Figure S1

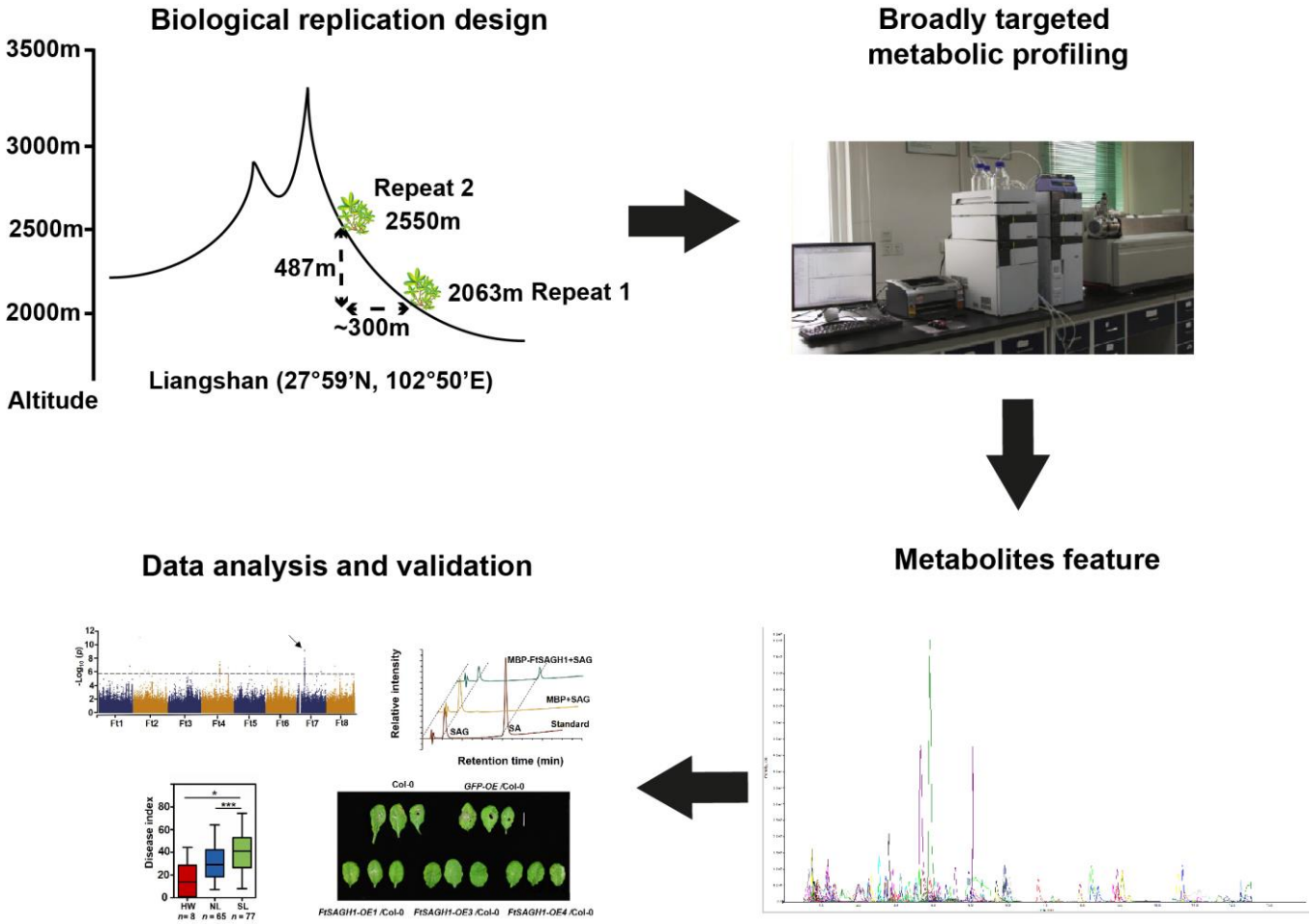


Figure S1. Workflow of this study.

Workflow of this study, which including plant cultivation (the sowing was 10 days later at higher altitude), LC-MS/MS, metabolite extraction, data processing, and data mining.

Figure S2

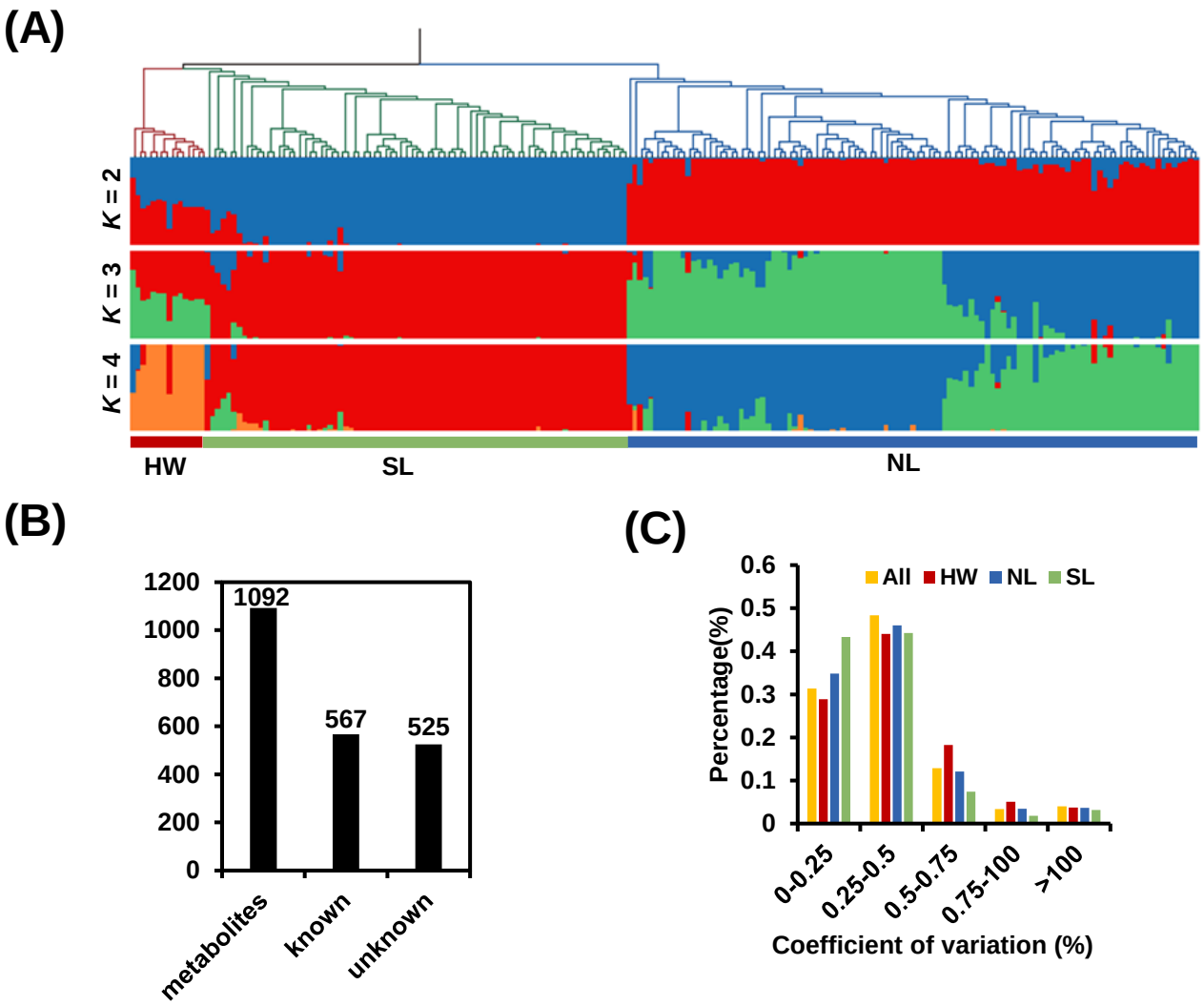


Figure S2. Characters of metabolome in three groups of Tartary buckwheat.

(A) The population structure analysis of 200 accessions calculated from whole genome resequencing SNPs. 200 accessions included 14 Himalaya wild accessions, 80 SL accessions and 106 NL accessions. (B) Quantity statistics of metabolites detected in Tartary buckwheat metabolome. (C) Coefficients of variation (CV) for HW, SL, NL and whole group, respectively.

Figure S3

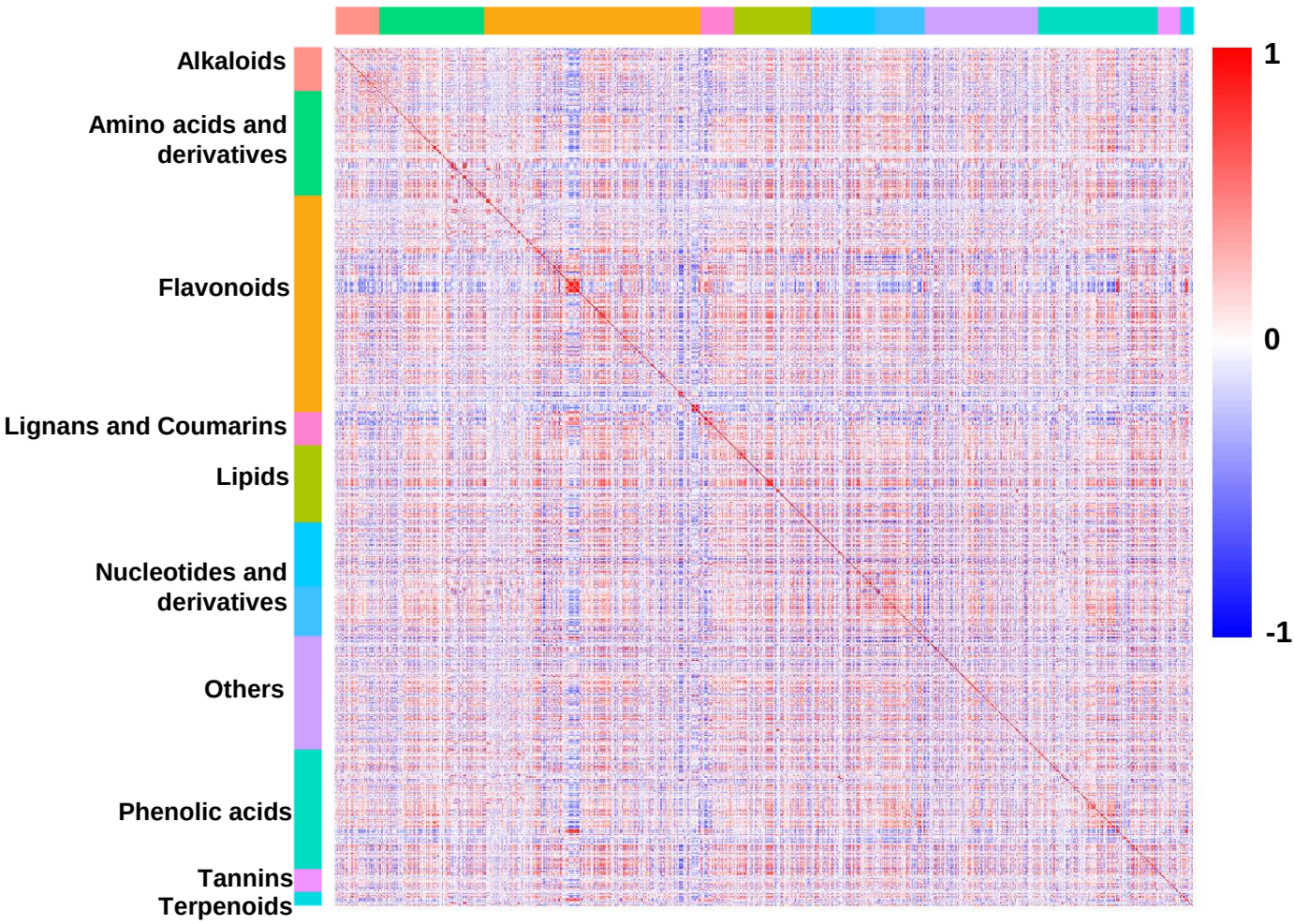


Figure S3. Correlation analyses of metabolites.

The metabolites were clustered based on their distribution patterns among the 200 Tartary buckwheat accessions; the red-to-blue colors for each cell represent the two corresponding metabolites sharing high-to-low similarly distributed patterns (indicated by the coefficient values) among buckwheat accessions.

