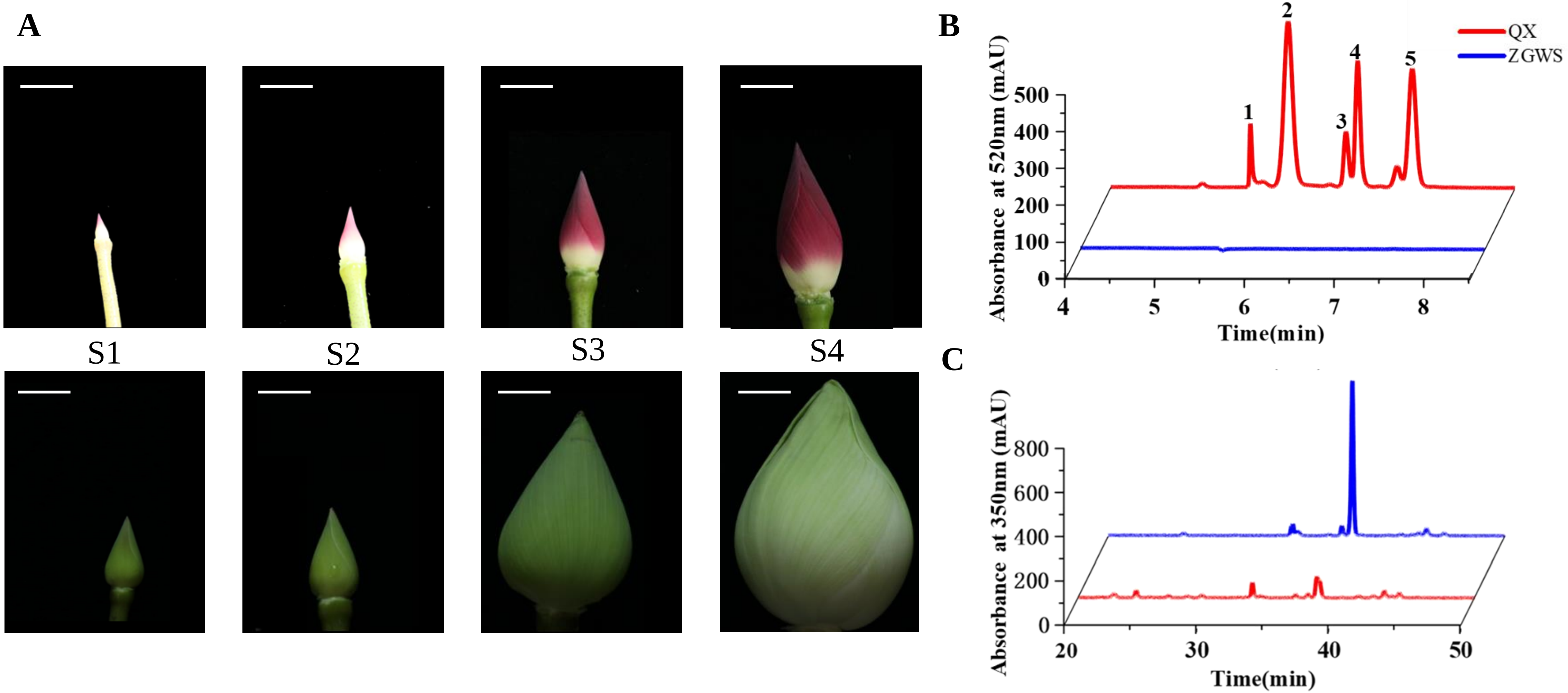
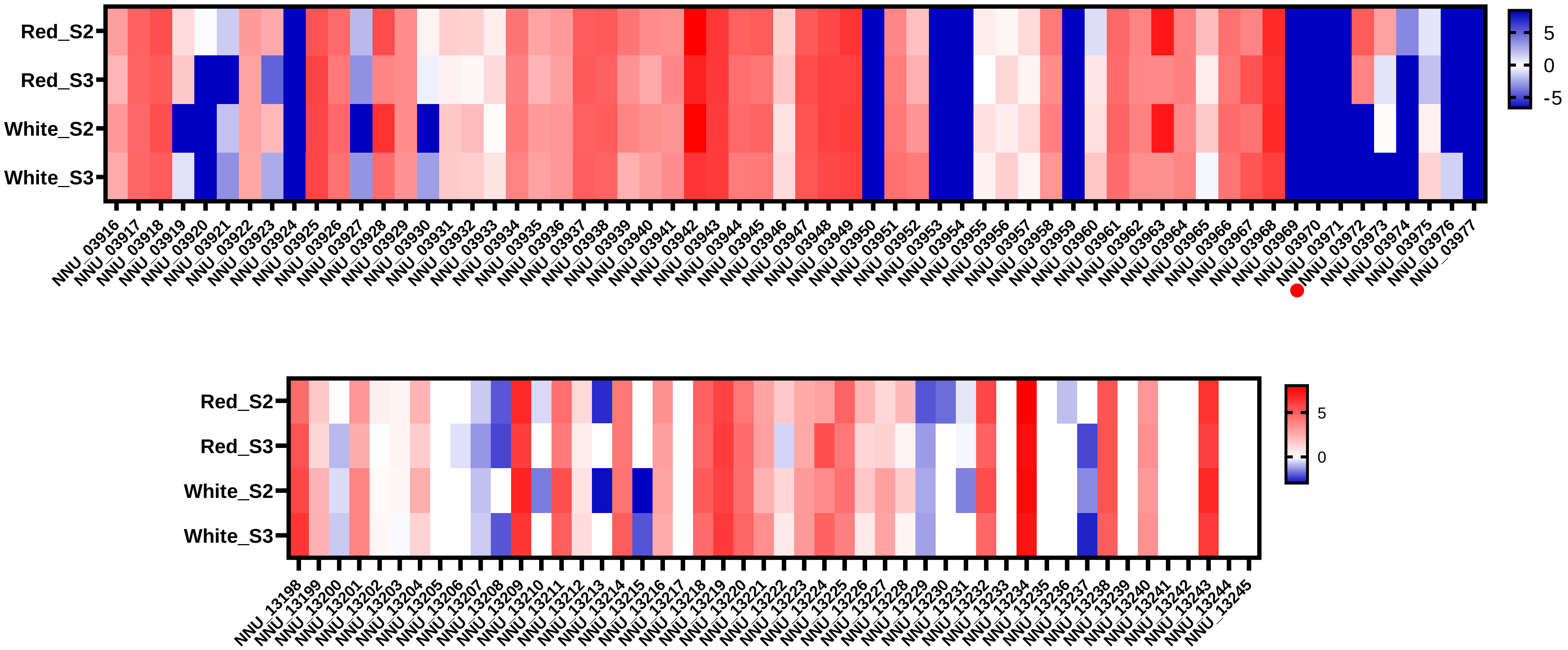
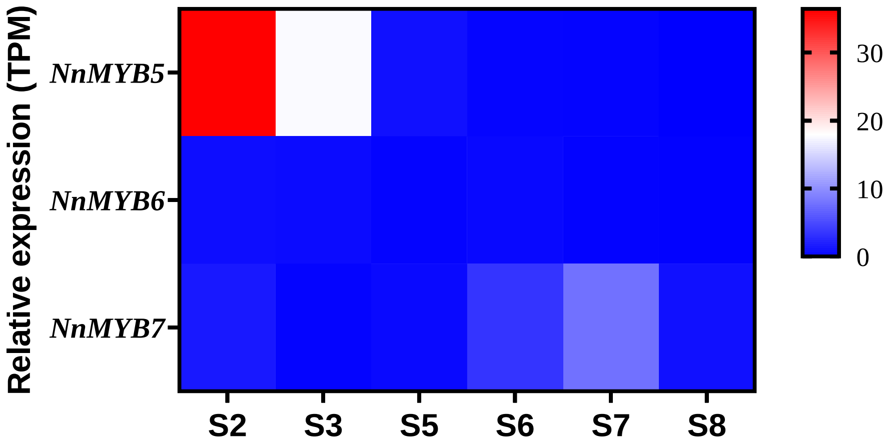


