H. Lin, Y. Yao, P. Sun, L. Feng, S. Wang, Y. Ren, X. i. Yu, Z. Xi, J. Liu, *BMC Biol.* 2023, *21*, 87.
[40] K. Zhang, Y. He, X. Lu, Y. Shi, H. Zhao, X. Li, J. Li, Y. Liu, Y. Ouyang, Y. u. Tang, X. Ren, X. Zhang, W. Yang, Z. Sun, C. Zhang, M. Quinet, Z. Luthar, M. Germ, I. Kreft, D. Janovská, V. Meglic, B. Pipan, M. I. Georgiev, B. Studer, M. A. Chapman, M. Zhou, *Mol. Plant* 2023, *16*, 1427.
[41] Y. Li, Z. Wang, M. Zhu, Z. Niu, M. Li, Z. Zheng, H. Hu, Z. Lu, J. Zhang, D. Wan, Q. Chen, Y. Yang, *Commun Biol* 2023, *6*, 867.
[42] A. L. Lind, J. H. Wisecaver, C. Lameiras, P. Wiemann, J. M. Palmer, N. P. Keller, F. Rodrigues, G. H. Goldman, A. Rokas, *PLoS Biol.* 2017, *15*, e2003583.
[43] Z. Liu, J. Cheema, M. Vigouroux, L. Hill, J. Reed, P. Paajanen, L. Yant, A. Osbourn, *Nat. Commun.* 2020, *11*, 5354.
[44] C. Zhan, L. Lei, Z. Liu, S. Zhou, C. Yang, X. Zhu, H. Guo, F. Zhang, M. Peng, M. Zhang, Y. Li, Z. Yang, Y. Sun, Y. Shi, K. Li, L. Liu, S. Shen, X. Wang, J. Shao, X. Jing, Z. Wang, Y. Li, T. Czechowski, M. Hasegawa, J. Luo, *Nat. Plants* 2020, *6*, 1447.
[45] A. Rai, H. Hirakawa, R. Nakabayashi, S. Kikuchi, K. Hayashi, M. Rai, H. Tsugawa, T. Nakaya, T. Mori, H. Nagasaki, R. Fukushi, Y. Kusuya, H. Takahashi, H. Uchiyama, A. Toyoda, S. Hikosaka, E. Goto, K. Saito, M. Yamazaki, *Nat. Commun.* 2021, *12*, 405.
[46] S. Yeaman, M. C. Whitlock, *Evolution* 2011, *65*, 1897.
[47] L. M. Schneider, N. M. Adamski, C. E. Christensen, D. B. Stuart, S. Vautrin, M. Hansson, C. Uauy, P. von Wettstein-Knowles, *J. Exp. Bot.* 2016, *67*, 2715.
[48] A. A. Brakhage, *Nat. Rev. Microbiol.* 2013, *11*, 21.
[49] G. C. Uguru, K. E. Stephens, J. A. Stead, J. E. Towle, S. Baumberg, K. J. McDowall, *Mol. Microbiol.* 2005, *58*, 131.
[50] J. Lang, Y. Fu, Y. Zhou, M. Cheng, M. Deng, M. Li, T. Zhu, J. Yang, X. Guo, L. Gui, L. Li, Z. Chen, Y. Yi, L. Zhang, M. Hao, L. Huang, C. Tan, G. Chen, Q. Jiang, J. Wang, (2021). Myb10-D confers PHS-3D resistance to pre-harvest sprouting by regulating NCED in ABA biosynthesis pathway of wheat. New Phytologist, *230*(5), 1940. *Portico*. https://doi.org/10.1111/nph.17312
[51] H. Kim, N. Lee, Y. Kim, G. Choi, (2024). The phytochrome-interacting factor genes PIF1 and PIF4 are functionally diversified due to divergence of promoters and proteins. The Plant Cell, *36*(8), 2778. https://doi.org/10.1093/plcell/koae110
[52] Q. Sun, Z. He, D. Feng, R. Wei, Y. Zhang, J. Ye, L. Chai, J. Xu, Y. Cheng, Q. Xu, X. Deng, (2024). An abscisic acid-responsive transcriptional regulatory module CsERF110-CsERF53 orchestrates citrus fruit coloration. Plant Communications, 101065. https://doi.org/10.1016/j.xplc.2024.101065