Figure S4

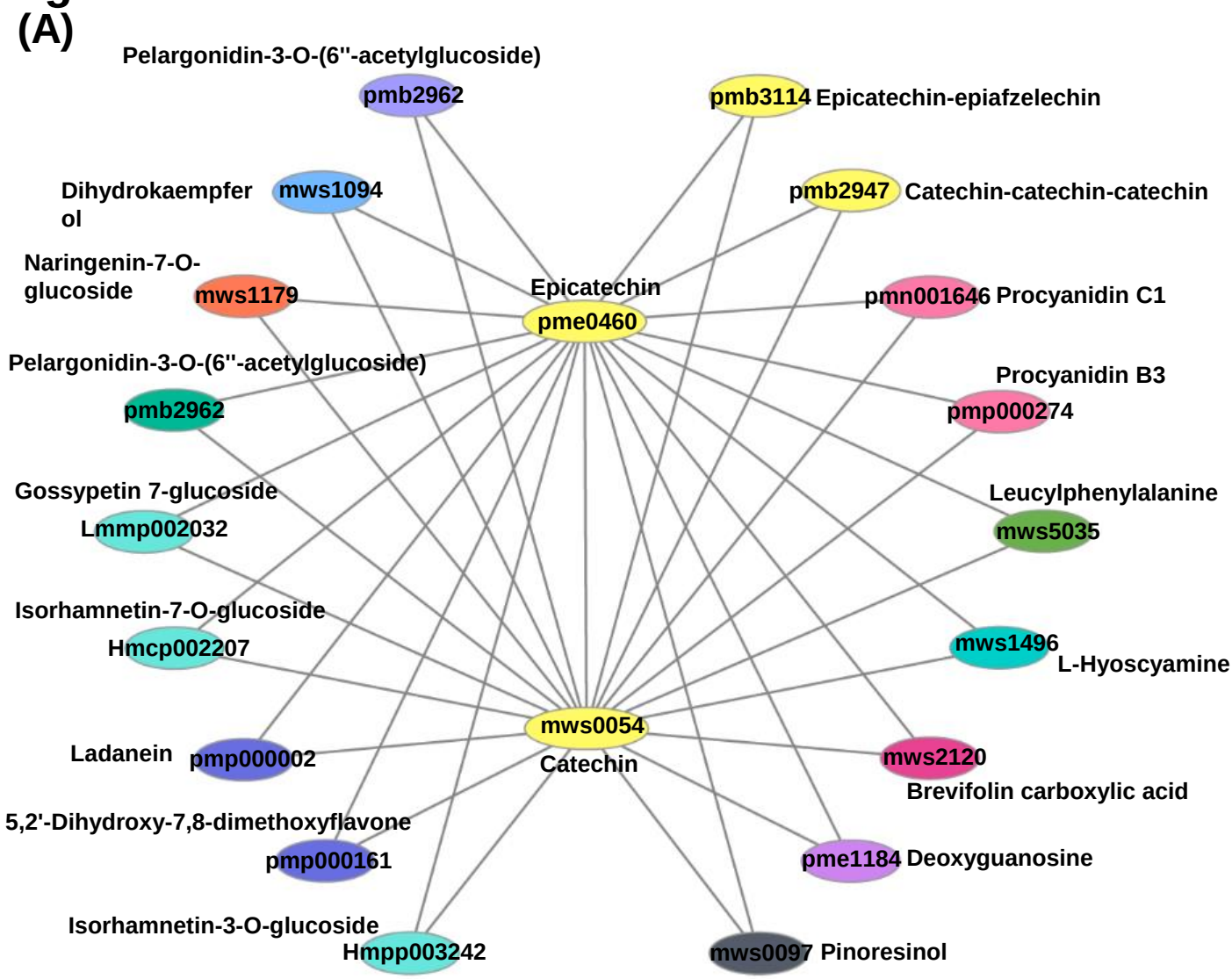
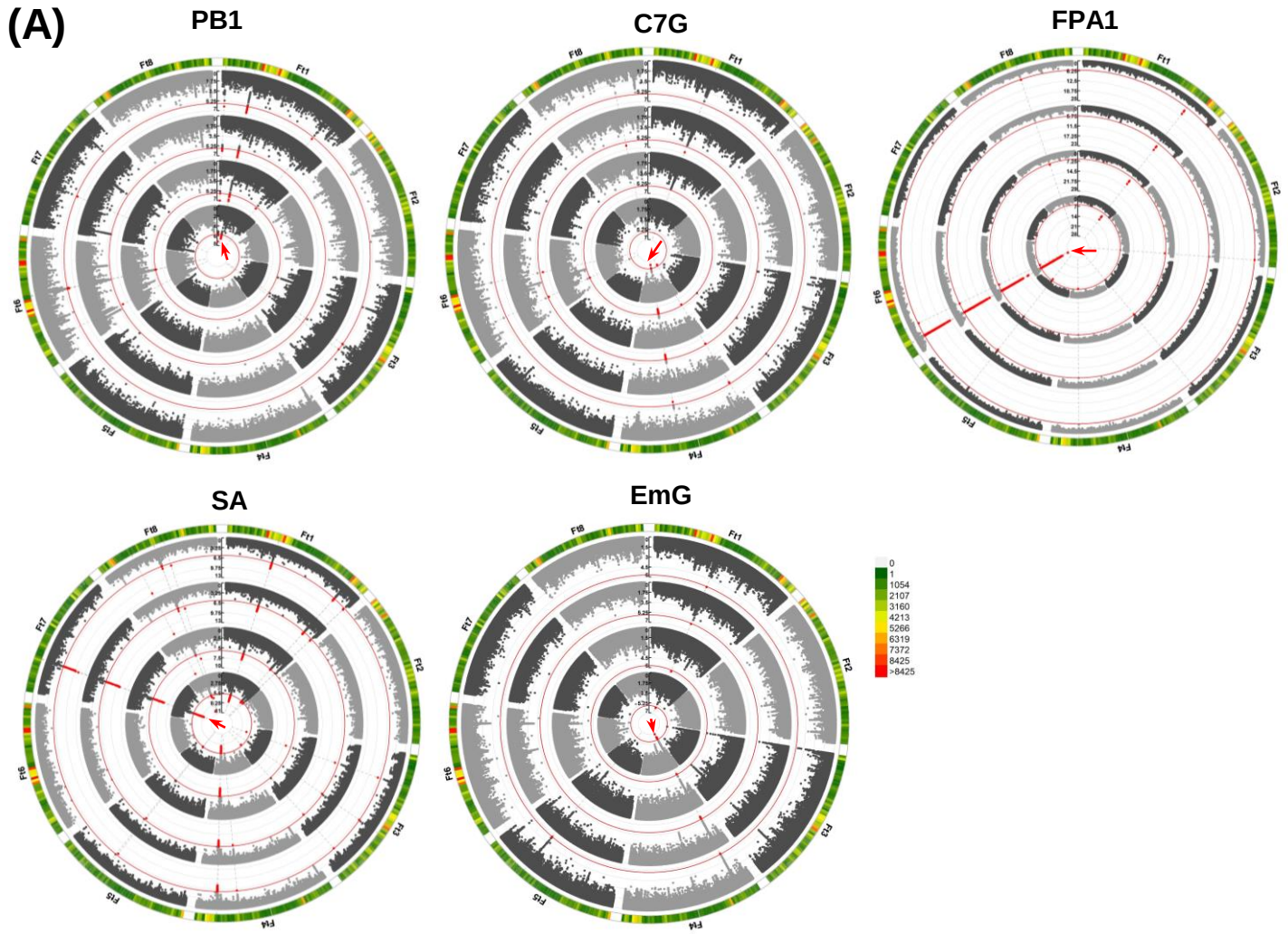

(A)
Pelargonidin-3-O-(6''-acetylglucoside)
pmb2962
pmb3114
Epicatechin-epiafzelechin
Dihydrokaempferol
mws1094
pmb2947
Catechin-catechin-catechin
Naringenin-7-O-glucoside
mws1179
Epicatechin
pme0460
pmn001646
Procyanidin C1
Pelargonidin-3-O-(6''-acetylglucoside)
pmb2962
Procyanidin B3
pmp000274
Gossypetin 7-glucoside
Lmmp002032
Leucylphenylalanine
mws5035
Isorhamnetin-7-O-glucoside
Hmcp002207
mws1496
L-Hyoscyamine
mws0054
Catechin
Ladanein
pmp000002
mws2120
Brevifolin carboxylic acid
5,2'-Dihydroxy-7,8-dimethoxyflavone
pmp000161
pme1184
Deoxyguanosine
Isorhamnetin-3-O-glucoside
Hmpp003242
mws0097
Pinoresinol
(B)
Kaempferol-3-O-glucosyl-rhamnosyl-glucoside-glucoside
Quercetin-3-O-robinobioside
Nicotiflorin
Quercetin-3-O-rutinoside (Rutin)
(C)
D-(+)-Melezitose
Fagopyritol B2
Fagopyritol B3
Fagopyritol A2
Maltotetraose
Fagopyritol A3
1,1-Kestotetraose
Panose
Alkaloids
Amino acids and derivatives
Flavonoids
Lipids
Lignans and Coumarins
Organic acids
Tannins
Nucleotides and derivatives
Phenolic acids
Others

Figure S4. Networks for important substances in Tartary buckwheat. Networks for catechin and epicatechin **(A)** rutin **(B)** and fagopyritols **(C)** in Tartary buckwheat. Lines indicated the linking metabolites shared similar distribution patterns among the 200 buckwheat accessions, wherein the thicker lines represented higher coefficient values. The correlation values for metabolites were listed in **Table S4**.

Figure S5


(A)
PB1
C7G
FPA1
SA
EmG
0
1
1054
2107
3160
4213
5266
6319
7372
8425
>8425

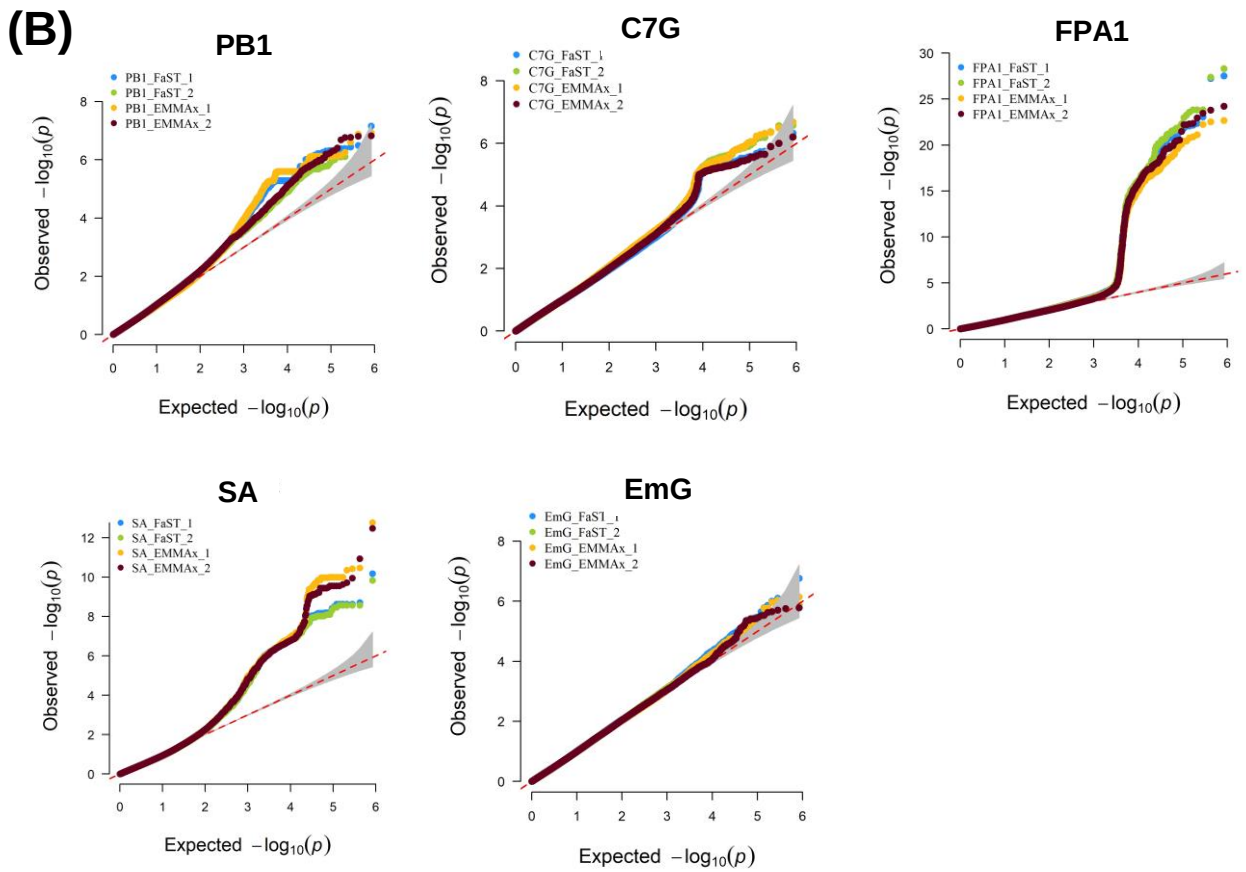

(B)
PB1
PB1_FaST_1
PB1_FaST_2
PB1_EMMAx_1
PB1_EMMAx_2
C7G
C7G_FaST_1
C7G_FaST_2
C7G_EMMAx_1
C7G_EMMAx_2
FPA1
FPA1_FaST_1
FPA1_FaST_2
FPA1_EMMAx_1
FPA1_EMMAx_2
SA
SA_FaST_1
SA_FaST_2
SA_EMMAx_1
SA_EMMAx_2
EmG
EmG_FaST_1
EmG_FaST_2
EmG_EMMAx_1
EmG_EMMAx_2
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

Figure S5. Mul-Manhattan plots and Quantile-Quantile (Q-Q) plots of the two biological replicates main metabolites based on the FaST-LMM and EMMAx model.

(A) Genome-wide *p*-values for the two biological replicates main metabolites based on the FaST-LMM and EMMAx model. The dashed line indicates the threshold $-\log_{10} p = 5$. Arrows indicate mGWAS signals overlapped between EMMAx and FaST-LMM analysis. (B) Q-Q plot of the expected null distribution and the observed *p*-value for the two biological replicates main metabolites using the FaST-LMM and EMMAx model.

Figure S6

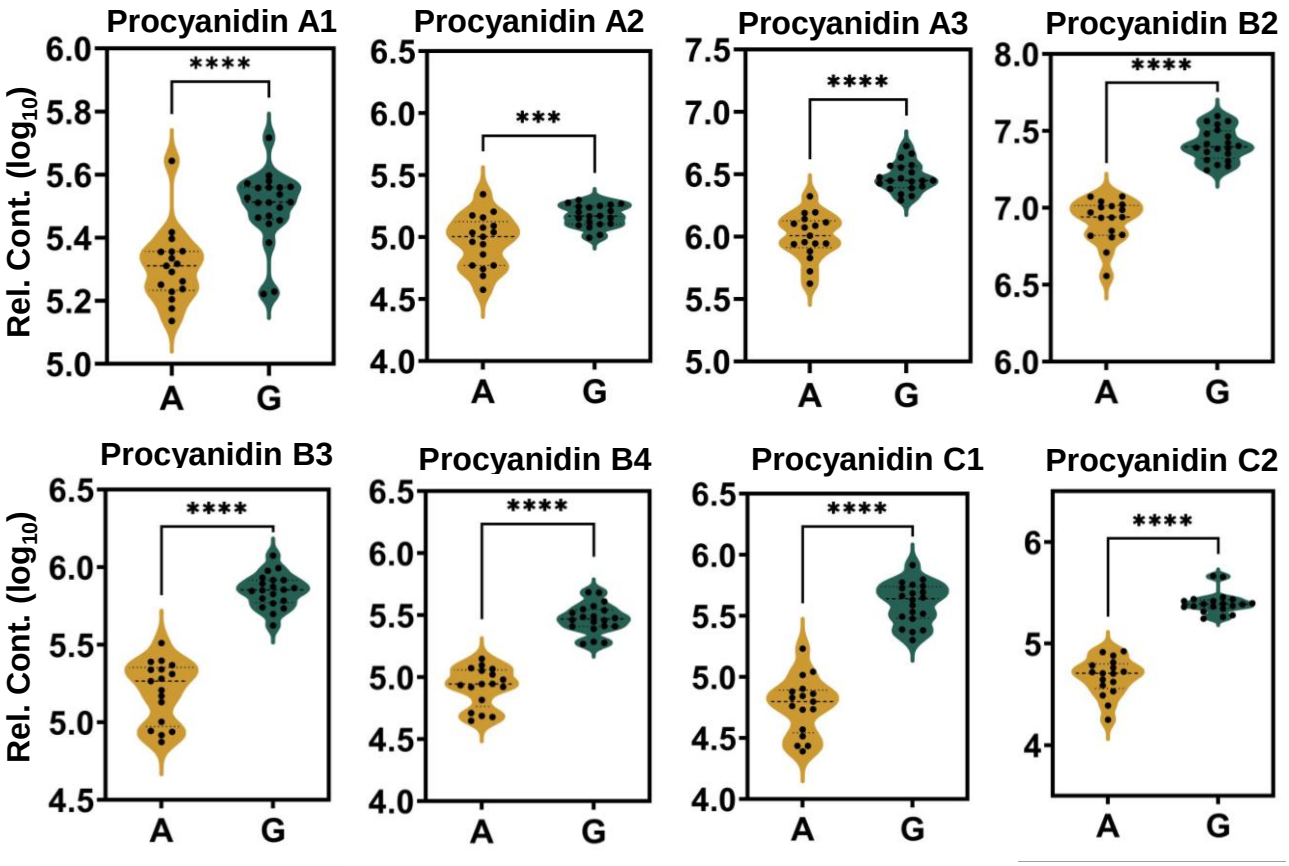
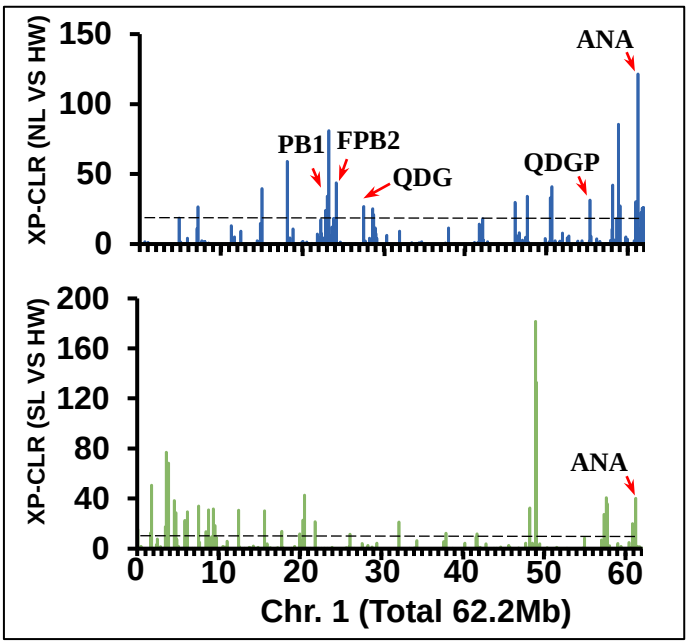
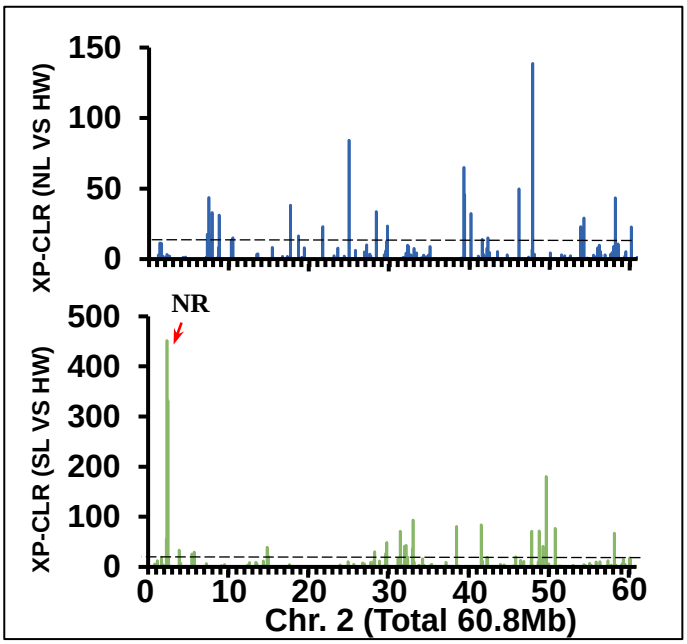