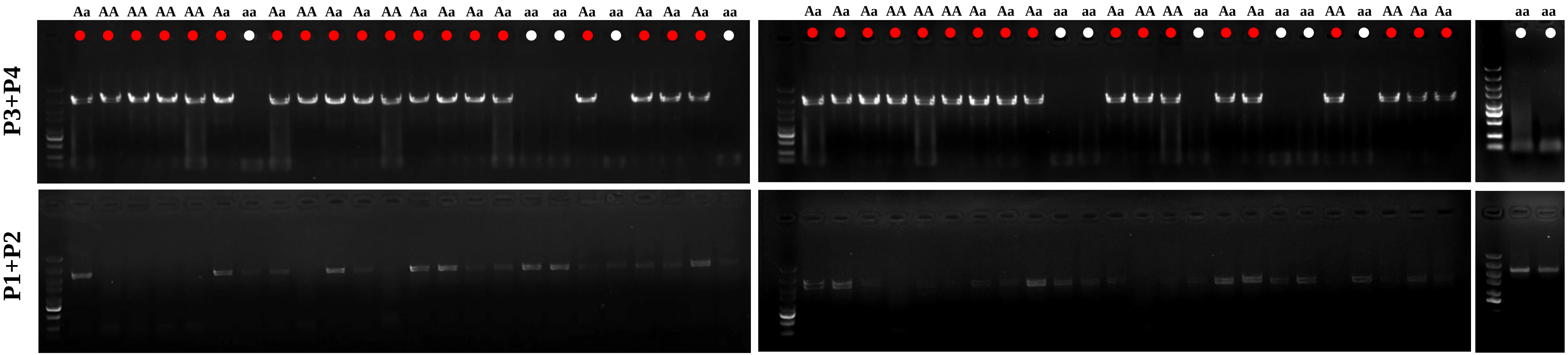
Supplemental Figure S1. Flower development and anthocyanin accumulation in the red and white lotus flowers. A. Four flower developmental stages (S1-S4) in the *N. nucifera* cv. ‘QX’ (top) and ‘ZGWS’ (bottom), scale bar = 500 mm. B. High performance liquid chromatography (HPLC) chromatogram of anthocyanin profiles in the red and white petals. Peaks 1–5 were detected in the red petals at 520 nm, which corresponded to delphinidin 3-*O*-glucoside, cyanidin 3-*O*-glucoside, petunidin 3-*O*-glucoside, peonidin 3-*O*-glucoside, and malvidin 3-*O*-glucoside, respectively. C. HPLC chromatogram of flavonoids in the red and white petals detected at 350 nm.