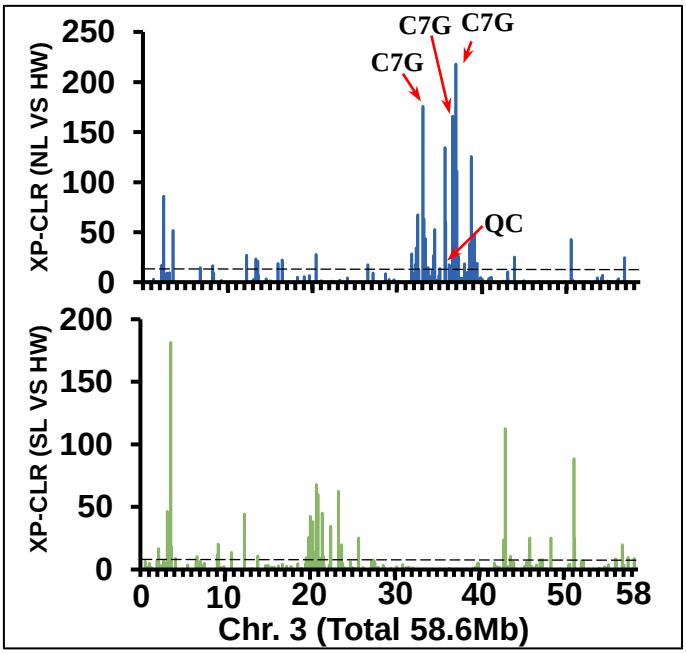
Figure S6. FtMYB43 transcription factor is involved in procyanidins biosynthesis in Tartary buckwheat.

Relative content of procyanidins in different alleles accessions of SNP Ft1:13,164,411. The metabolic data of procyanidins was $\log_{10}$ -transformed. *nA*=17,*nG*=20 for all plots.

Figure S7


XP-CLR (NL VS HW)
150
100
50
0
ANA
PB1
FPB2
QDG
QDGP
XP-CLR (SL VS HW)
200
160
120
80
40
0
ANA
0
10
20
30
40
50
60
Chr. 1 (Total 62.2Mb)


XP-CLR (NL VS HW)
150
100
50
0
XP-CLR (SL VS HW)
500
400
300
200
100
0
NR
0
10
20
30
40
50
60
Chr. 2 (Total 60.8Mb)


XP-CLR (NL VS HW)
250
200
150
100
50
0
C7G
C7G
C7G
QC
XP-CLR (SL VS HW)
200
150
100
50
0
0
10
20
30
40
50
58
Chr. 3 (Total 58.6Mb)

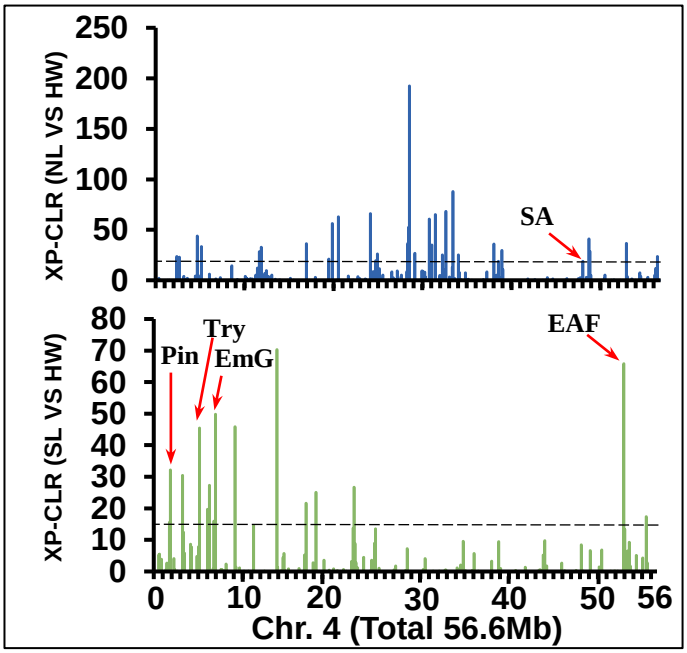

XP-CLR (NL VS HW)
250
200
150
100
50
0
SA
XP-CLR (SL VS HW)
80
70
60
50
40
30
20
10
0
Pin
Try
EmG
EAF
0
10
20
30
40
50
56
Chr. 4 (Total 56.6Mb)

Figure S7 (continued)

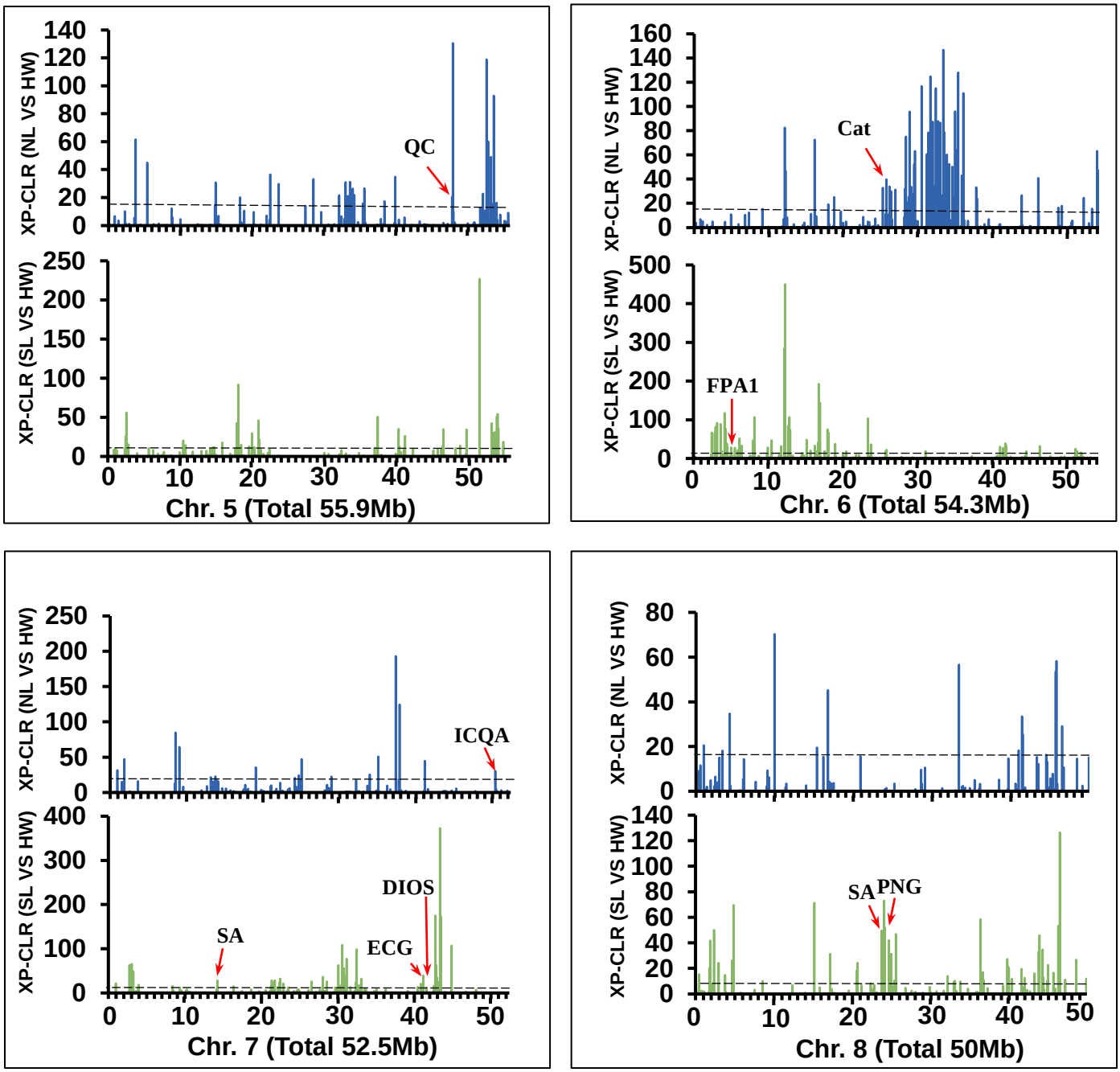


Figure S7. Selective signals in domestication of NL (blue) and in SL (green) on each chromosome. The cross population composite likelihood ratio (XP-CLR) between NL and HW (upper), and between SL and HW (lower) on each chromosome. The black dashed horizontal lines indicate top 10% threshold for each chromosome (threshod with 18 in NL/HW and 14 in SL/HW). The red arrows indicate the position of mGWAS signal loci in the sweeps. PB1, procyanidin B1; FPB2, fagopyritol B2; QDG, quercetin-3,7-*O*-β-D-diglucoside; QDGP, quercetin-3-*O*-(2",3"-digalloyl)-β-D-glucopyranoside; ANA, anchoic acid; NR, nicotinate D-ribonucleoside; C7G, (+)-catechin-7-*O*-glucoside; QC, quercetin; Pin, pinoresinol; Try, tryptamine; EmG, emodin-8-*O*-β-D-glucoside; SA, salicylic acid; EAF, epiafzelechin; FPA1, fagopyritol A1; Cat, catechin; ECG, epicatechin glucoside; DIOS, diosmetin; ICQA, isochlorogenic acid A and PNG, pyridoxine-5'-*O*-glucoside. The lengths of different chromosomes were annotated in the figure.

Figure S8

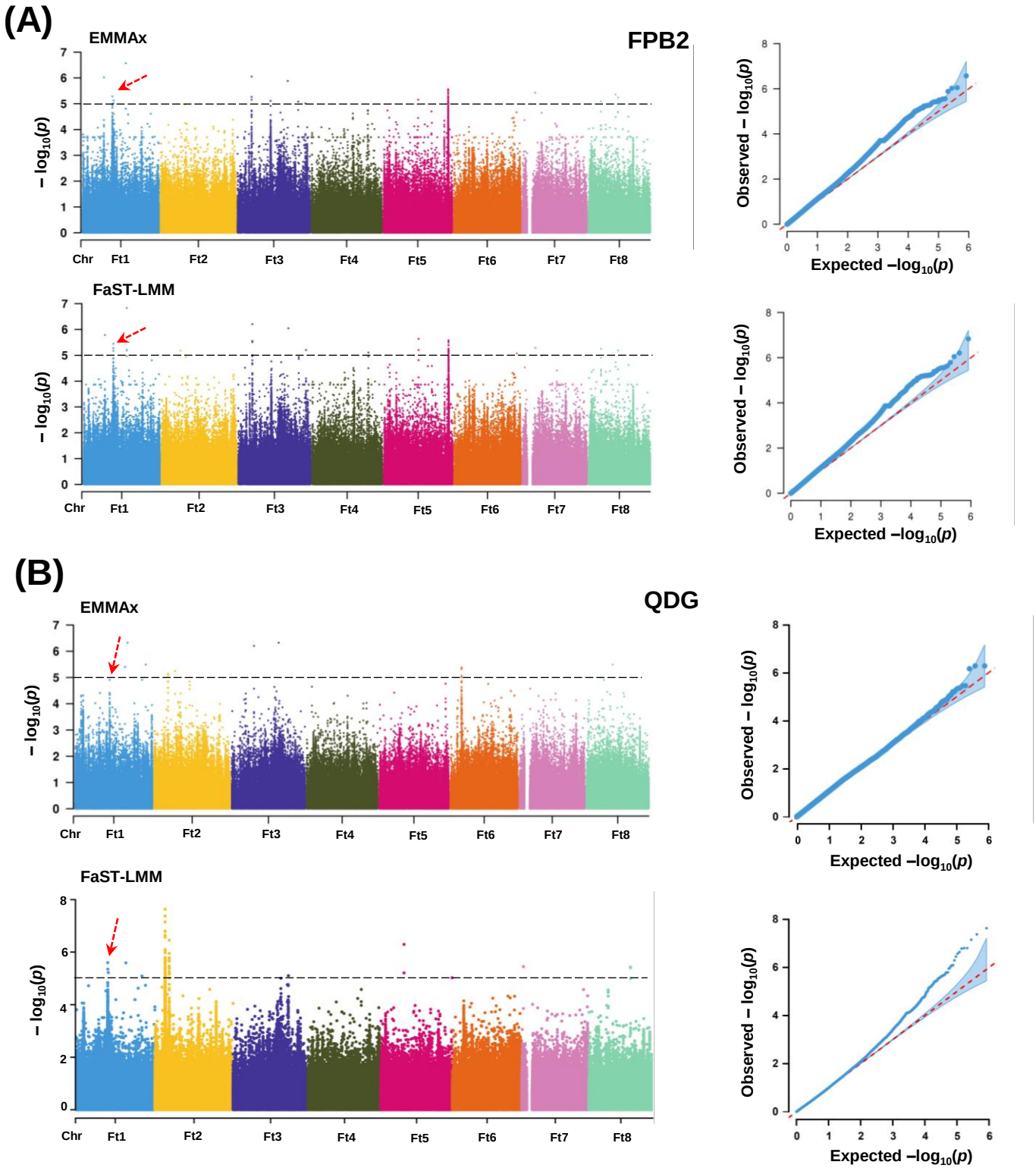

(A)
EMMAx
FPB2
FaST-LMM
(B)
EMMAx
QDG
FaST-LMM
– log10(p)
Observed – log10(p)
Expected –log10(p)
Chr
Ft1
Ft2
Ft3
Ft4
Ft5
Ft6
Ft7
Ft8

Figure S8 (continued)

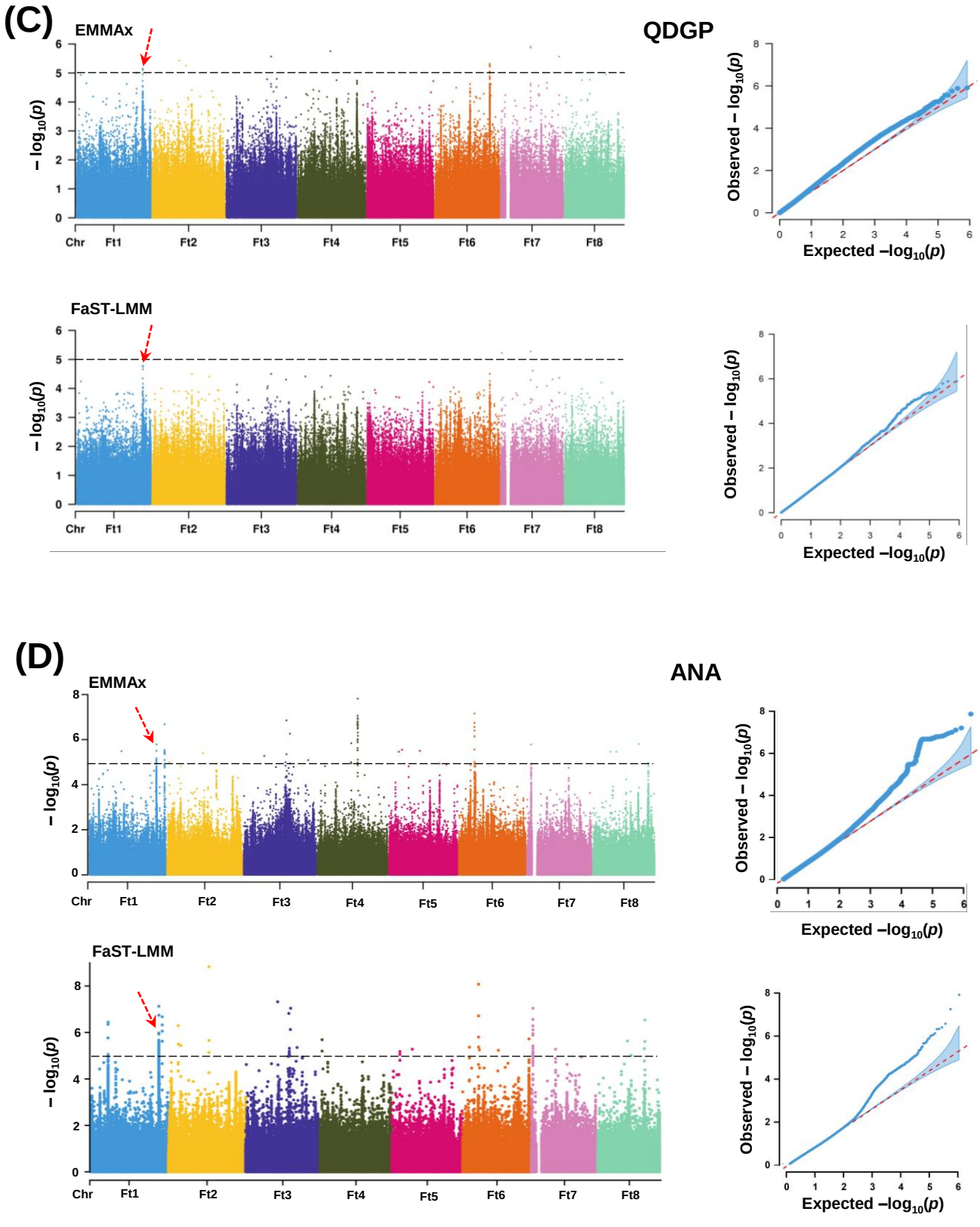
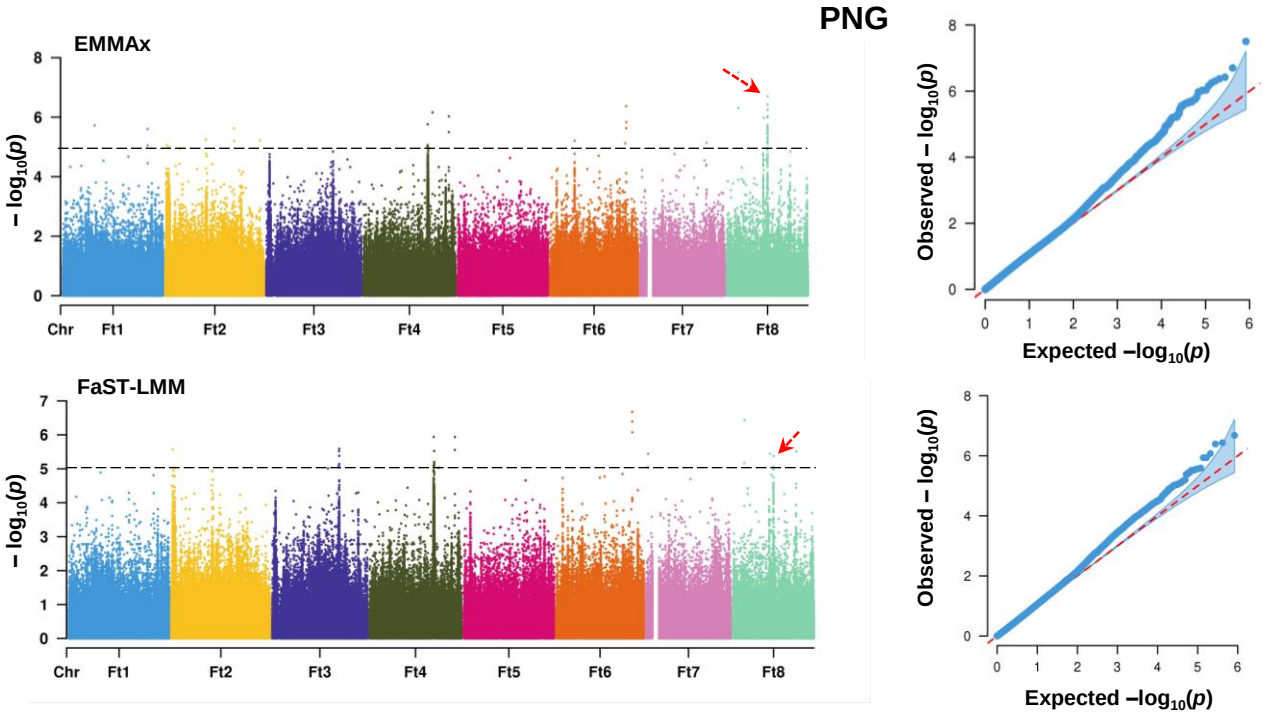


Figure S8 (continued)

(E)



(F)

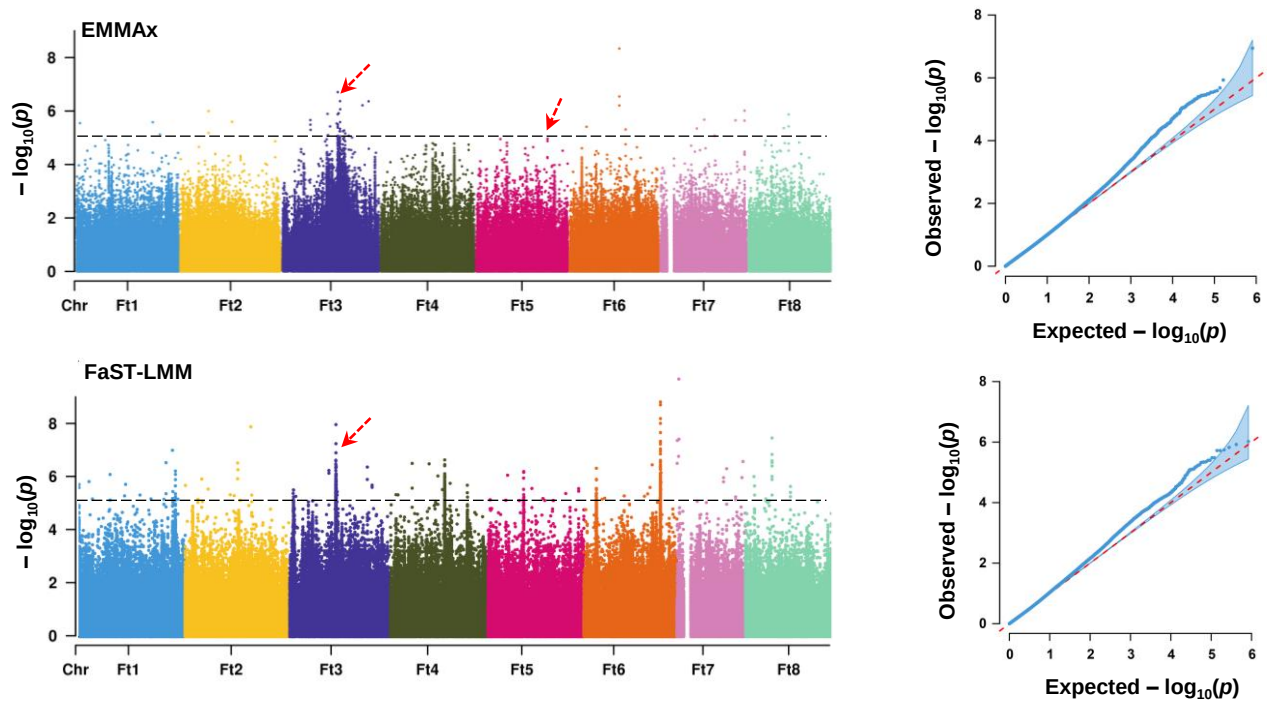
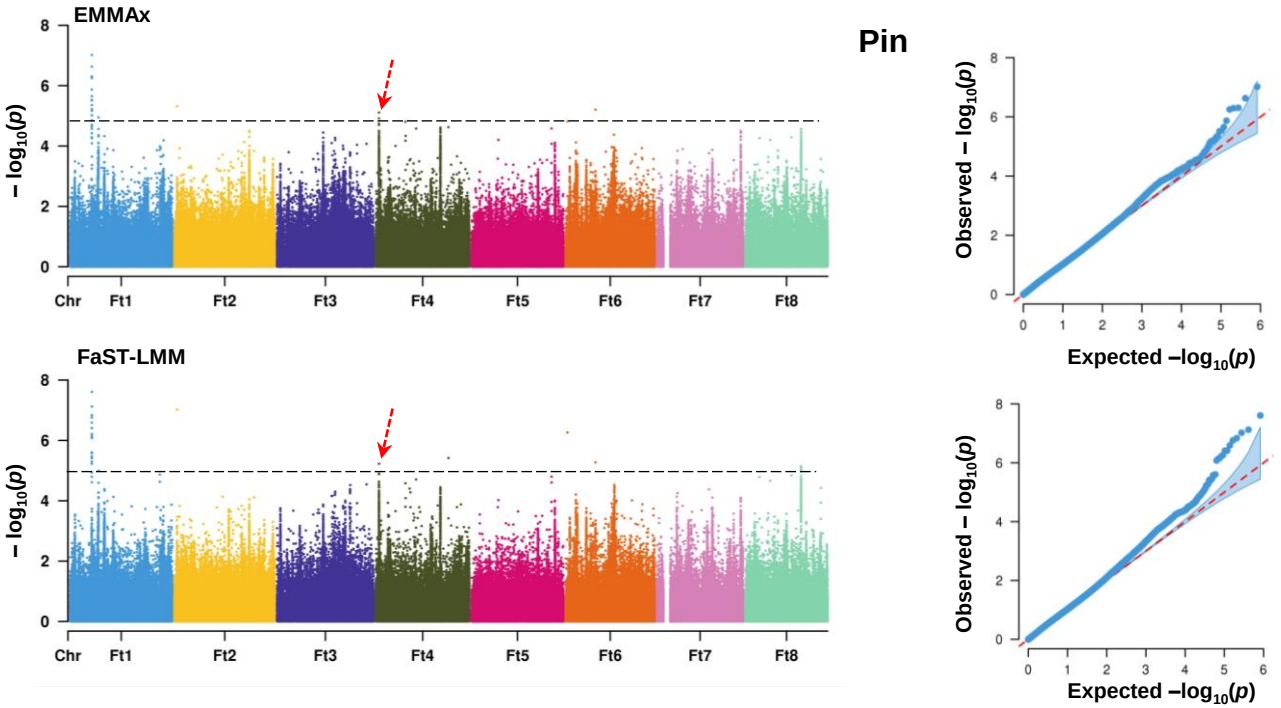


Figure S8 (continued)

(G)



(H)

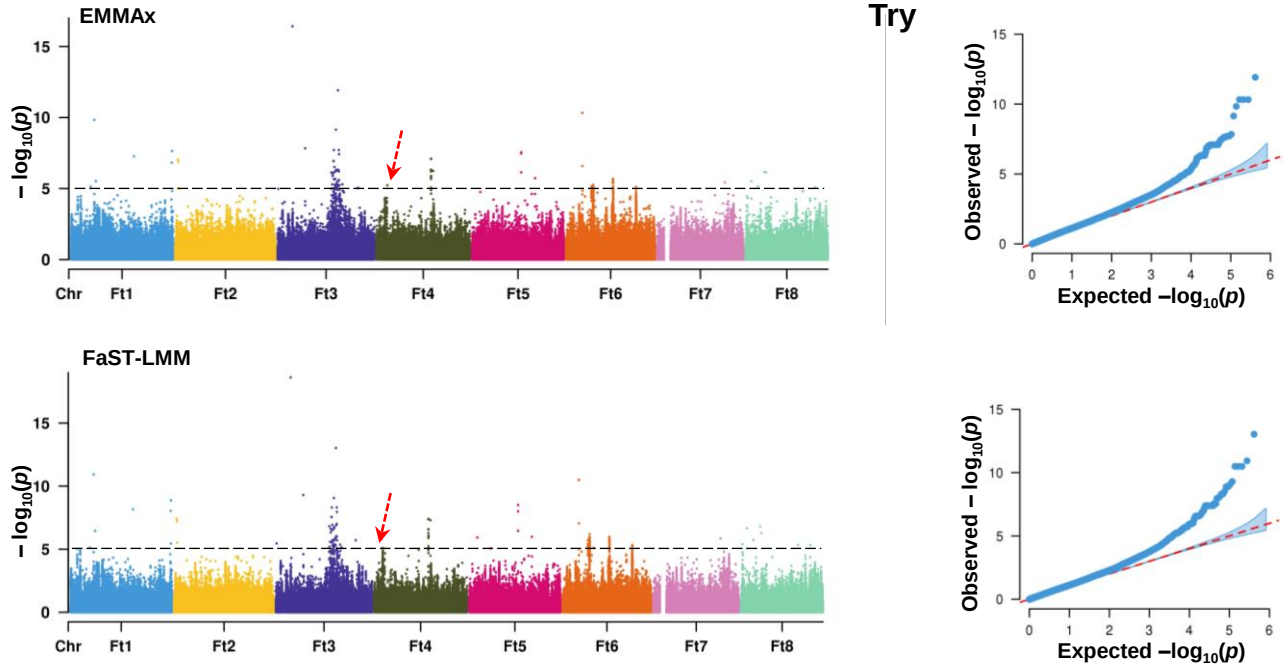


Figure S8 (continued)