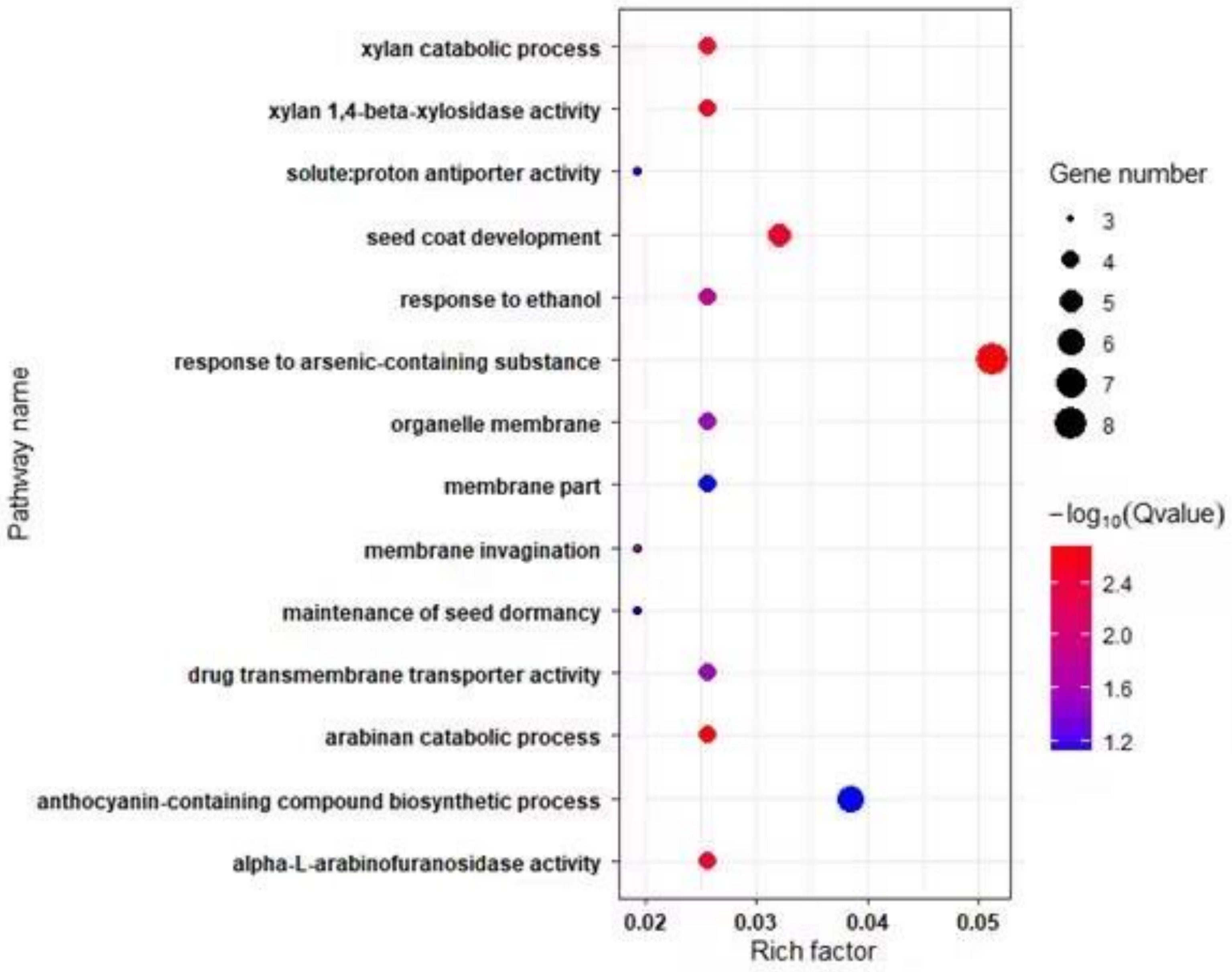
Supplemental Figure S2. The expression patterns of 110 genes in the two target intervals between the red and white petals at developmental stages 2 (S2) and stage 3 (S3). The heatmaps illustrate the digital expression level log2 (Transcripts per million) in four samples. The red dot in the upper panel indicates the red flower specific expressed *NnMYB5* (*NNU_03972*) gene.



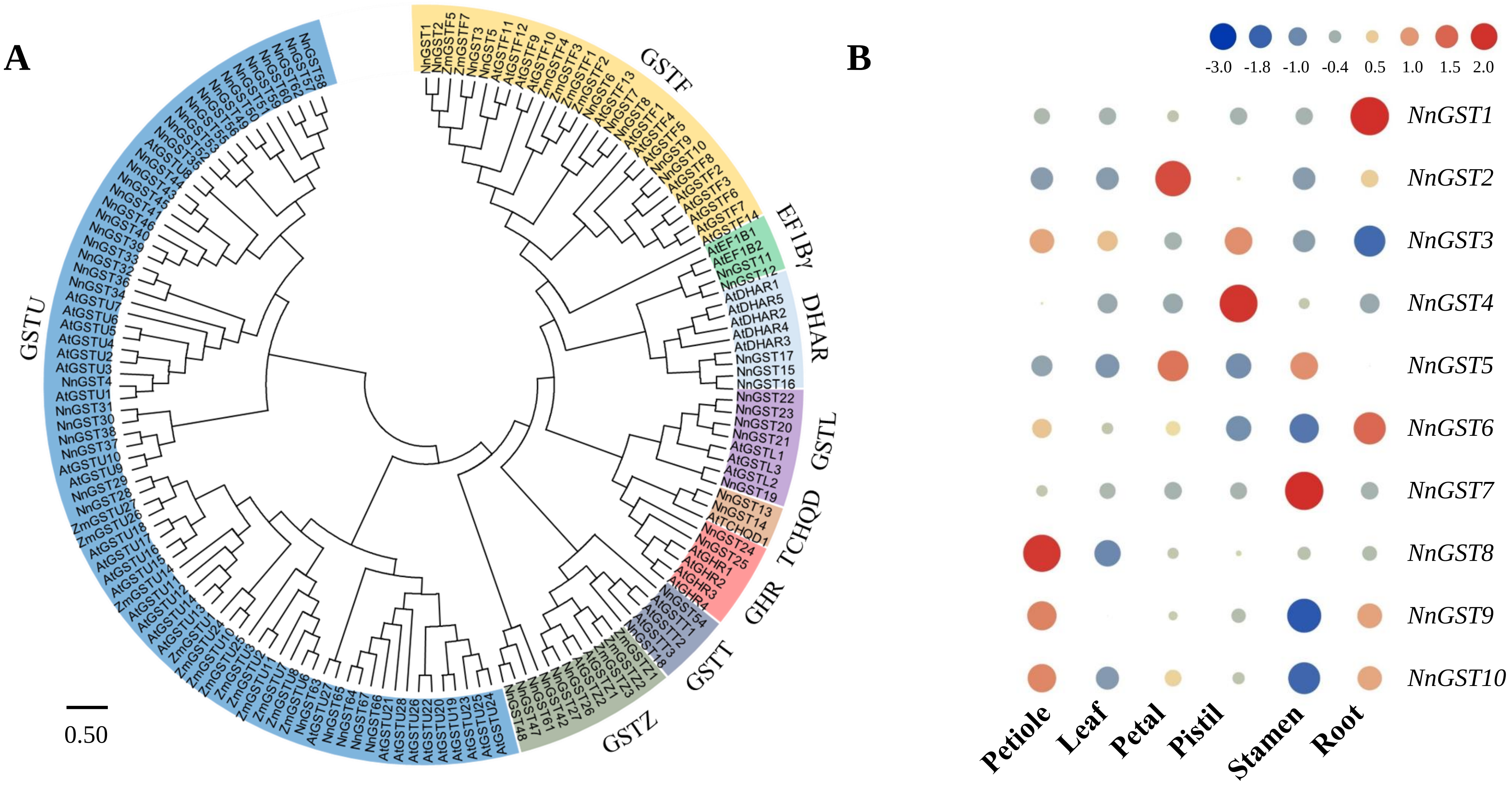
Supplemental Figure S3. Digital expression levels of three genes in the 80-kb PAV in red petals at different developmental stages. Transcripts per million (TPM) was used to measure the gene expression level. S2 and S3 are petals at developmental stage 2 and 3 as described in this article, S5 to S8 are petals at developmental stage 5 to stage 8 from flowering flowers reported previously(Liu et al., 2022).