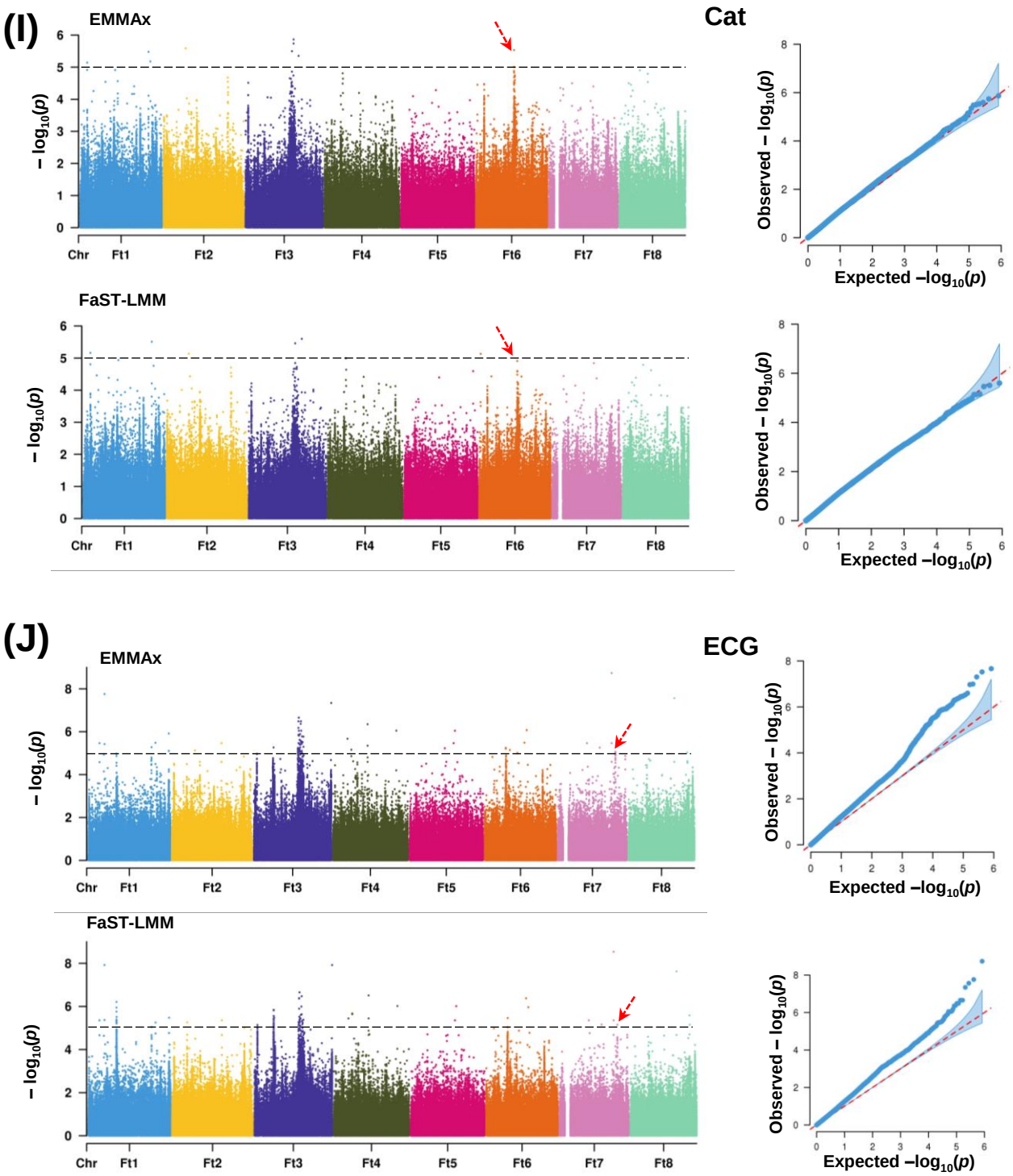
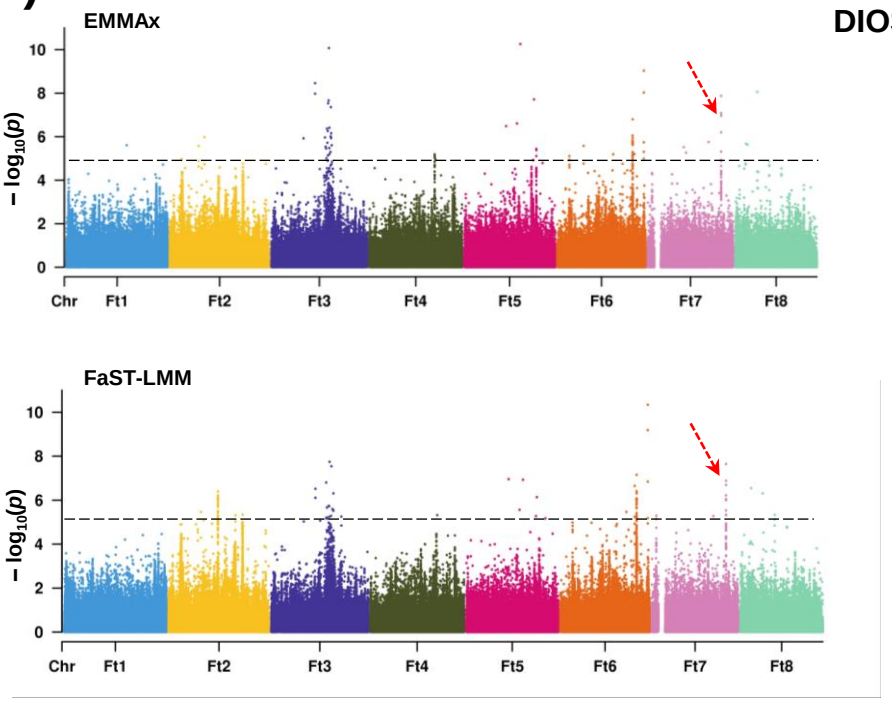


Figure S8 (continued)

(K)



(L)

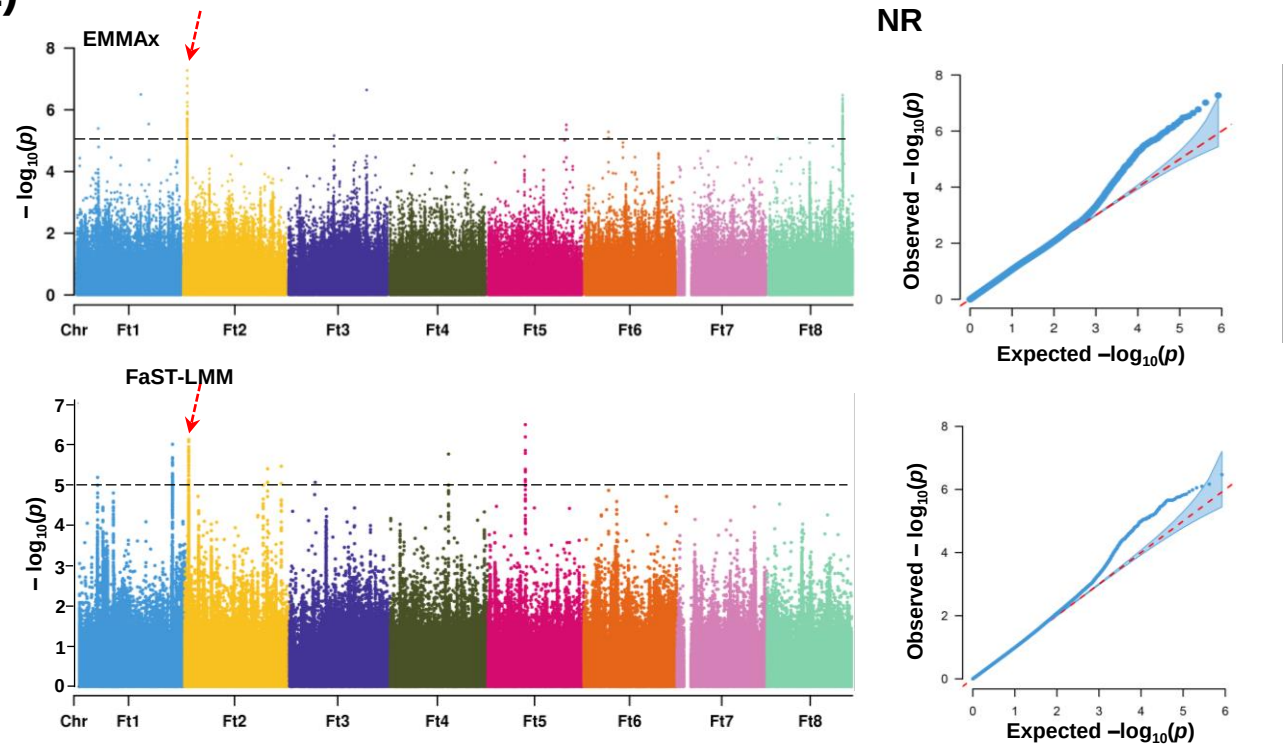


Figure S8 (continued)

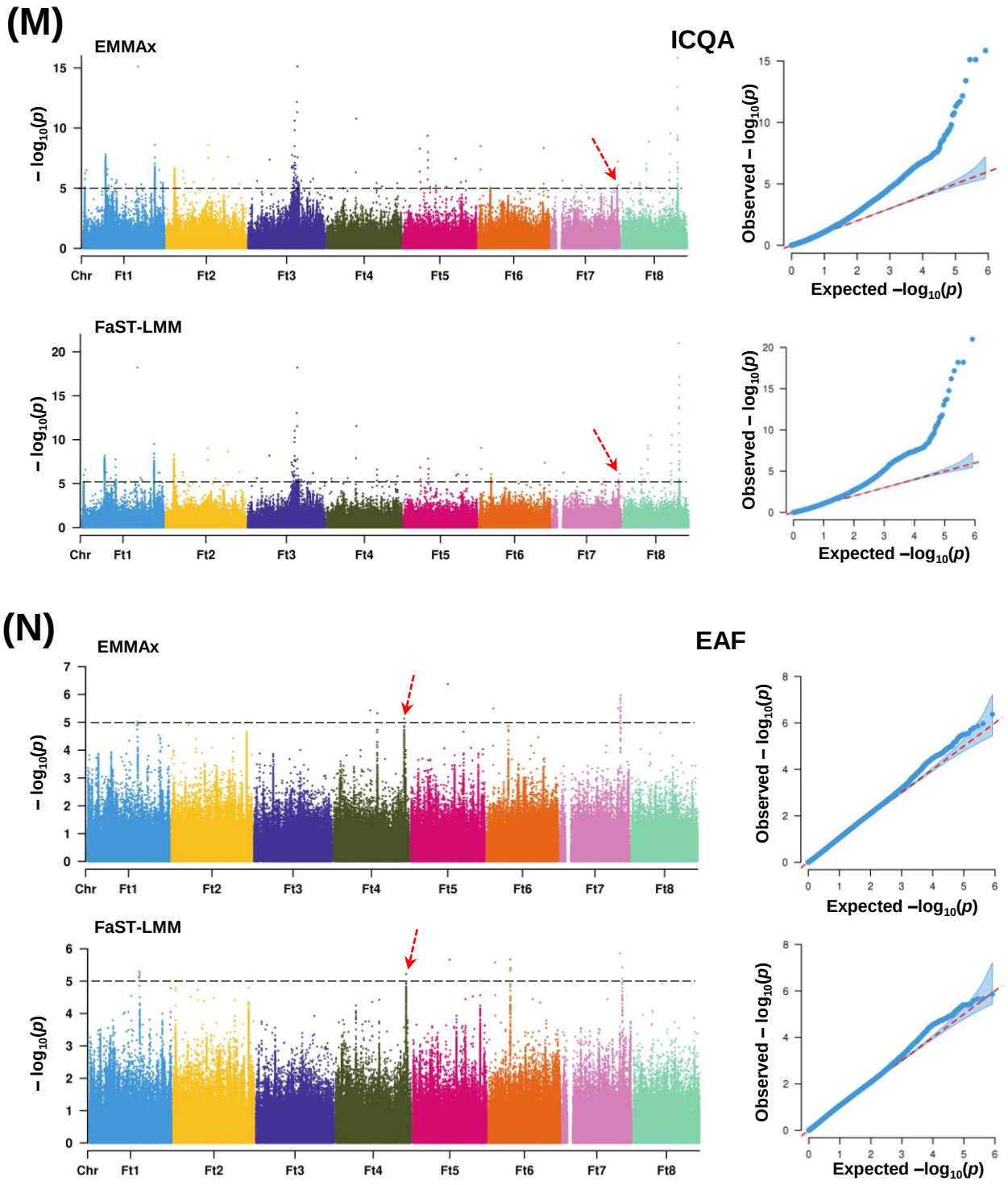


Figure S8. Manhattan plots and Quantile-Quantile (Q-Q) plots of the domesticated metabolites on the basis of different model.

Upper, genome-wide *p*-values and Q-Q plot of the expected null distribution and the observed *p*-value using the EMMAx model. Lower, genome-wide *p*-values and Q-Q plot of the expected null distribution and the observed *p*-value using the FaST-LMM model. The dashed line indicates the threshold $-\log_{10} p = 5$. Arrows indicate mGWAS signals overlapped with selective signals in domestication of SL and NL.

Figure S9

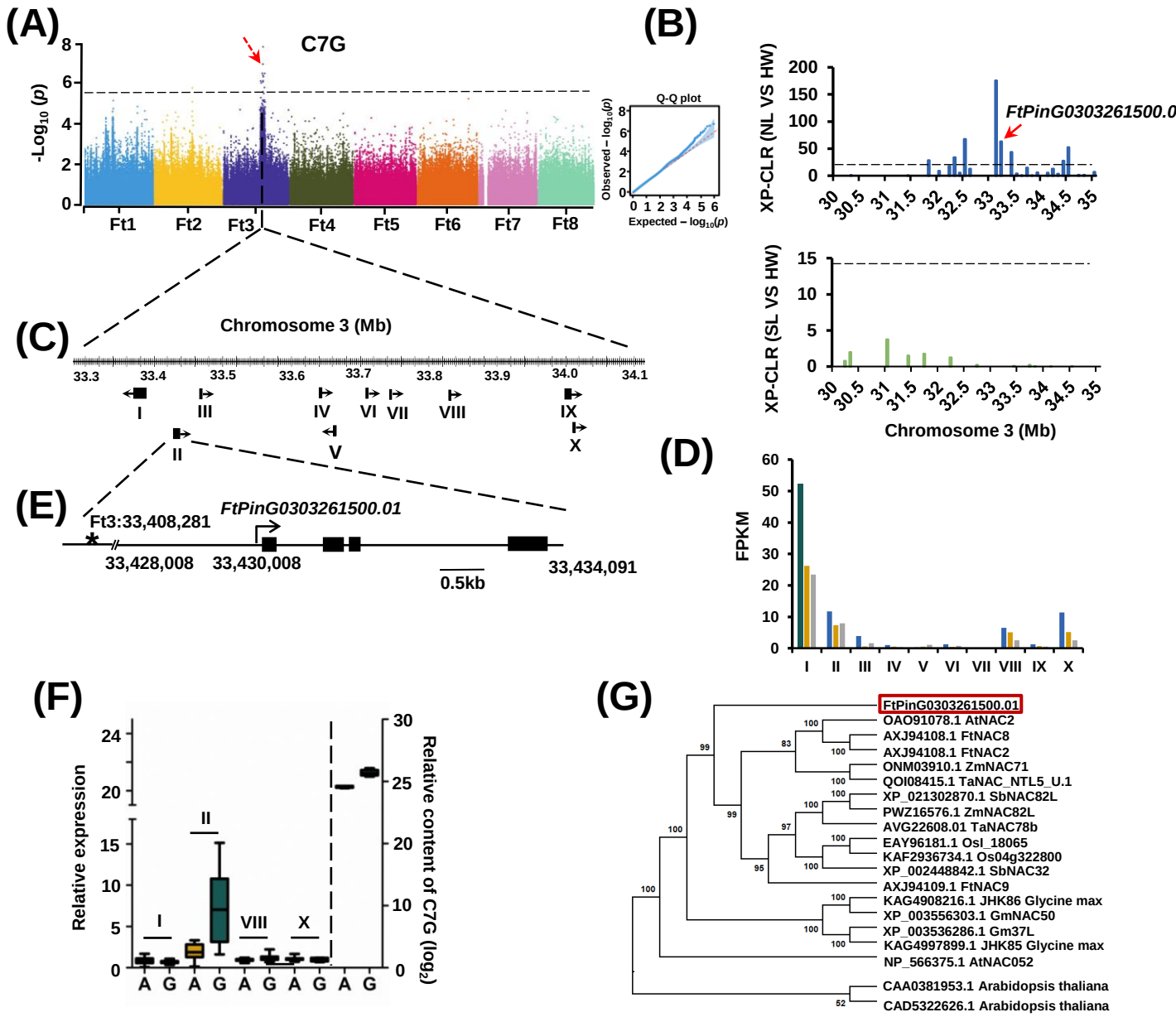


Figure S9. FtNAC is involved in regulating C7G biosynthesis.

(A) Manhattan plot and QQ plot display the GWAS results of C7G content. The dashed line indicates the threshold $-\log_{10}P = 5.68$. (B) The cross population composite likelihood ratio (XP-CLR) between NL and HW (upper), and between SL and HW (lower) on chromosome 3. The black dashed horizontal lines indicate top 10% threshold for entire chromosome 3 (threshod with 18 in NL/HW and 14 in SL/HW). The red arrows indicate the position of *FtPinG0303261500.01* in the sweeps. (C) Schematic representation of genes in the association region. I, *FtPinG0380001007.01*; II, *FtPinG0303261500.01*; III, *FtPinG0303262100.01*; IV, *FtPinG0380001010.01*; V, FtPinG0303264200.01; VI, *FtPinG0303265700.01*; VII, *FtPinG0303266500.01*; VIII, *FtPinG0303268500.01*; IX, *FtPinG0380001016.01*; X, *FtPinG0380001018.01*. (D) Lead SNP adjacent genes expression in different developmental stages of Tartary Buckwheat. (E) Gene model of *FtPinG0303261500.01*. The star (*) represents the lead SNP site. (F) Box plot indicates the relative expression of only *FtPinG0303261500.01* (left) was positively correlated with C7G content (right). *n*Hap.A=12, *nHap.G*=12. (G) Phylogenetic analysis of the *FtPinG0303261500.01*. The neighbor-joining tree was constructed using aligned full-length amino acid sequences. Bootstrap values from 1,000 replicates are indicate at each node. Bar = 0.1 amino acid substitutions per site.

Figure S10

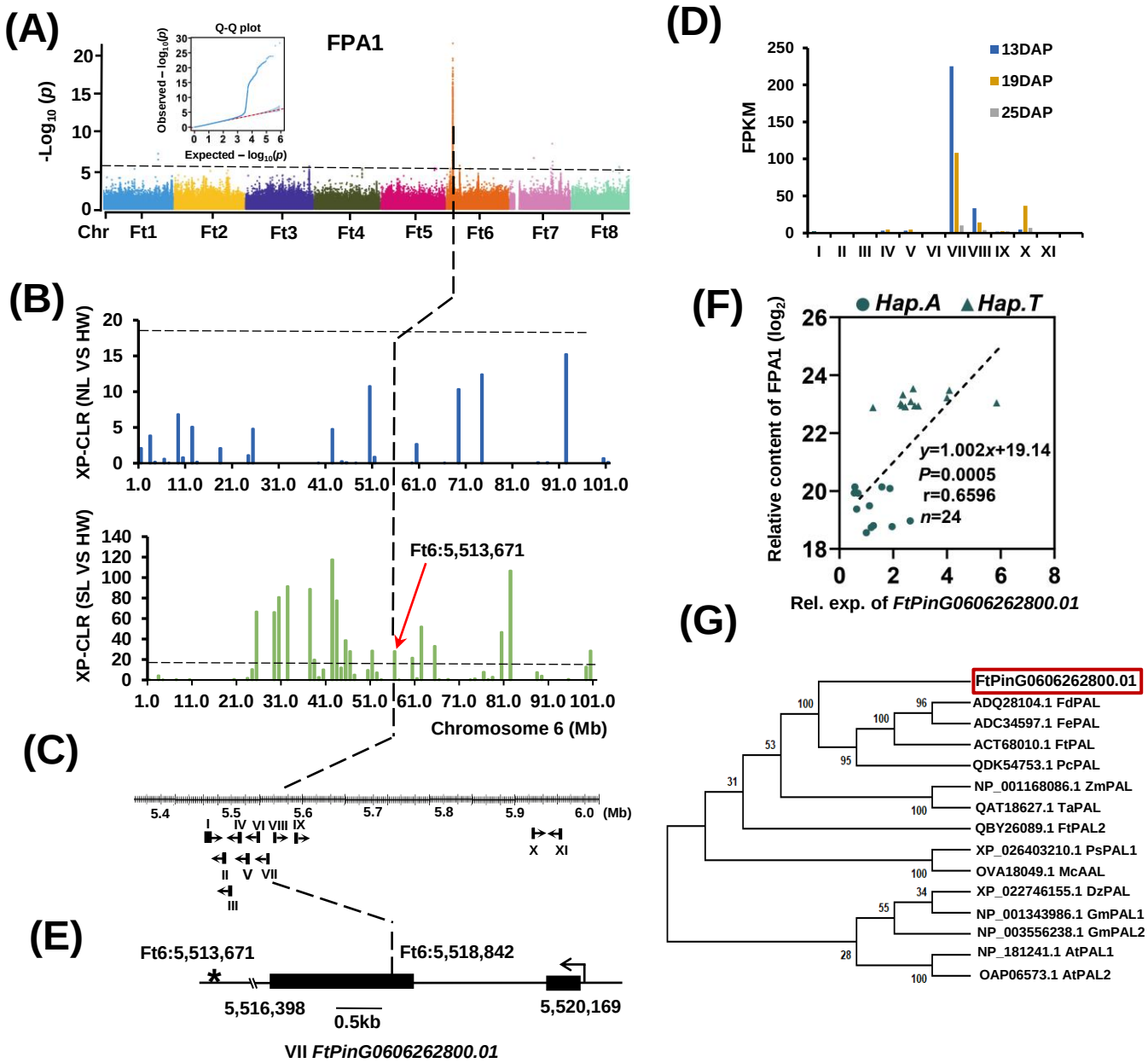


Figure S10. FtPAL plays an important role in FPA1 biosynthesis.

(A) Manhattan plot and QQ plot display the GWAS results of FPA1 content. The dashed line indicates the threshold -$\log_{10} P$= 5.68. (B) The cross population composite likelihood ratio (XP-CLR) between NL and HW (upper), and between SL and HW (lower) on chromosome 6. The black dashed horizontal lines indicate top 10% threshold for entire chromosome 6 (threshod with 18 in NL/HW and 14 in SL/HW). The red arrows indicate the position of *FtPinG0606262800.01* in the sweeps. (C) Schematic representation of genes in the association region. I, *FtPinG0606261300.01*; II, *FtPinG0606261600.01*; III, *FtPinG0606262000.01*; IV, *FtPinG0606262200.01*; V, *FtPinG0606262400.01*; VI, *FtPinG0606262600.01*; VII, *FtPinG0606262800.01*; VIII, *FtPinG0606263900.01*; IX, *FtPinG0606264300.01*; X, *FtPinG0606274800.01*; XI, *FtPinG0606275900.01*. (D) Lead SNP adjacent genes expression in different developmental stages of Tartary Buckwheat. (E) Gene model of *FtPinG0606262800.01*. (F) Correlation between the relative contents of FPA1 and the transcription level of *FtPinG0606262800.01* in 24 buckwheat varieties, including 12 *Hap.A* and 12 *Hap.T*. *P* value was calculated using Student's t-test.(G) Phylogenetic analysis of the *FtPinG0606262800.01*. The neighbor-joining tree was constructed using aligned full-length amino acid sequences. Bootstrap values from 1,000 replicates are indicate at each node. Bar = 0.1 amino acid substitutions per site.

Figure S11

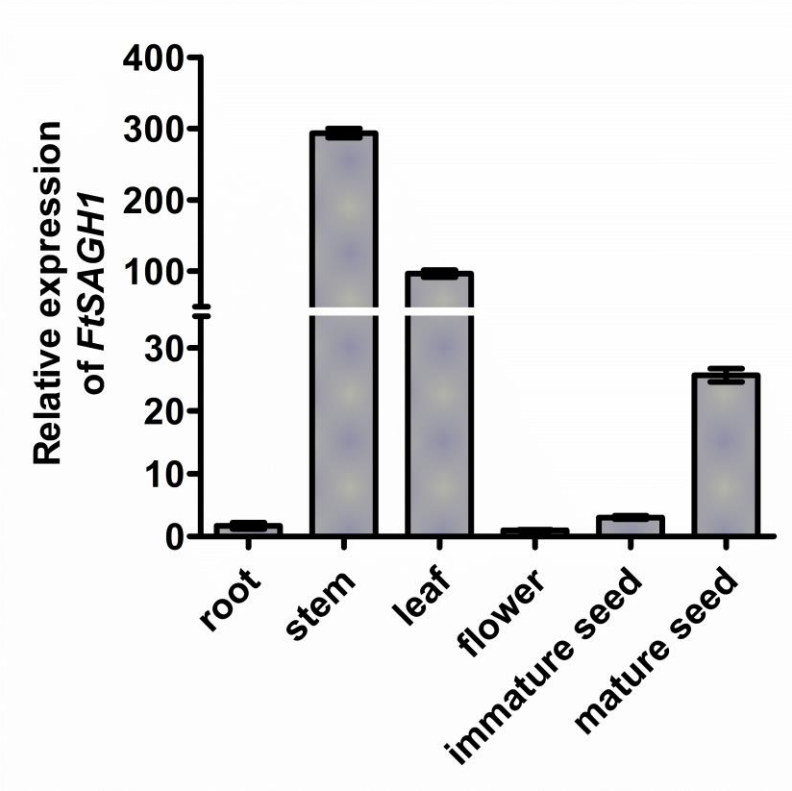


Figure S11. Identification of *FtSAGH1* tissue-specific expression by qRT-PCR

Figure S12

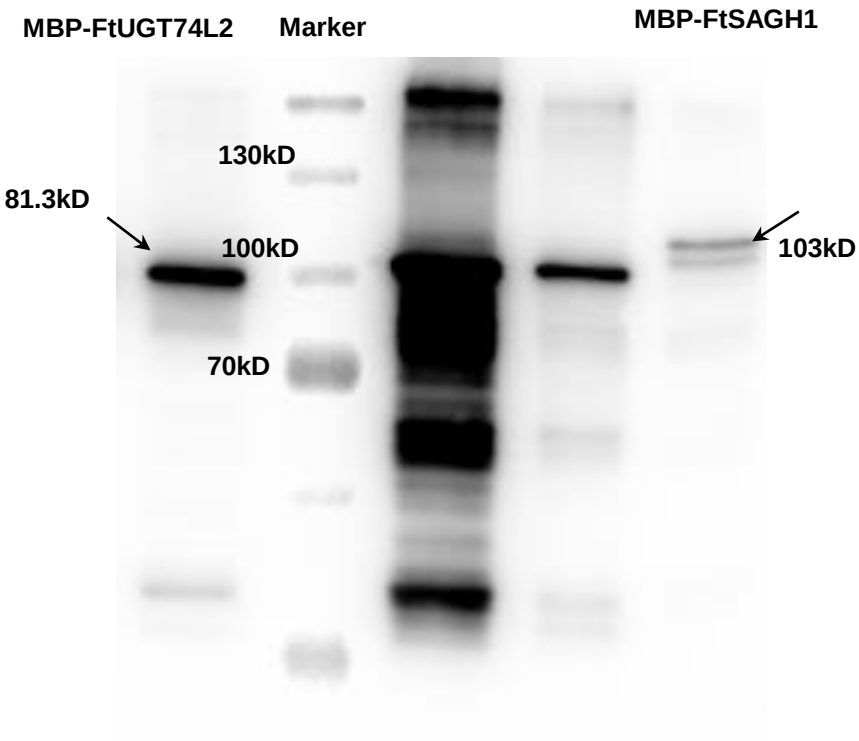


Figure S12. Western blotting with MBP antibody detected the expression of purified recombinant MBP-FtSAGH1 and MBP-FtUGT74L2.

Figure S13

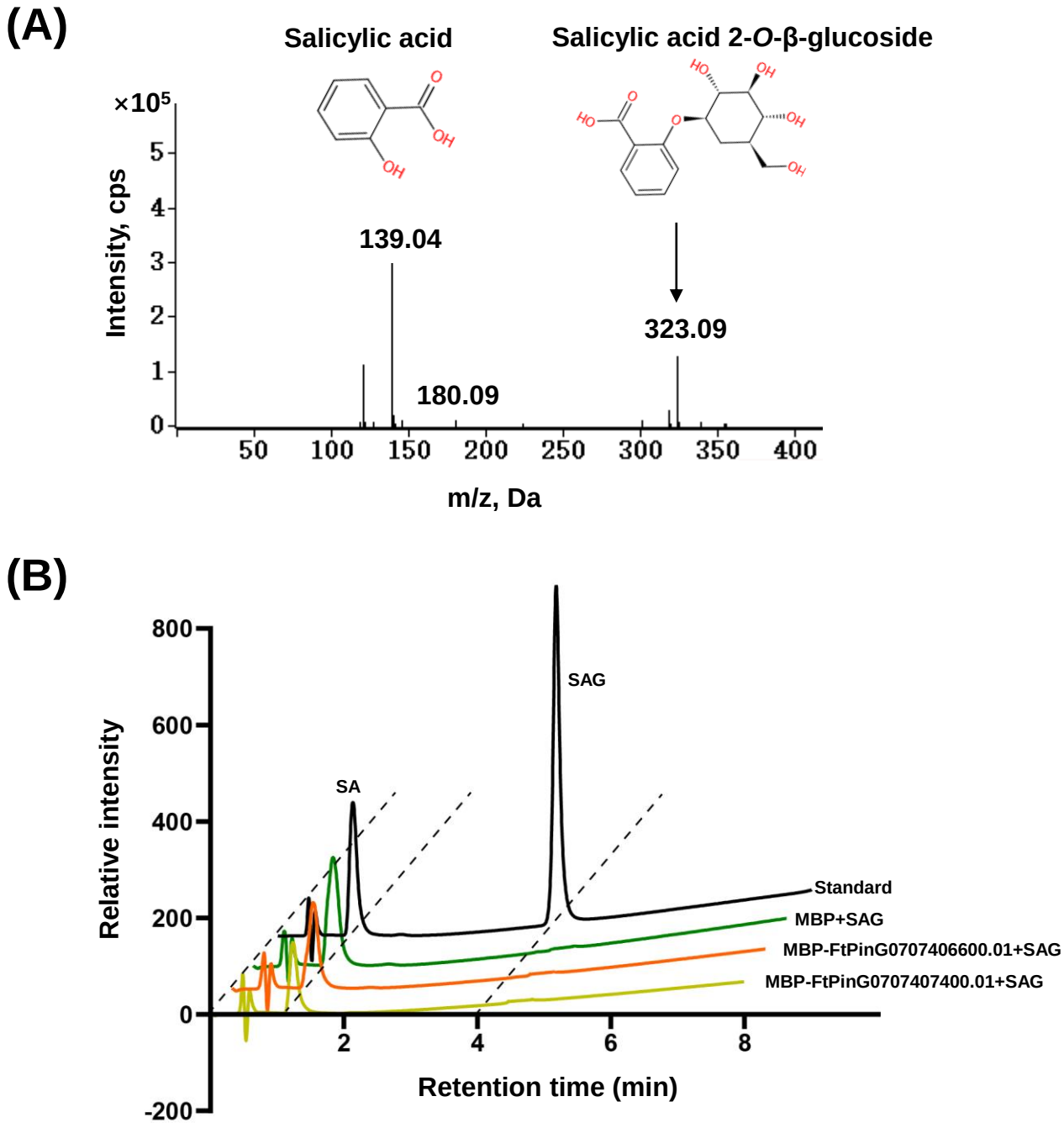


Figure S13. LC-MS/MS fragmentation of salicylic acid and salicylic acid 2-*O*-β-glucoside from enzymatic assays and other two candidate genes enzymatic assay in *vitro*.