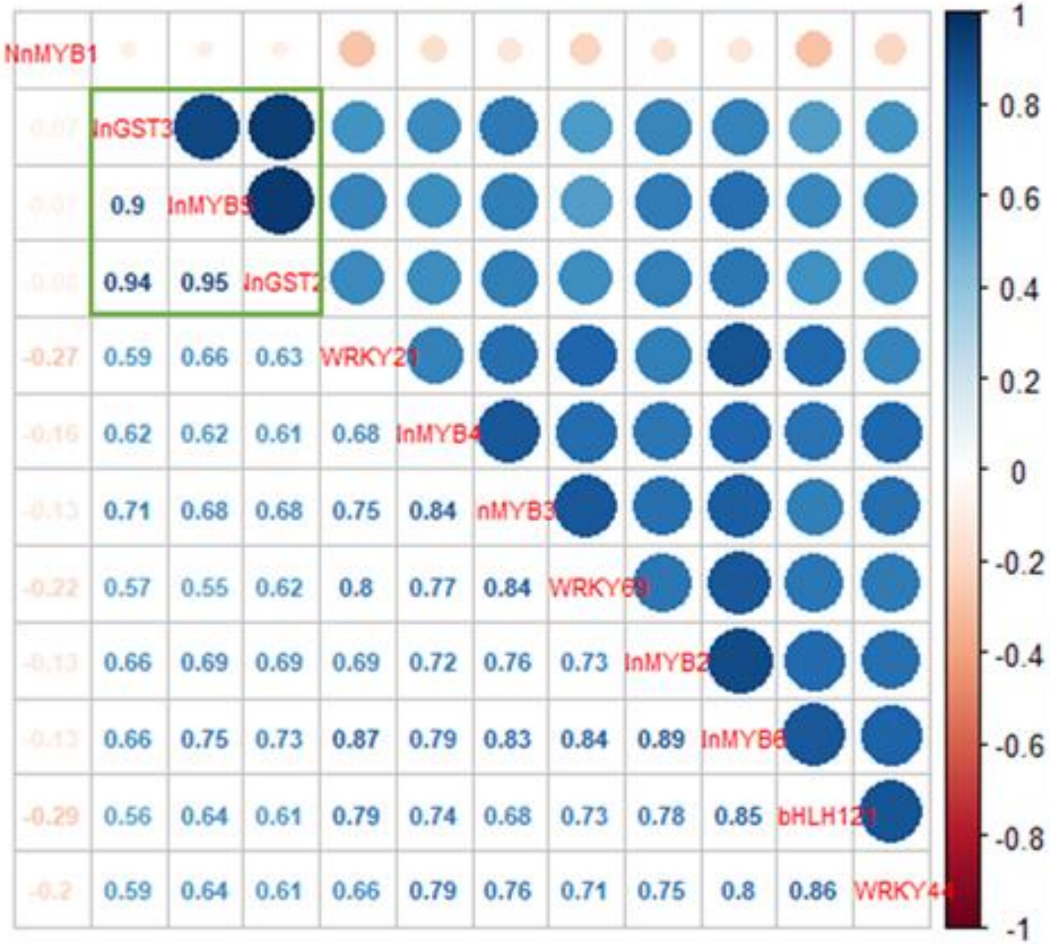
Supplemental Figure S4. PCR genotype verification of the F_2 individuals used in this study. Primers P1/P2 and P3/P4 were used to amplify the gDNA of lotus F_2 individuals, and 13, 24, and 13 individuals with AA, Aa, and aa genotypes, respectively, were identified. Red and white dots represent lotus individuals with the red and white flowers.



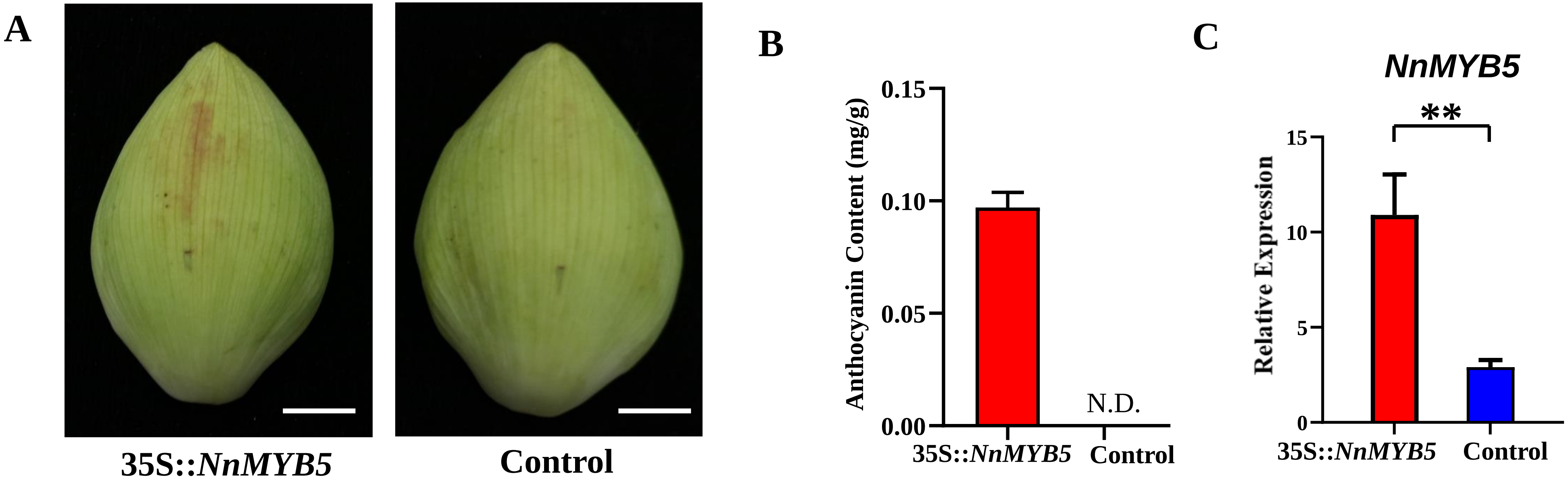
Supplemental Figure S5. GO enrichment of the 170 DEGs.