(A) Chemical structures and LC-MS/MS fragmentation of salicylic acid and salicylic acid 2-O-β-glucoside from enzymatic assays. (B) FtPinG0707406600.01 and FtPinG0707407400.01 enzymatic assay in *vitro*. The different reaction curves by LC-MS were showed.

Figure S14

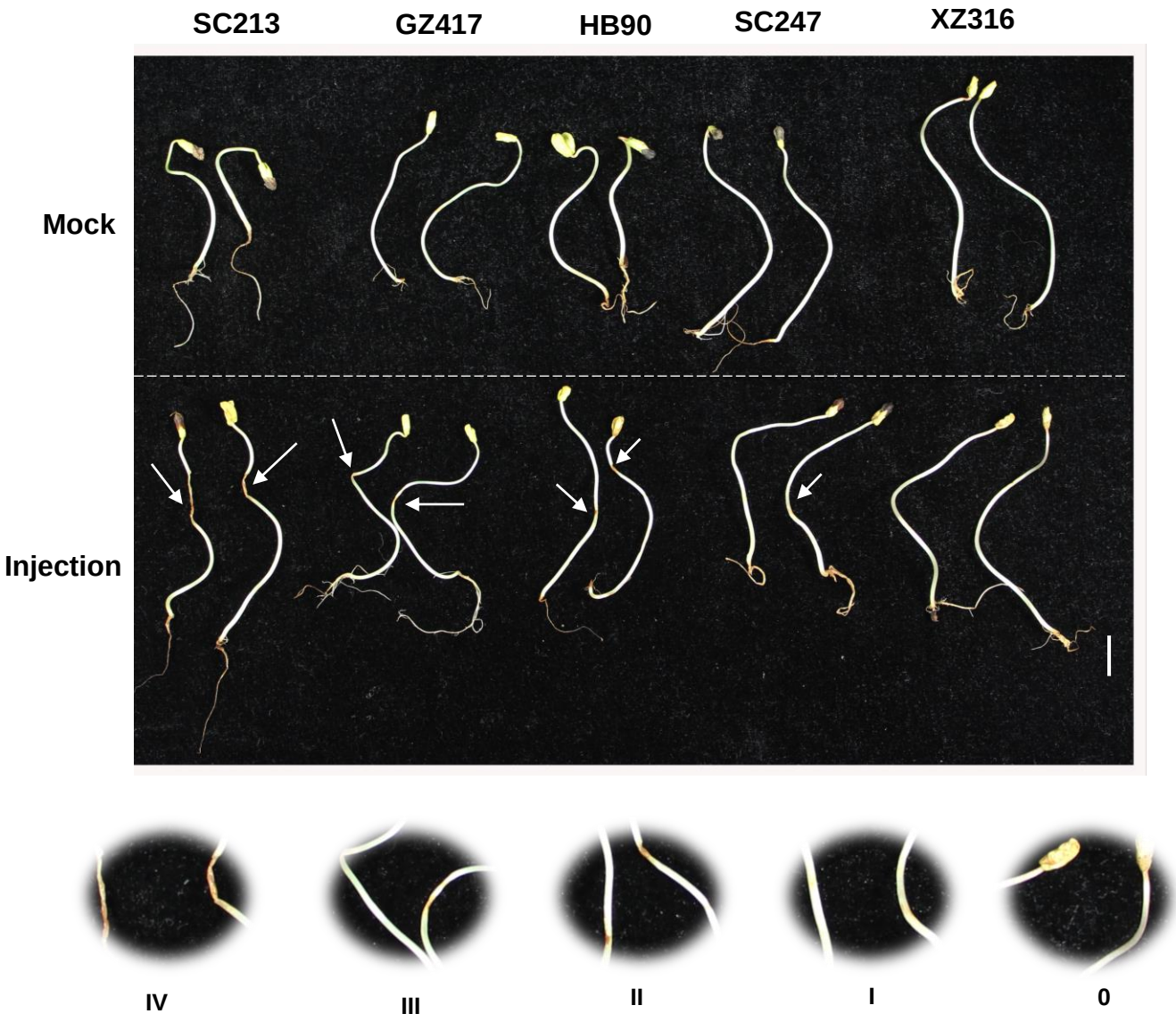


Figure S14. Evaluation of disease resistance to *Rhizoctonia solani* in different buckwheat accessions.

Standard for resistance grades of buckwheat were given according to the damping-off. 0 grade, symptomless; I grade, a slight spot or discoloration; II grade, patches or discoloration or decay on the base of stem; III grade, patches around the whole stem or root, discolored and rotten; IV grade, buckwheat wilts, falls or dies. Bar = 1 cm.

Figure S15

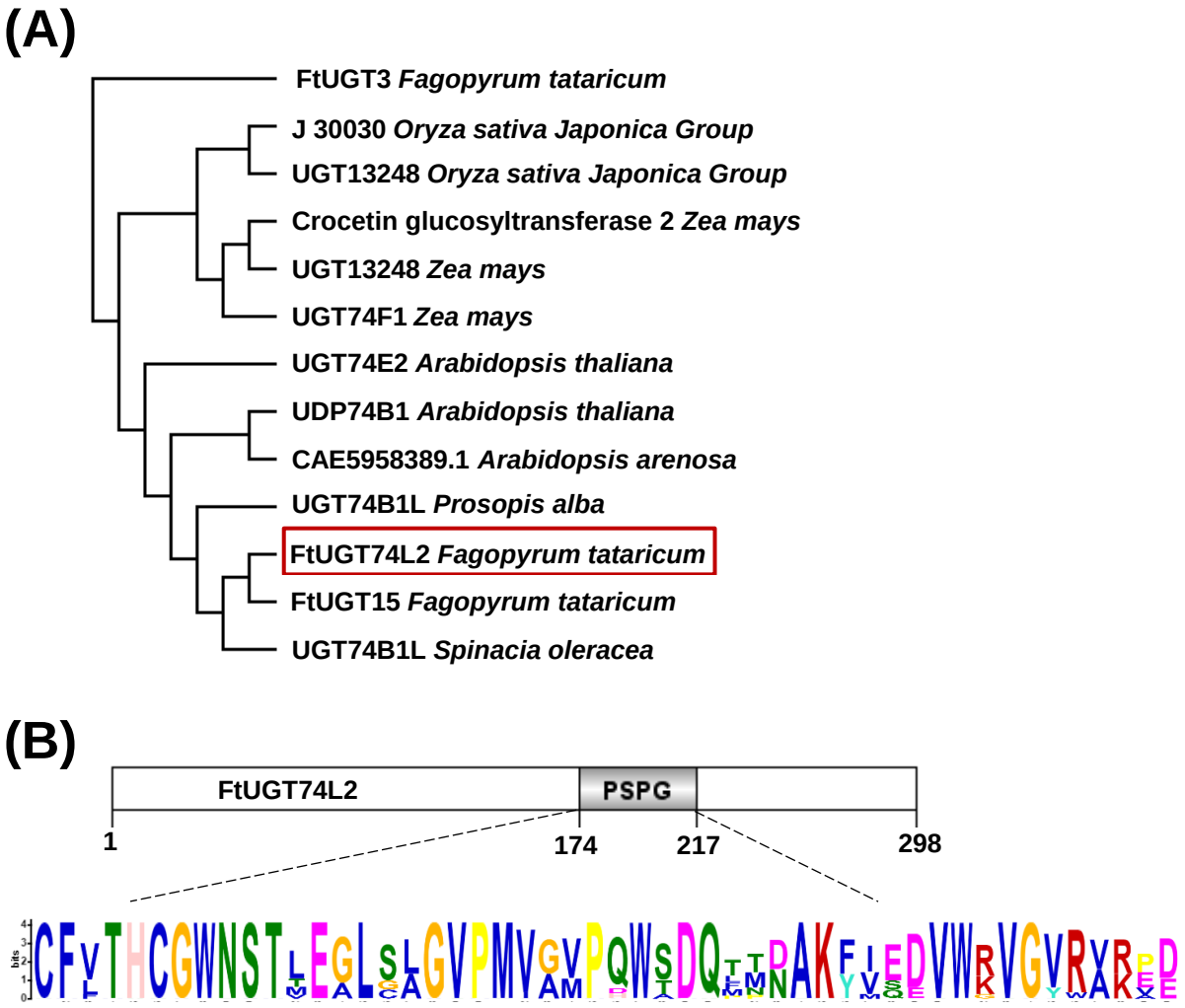


Figure S15. The analysis of the proposed function of FtUGT74L2.

(A) Phylogenetic analysis of the FtUGT74L2. The neighbor-joining tree was constructed using aligned full-length amino acid sequences. Bootstrap values from 1,000 replicates are indicate at each node. Bar = 0.1 amino acid substitutions per site. (B) The conserved sequence of the plant secondary product glycosyltransferase (PSPG) box in UGTs identified so far.

Figure S16

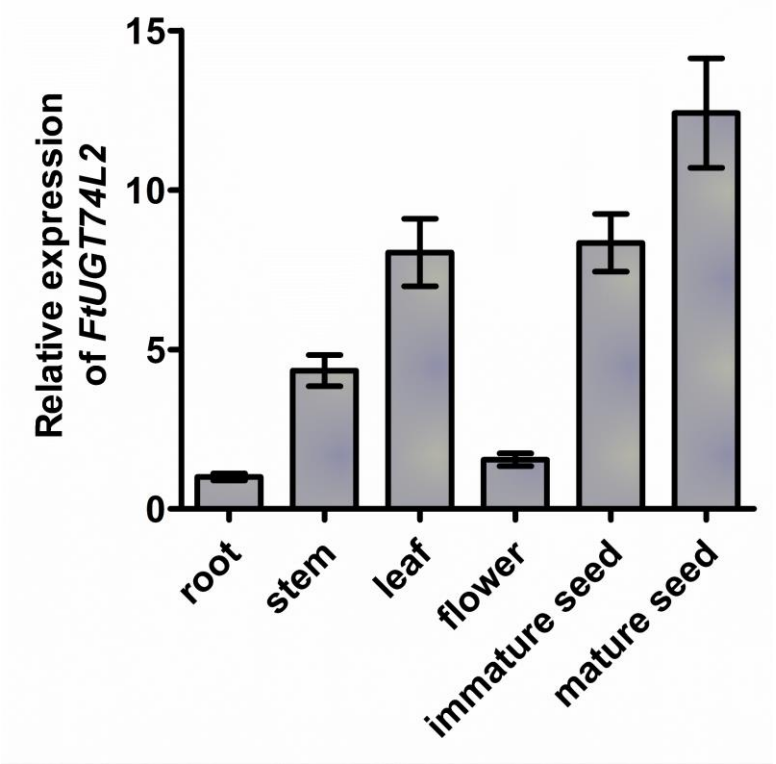


Figure S16. Identification of *FtUGT74L2* tissue-specific expression by qRT-PCR.

Figure S17

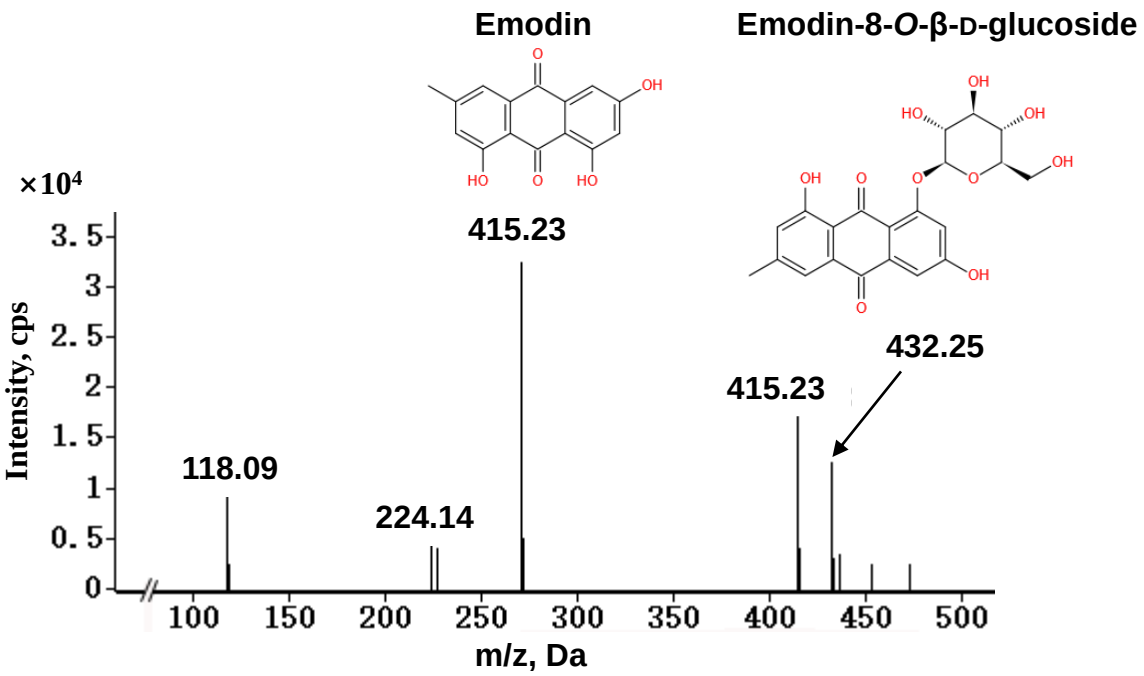


Figure S17. Chemical structures and LC-MS/MS fragmentation of emodin and emodin-8-*O*-β-D-glucoside from enzymatic assays.