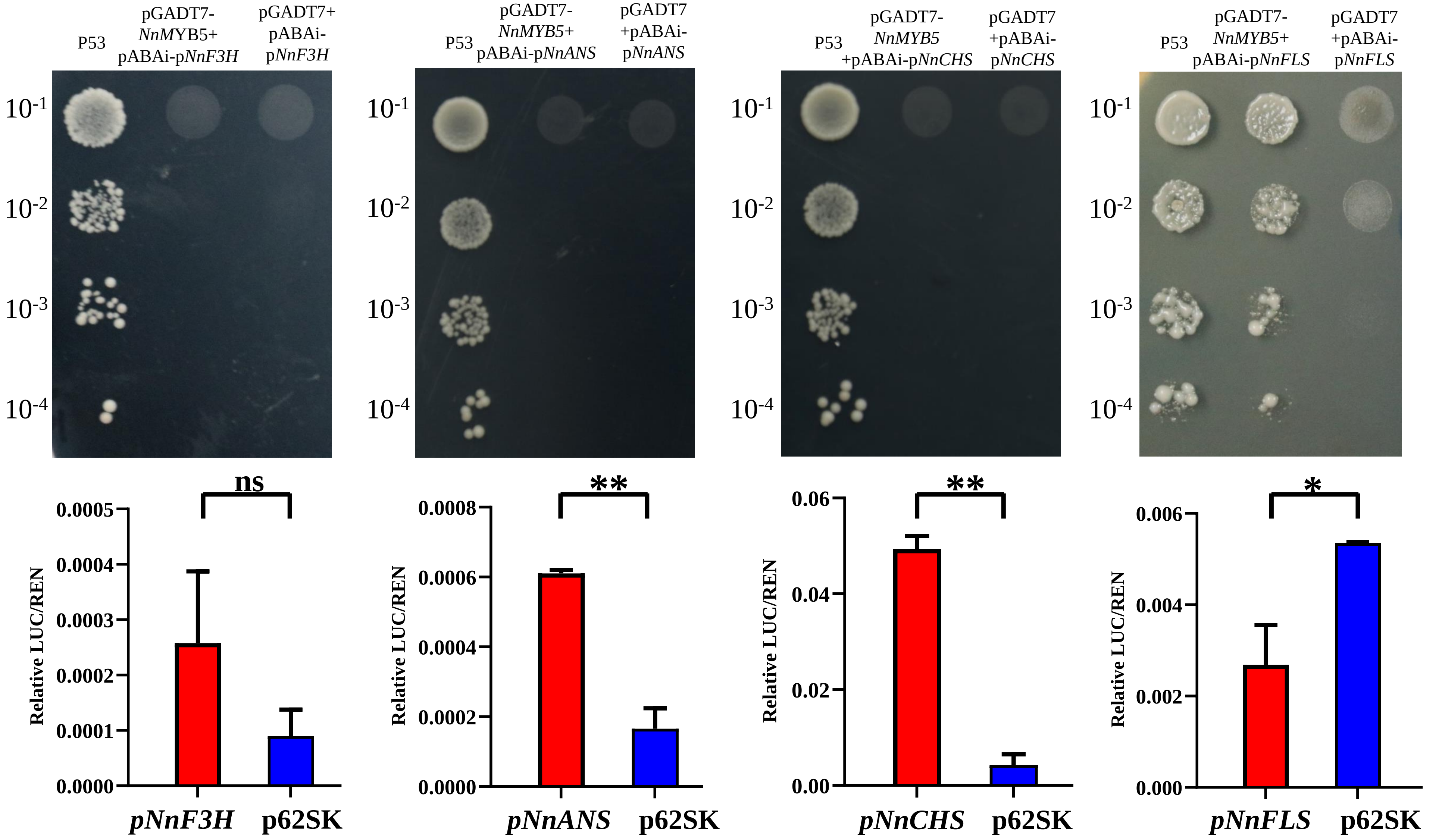
Supplemental Figure S6. Phylogenetic analysis and expression patterns of GSTs in *N. nucifera*. A. Phylogenetic tree of GSTs from *A. thaliana*, *Z. mays*, and *N. nucifera*. The 62 *Arabidopsis*, 20 maize, and 67 lotus GST protein sequences were aligned using ClustalW. The phylogenetic tree was constructed based on the neighbor-joining algorithm, with 1000 bootstrap replicates. The accession numbers of the nine GST classes are listed in Table S5. B. Expression patterns of the ten phi class *GSTs* in *N. nucifera*. The $\log_2$(FPKM) expression levels of *GSTs* were analyzed using our unpublished RNA-seq data (Heyun Song, Heng Sun and Mei Yang, unpublished data) of petiole, leaf, petal, stamen, and root. Color and size of the dots indicate the value of $\log_2$(FPKM): positive values are in red and negative values are in blue. And the larger the absolute values are, the larger the dots size are.



Supplemental Figure S7. Pearson correlation analysis between candidate TFs and *NnGST2*. Color towards blue and red represent positive and negative correlations, respectively, while dot size represents the correlation value. The number in each cell show the correlation value as well. The green box showed a high correlation between *NnMYB5* and *NnGST2* and *NnGST3*.



Supplemental Figure S8. Transient expression of *NnMYB5* induce anthocyanin accumulation in the white lotus petals. A. The phenotype of lotus petal after transient *NnMYB5* gene expression. Scale bar = 1cm. B. Anthocyanins content in the transgenic 35S::*NnGST2* lines and control petals. N.D, Not detected. C. Transcript levels of *NnMYB5* in the transgenic 35S::*NnMYB5* lines and control petals. Data are mean ± SD ($n = 3$). *t*-test, ** $P < 0.01$ in B and C.



Supplemental Figure S9. The yeast one-hybrid (Y1H) and dual-luciferase reporter assay (LUC) validation of NnMYB5 regulatory activity on four anthocyanin synthesis genes. NnMYB5 could activate the promoter of *NnFLS*, but with no effects on the promoters of *NnF3H*, *NnANS*, and *NnCHS*. From top to bottom, the yeast concentrations at OD600 = 0.1, 0.01, 0.001, and 0.0001, respectively. P53 and pGADT7+pABAi-p*NnF3H/ANS/CHS/FLS* were used as positive and negative control in Y1H. p62SK was negative control in LUC. Data are mean ± SD (n = 3). t-test, **$P < 0.01$, *$P < 0.05$.

Supplemental Table S1. Primers used in this study.

Primer	**Sequence (3’-5’)**
NnGST2-QPCR-F	GCAGTTTGCCCTCAGAGAGT
NnGST2-QPCR-R	GCCCTTGACTCAAAGAGCCT
NnMYB5-QPCR-F	CAAAACGCGGTTGAGCAAGA
NnMYB5-QPCR-R	GAGATCGTCTGTAGGCGGTG
NnCHS-QPCR-F	CTGTGTCGACCAGAGCACAT
NnCHS-QPCR-R	CTCCTCGGTCAAGTGCATGT
NnF3H-QPCR-F	CATTCGCCGAGATGTACCGA
NnF3H-QPCR-R	TTTTGCTTTCCGCTTCCAGC
NnFLS-QPCR-F	ATCGACACCATCCCTCCTGA
NnFLS-QPCR-R	TTGACGACCTGGAAGATGCC
NnANS-QPCR-F	GGCTCAACCACTTGCAGTTG
NnANS-QPCR-R	AGCTCTTCTTGAGGACGCAC
AtANS-QPCR-F	TGGAAGCTCACACCGATGTAAG
AtANS-QPCR-R	CGGAGACTCAACACTCACCATC
AtCHI-QPCR-F	TCACGAAGCTTCATGTAGACTCC
AtCHI-QPCR-R	CCTTGGATATCAAGGCCTCGGA
AtF3H-QPCR-F	CAAGACATGGATTACGGTTCAGC
AtF3H-QPCR-R	ACGTGGCTATGGATAATCTGCTC
AtF3'H-QPCR-F	GATCTTGACGGTGACGGAGGAA
AtF3'H-QPCR-R	GCGATGTGTGGTAACGAGAGTG
AtFLS-QPCR-F	ATGCGTCAATTACAGATTCTGGCCT
AtFLS-QPCR-R	TTCTTTCAACGCATCACGCTTTAAC
pNnANS-pABAI-F	TTGAATTCGAGCTCGGTACCTTTGGTCTTGGGGCCTTAGG
pNnANS-pABAI-R	GAGCACATGCCTCGAGTGGATCTGCTATAAACACAGCT
pNnFLS-pABAI-F	TTGAATTCGAGCTCGGTACCACTTGCAGCAATAGGTAGGC
pNnFLS-pABAI-R	GAGCACATGCCTCGAGAAGCTCCAGATAGGTTTCTAT
pNnF3H-pABAI-F	TTGAATTCGAGCTCGGTACCTCGACTGGCCCAATTCAAAA
pNnF3H-pABAI-R	GAGCACATGCCTCGAGTTGCGTGTCACGGTTATG
pNnCHS-pABAI-F	TTGAATTCGAGCTCGGTACCTTTGCCTCAACCCAAACCTG
pNnCHS-pABAI-R	GAGCACATGCCTCGAGCGCCTTGAGCTAGCTGAGAG
pNnGST2-pABAI-F	TTGAATTCGAGCTCGGTACCGTTGTTCCAAACCACACATTCT
pNnGST2-pABAI-R	GAGCACATGCCTCGAGTTTTCTGTGCAAAAGTGGAGACA
AD-NnMYB5-F	CAGATTACGCTCATATGATGGATGGTGGTTTGGGTTTG
AD-NnMYB5-R	TGCTTGGGTGGAATTCTCAATAACTCCACCACCTATGA
pNnFLS-p0800-F	CTTGATATCGAATTCCTGCAGACTTGCAGCAATAGGTAGGC
pNnFLS-p0800-R	CGCTCTAGAACTAGTGGATCCAAGCTCCAGATAGGTTTCTAT
pNnANS1-p0800-F	CTTGATATCGAATTCCTGCAGTTTGGTCTTGGGGCCTTAGG
pNnANS1-p0800-R	CGCTCTAGAACTAGTGGATCCTGGATCTGCTATAAACACAGCT
pNnF3H-p0800-F	CTTGATATCGAATTCCTGCAGTCGACTGGCCCAATTCAAAA
pNnF3H-p0800-R	CGCTCTAGAACTAGTGGATCCTTGCGTGTCACGGTTATG
pNnCHS-p0800-F	CTTGATATCGAATTCCTGCAGTTTGCCTCAACCCAAACCTG

pNnCHS-p0800-R	CGCTCTAGAACTAGTGGATCCCGCCTTGAGCTAGCTGAGAG
pNnGST2-p0800-F	CTTGATATCGAATTCCTGCAGGTTGTTCCAAACCACACATTCT
pNnGST2-p0800-R	CGCTCTAGAACTAGTGGATCCTTTTCTGTGCAAAAGTGGAGACA
NnMYB5-p62sk-F	CGCTCTAGAACTAGTGGATCCATGGATGGTGGTTTGGGTTTG
NnMYB5-p62sk-R	CTTGATATCGAATTCCTGCAGTCAATAACTCCACCACCTATGA
P1	GGAAGGATTATGCAATAGATCCGAGC
P2	CGAACTGGTTCAGACCCAAAATTAATTATT
P3	GAAGTTTTTGGAGTTGATTACGTGTGAC
P4	TCAATAACTCCACCACCTATGATTCAT

Supplemental Table S2. List of natural lotus accessions tested in Fig. 2b. The P1/P2 and P3/P4 primers were used for PCR verification.