Figure S18

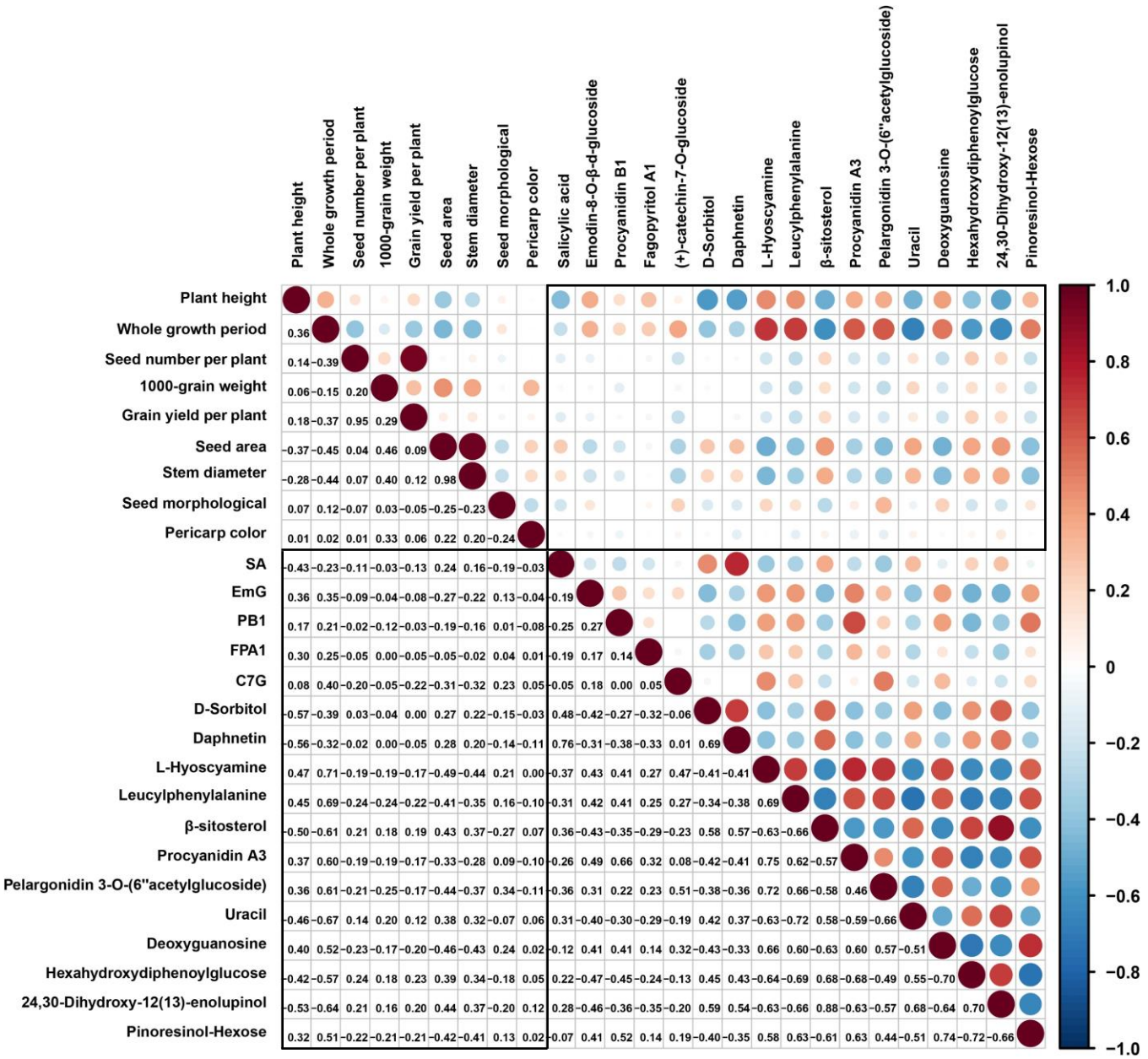


Figure S18. Correlation analyses of agronomic traits and partial metabolites.

The correlation between agronomic traits and metabolites of 200 Tartary buckwheat accessions on the basis of Spearman's correlation coefficients were presented. The red-to-blue colors for each cell represent the two corresponding metabolites sharing high-to-low similarly distributed patterns (indicated by the coefficient values) among buckwheat accessions.

Figure S19

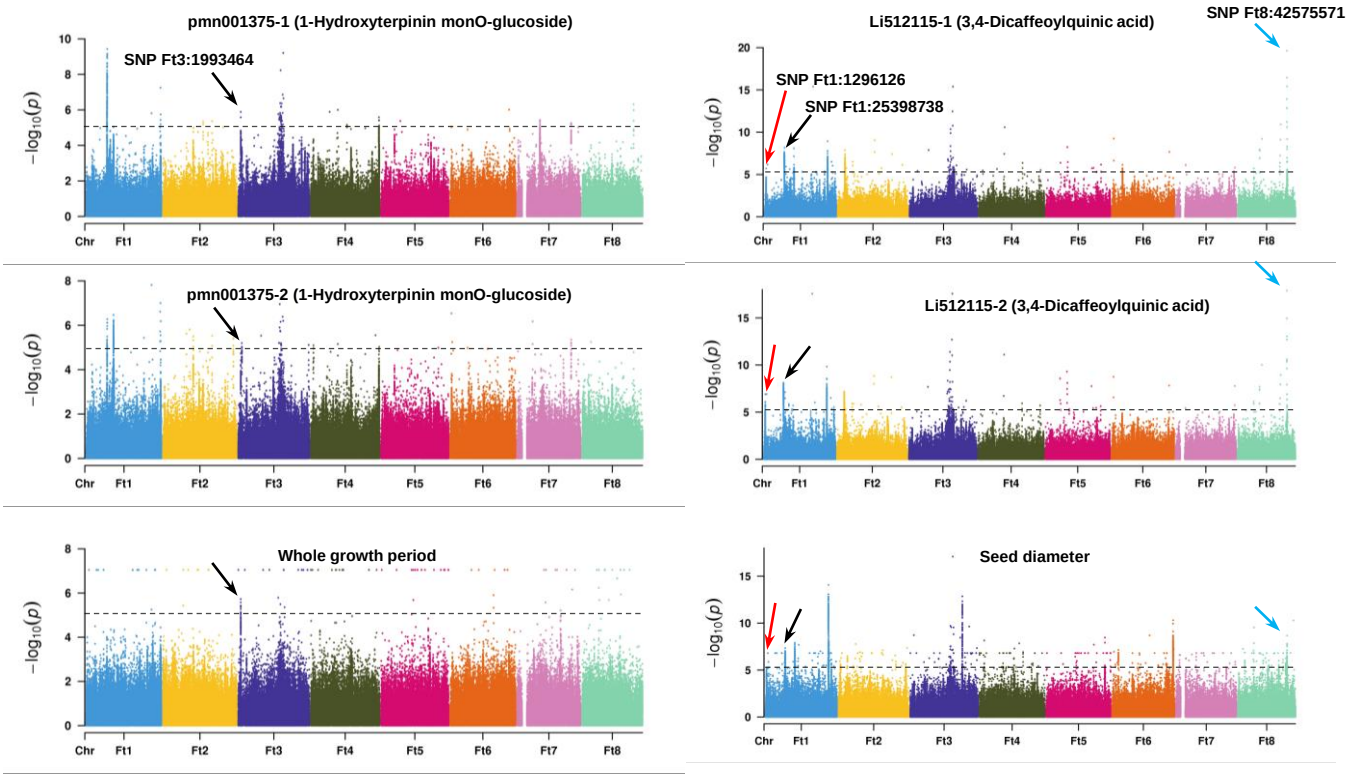


Figure S19. Manhattan plots display the mGWAS of 1-Hydroxyterpinin monO-glucoside and pGWAS of whole growth period (left), and mGWAS of 3,4-Dicaffeoylquinic acid and pGWAS of seed diameter (right). The dashed line indicates the threshold -$\log_{10}P$=5. The same color arrows indicate the same loci.

Figure S20

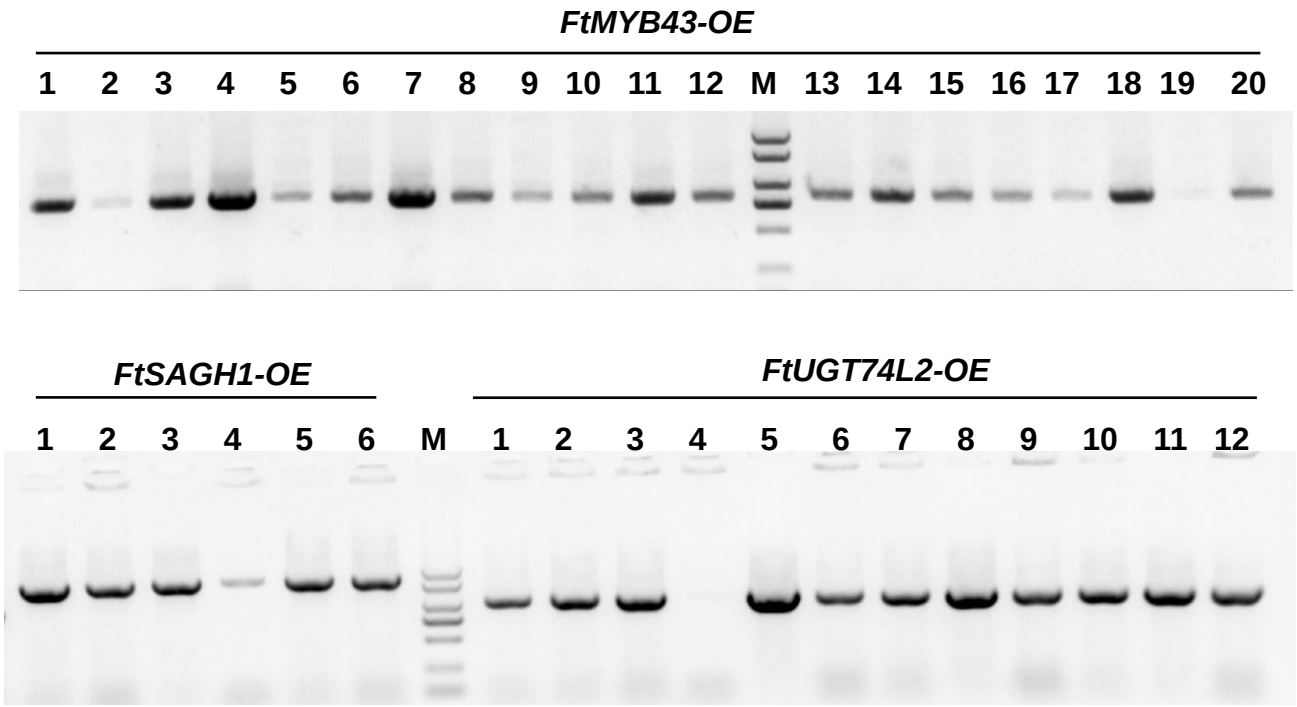


Figure S20. All related transgenic hairy roots were verified by PCR.

The forward primers were paired with the complementary sequences of plasmid (pCAMBIA 1307 for *FtMYB43-OE*, pCAMBIA 1302 for *FtSAGH1-OE* and *FtUGT74L2-OE*), and the reverse primers were paired with the complementary sequences of template respectively.

Figure S21

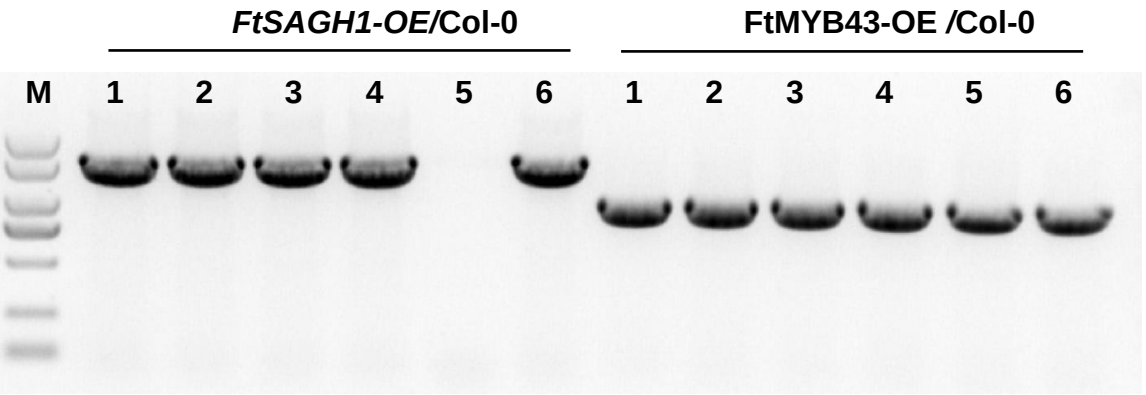


Figure S21. Ectopic expression of *FtSAGH1* and *FtMYB43* in *Arabidopsis* were verified by PCR.

Figure S22

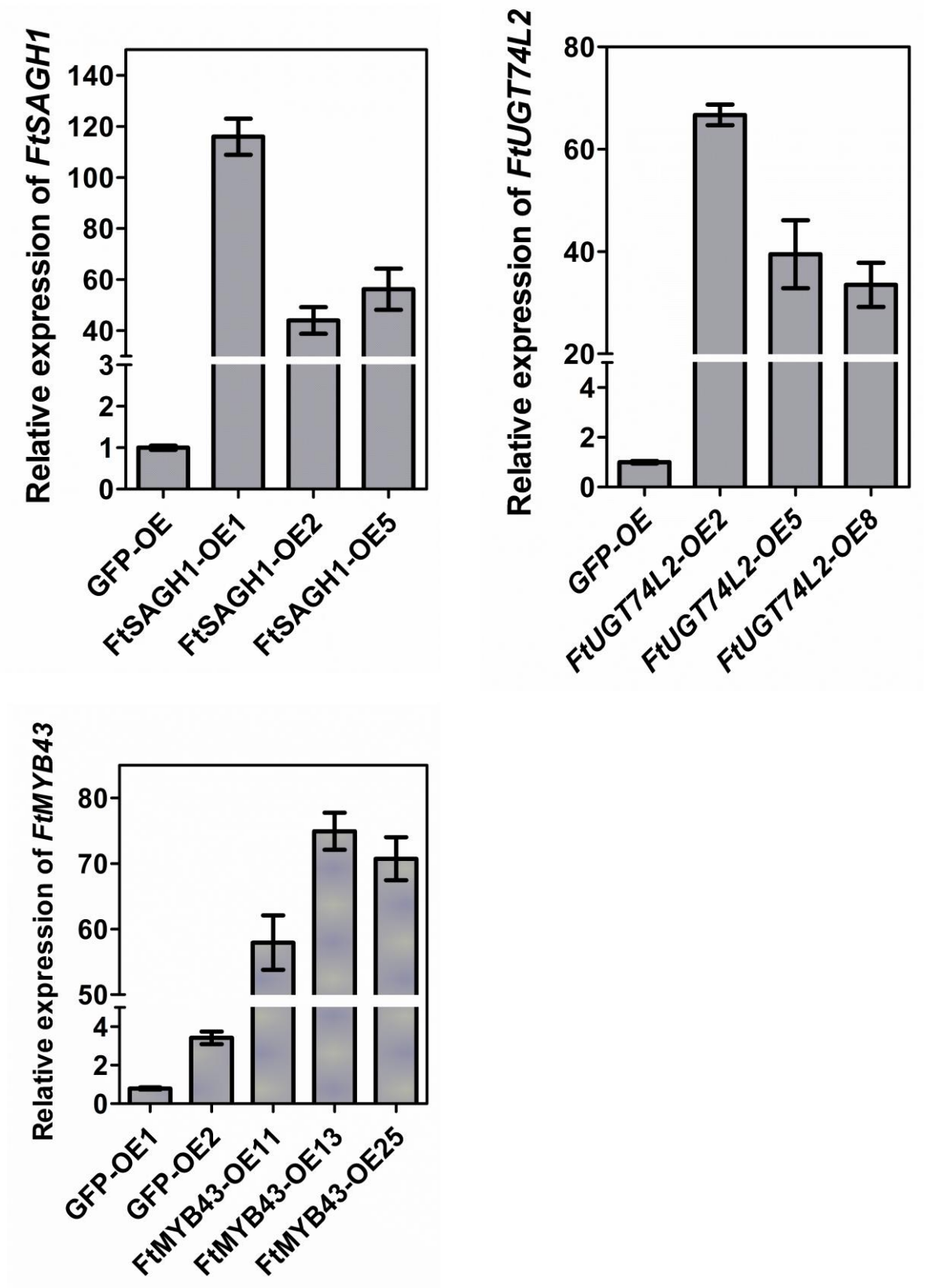


Figure S22. All related transgenic hairy roots were verified by qRT-PCR.