No.	**Lotus Accession**	**Flower color**	**P1/P2***	**P3/P4****
1	Marker			
2	Zao Bai Xue	White	Yes	No
3	Guang Han Gong	White	Yes	No
4	Bai Ge	White	Yes	No
5	Xiao Bi Lian	White	Yes	No
6	Xi Wang	White	Yes	No
7	Dan Yue	White	Yes	No
8	Bai Ying Lian	White	Yes	No
9	Qing Xia	White	Yes	No
10	Qing He	White	Yes	No
11	Shui Jing Bai	White	Yes	No
12	Luan Feng Yu	White	Yes	No
13	Tai Zhen Chu Yu	White	Yes	No
14	Jiang Fu Qu	Red	No	Yes
15	Qiu Hong Yang	Red	No	Yes
16	Zhi Zun Ling Xiao	Red	No	Yes
17	Wu Zhi Zi 2 Hao	Red	No	Yes
18	Jin Ling Huo Du	Red	No	Yes
19	Jun Jie	Red	No	Yes
20	Hong Tong Shuai	Red	No	Yes
21	Dong Fang Hong	Red	No	Yes
22	Qian Tang Hong Lou	Red	No	Yes
23	Hong Deng Long	Red	No	Yes
24	Jian Xuan 17	Red	No	Yes
25	Chinese Ancient Lotus	Red	No	Yes
26	Hong Hua Jian Lian***	Red	Yes	Yes
27	Cun San Lian***	Red	Yes	Yes
28	Negative control			
29	Marker			

*The P1/P2 primers cover 2 kb sequence for white flower genotypes with the 80 kb deletion and 82 kb for red flower genotypes with the 80-kb sequence present, which can not be amplified by regular polymerase used in PCR reaction.

**The P3/P4 primers cover 3.6 kb sequence of the NnMYB5 gene in the deleted region, which is present in red flower genotypes, but deleted in the 80-kb PAV in white flowers.

***These two are hybrid cultivars with heterozygous genotypes.

Supplemental Table S3. Summary statistics of clean reads and RNA-seq mapping results.

Samples	Total Reads	Mapped Rate	Uniq Mapped Reads	Uniq Mapped Reads Ratio
QX2-1	24334644	92.29%	18841718	77.43%
QX2-2	23820860	91.22%	18067443	75.85%
QX2-3	24641784	91.04%	18617278	75.55%
QX3-1	24797114	90.80%	18711755	75.46%
QX3-2	23401357	90.46%	17675944	75.53%
QX3-3	24112698	91.25%	18469458	76.60%
ZGWS2-1	19382089	91.23%	16407032	85.91%
ZGWS2-2	23976710	90.95%	18567077	77.44%
ZGWS2-3	22956946	92.82%	18178982	79.19%
ZGWS4-1	23277011	91.92%	18278177	78.52%
ZGWS4-2	24277354	91.77%	18859523	77.68%
ZGWS4-3	24316265	90.73%	18697593	76.89%

Supplemental Table S4. List of the 170 overlapping DEGs between the white and red petal at developmental stages 2 and 3.

Gene	baseMean	log2Fold Change	lfcSE	stat	pvalue	padj	gene_name
NNU_23566	816.369	-13.072	1.193	-10.122	4.41E-24	8.70E-21	cysteine proteinase inhibitor 12-like
NNU_16737	426.109	-12.453	1.192	-9.608	7.37E-22	9.70E-19	respiratory burst oxidase homolog protein B-like
NNU_00015	405.574	-12.201	1.212	-9.239	2.48E-20	2.94E-17	mitochondrial inner membrane protease subunit 2
NNU_10157	216.162	-11.719	1.221	-8.782	1.61E-18	1.52E-15	protein SRG1-like
NNU_24884	311.176	-11.686	1.220	-8.757	2.01E-18	1.83E-15	probable F-box protein At4g22030
NNU_22382	176.163	-11.194	1.279	-7.969	1.60E-15	1.08E-12	--
NNU_22683	224.385	-11.186	1.229	-8.285	1.18E-16	9.31E-14	serine/threonine-protein kinase SAPK2
NNU_26180	142.305	-10.586	1.202	-7.973	1.55E-15	1.08E-12	membrane-bound transcription factor site-1 protease-like
NNU_15469	116.931	-10.585	1.205	-7.954	1.80E-15	1.19E-12	--
NNU_22188	135.865	-10.422	1.237	-7.615	2.64E-14	1.45E-11	"putative protein TPRXL, partial"
NNU_00312	124.065	-10.407	1.227	-7.667	1.75E-14	1.01E-11	ABSCISIC ACID-INSENSITIVE 5-like protein 2
NNU_15713	94.774	-10.246	1.226	-7.541	4.67E-14	2.40E-11	uncharacterized protein LOC104590517
NNU_23714	98.694	-10.228	1.211	-7.623	2.48E-14	1.40E-11	putative wall-associated receptor kinase-like 16
NNU_15649	89.133	-10.169	1.261	-7.271	3.56E-13	1.48E-10	17.8 kDa class I heat shock protein-like
NNU_19240	100.187	-10.163	1.250	-7.333	2.26E-13	9.90E-11	unnamed protein product
NNU_02817	99.740	-10.102	1.215	-7.491	6.84E-14	3.32E-11	aspartic proteinase-like protein 2 isoform X3
NNU_23141	91.880	-9.928	1.213	-7.358	1.87E-13	8.48E-11	protein ROOT PRIMORDIUM DEFECTIVE 1-like
NNU_13735	63.364	-9.720	1.246	-7.000	2.57E-12	9.21E-10	probable beta-D-xylosidase 7
NNU_08149	67.092	-9.642	1.218	-7.097	1.27E-12	5.20E-10	--
NNU_03879	57.828	-9.507	1.255	-6.781	1.20E-11	3.88E-09	unnamed protein product

NNU_09307	68.318	-9.477	1.533	-5.531	3.19E-08	6.09E-06	--
NNU_15654	66.851	-9.406	1.226	-6.855	7.12E-12	2.41E-09	17.8 kDa class I heat shock protein-like
NNU_25211	60.975	-8.960	1.256	-6.337	2.34E-10	6.75E-08	uncharacterized protein LOC104609439
NNU_11856	34.013	-8.521	1.262	-5.961	2.51E-09	5.87E-07	myosin-17-like
NNU_22390	26.905	-8.372	1.298	-5.678	1.36E-08	2.80E-06	17.8 kDa class I heat shock protein-like
NNU_13311	26.251	-8.371	1.297	-5.685	1.31E-08	2.72E-06	protein TRANSPARENT TESTA 12 isoform X2
NNU_13738	22.142	-8.201	1.315	-5.476	4.35E-08	8.05E-06	60S ribosomal protein L27-3
NNU_23154	24.698	-8.133	1.275	-5.592	2.24E-08	4.46E-06	--
NNU_00214	25.227	-7.951	1.288	-5.396	6.81E-08	1.23E-05	WRKY transcription factor 44-like isoform X2
NNU_19360	21.097	-7.842	1.313	-5.210	1.89E-07	3.22E-05	uncharacterized protein LOC104590526
NNU_09901	118.715	-7.811	0.770	-8.845	9.18E-19	9.45E-16	ABC transporter A family member 8-like isoform X1
NNU_04203	27.154	-7.756	1.303	-5.185	2.16E-07	3.65E-05	probable LRR receptor-like serine/threonine-protein kinase At4g29180
NNU_14773	17.829	-7.696	1.364	-4.909	9.13E-07	0.00013427	serine/threonine-protein kinase 4-like isoform X3
NNU_06805	16.418	-7.676	1.372	-4.867	1.13E-06	0.00016575	--
NNU_20253	218.200	-7.134	0.722	-8.500	1.90E-17	1.55E-14	(S)-scoulerine 9-O-methyltransferase-like
NNU_00289	12.788	-7.097	1.396	-4.367	1.26E-05	0.00145736	uncharacterized protein LOC104590939
NNU_24947	190.275	-7.082	0.612	-9.939	2.82E-23	5.13E-20	geraniol 8-hydroxylase-like
NNU_22643	345.328	-6.988	1.100	-5.443	5.24E-08	9.61E-06	bet1-like protein At4g14600
NNU_21960	11.104	-6.845	1.465	-3.989	6.62E-05	0.00661584	uncharacterized protein LOC104611094
NNU_05557	11.530	-6.787	1.463	-3.955	7.67E-05	0.00750079	uncharacterized protein LOC104601004
NNU_18233	8.454	-6.549	1.576	-3.520	0.0004314	0.03404077	pectinesterase 2-like
NNU_08863	44.834	-5.746	0.835	-5.686	1.30E-08	2.72E-06	putative ATP-dependent RNA helicase C550.03c
NNU_00034	193.292	-5.671	0.371	-12.595	2.26E-36	8.91E-33	uncharacterized protein LOC104610799
NNU_08874	210.973	-5.585	0.722	-6.350	2.16E-10	6.30E-08	myeloid leukemia factor 1-like

NNU_13308	43.386	-5.442	0.932	-4.764	1.90E-06	0.00026095	protein TRANSPARENT TESTA 12 isoform X1
NNU_04612	147.265	-4.961	0.572	-6.924	4.40E-12	1.51E-09	protein GAMETE EXPRESSED 1
NNU_19771	3826.870	-4.758	0.233	-16.141	1.31E-58	1.56E-54	uncharacterized protein LOC104611349 isoform X3
NNU_15680	337.714	-4.679	0.527	-6.975	3.06E-12	1.08E-09	probable glutathione S-transferase
NNU_21748	227.803	-4.476	0.508	-6.838	8.01E-12	2.63E-09	probable rhamnose biosynthetic enzyme 1
NNU_19770	1678.850	-4.287	0.272	-12.077	1.39E-33	4.70E-30	uncharacterized protein LOC104611349 isoform X2
NNU_11700	490.666	-4.199	0.323	-9.903	4.03E-23	6.81E-20	--
NNU_06538	57.147	-3.830	0.569	-4.971	6.67E-07	0.00010327	"D-glycerate 3-kinase, chloroplastic-like isoform X2"
NNU_10613	34.255	-3.824	0.627	-4.503	6.70E-06	0.00082998	uncharacterized protein LOC104589622
NNU_06703	374.750	-3.725	0.283	-9.632	5.85E-22	8.66E-19	"pentatricopeptide repeat-containing protein At3g46790, chloroplastic-like"
NNU_03850	110.012	-3.647	0.562	-4.714	2.43E-06	0.00032847	uncharacterized protein LOC104589371
NNU_13516	66.381	-3.541	0.459	-5.537	3.09E-08	5.94E-06	--
NNU_11876	94.754	-3.524	0.597	-4.231	2.33E-05	0.00258961	probable beta-D-xylosidase 7
NNU_10305	201.397	-3.522	0.416	-6.066	1.31E-09	3.19E-07	uncharacterized protein LOC104602797 isoform X2
NNU_03849	42.981	-3.364	0.660	-3.581	0.0003427	0.02807007	uncharacterized protein LOC104589372
NNU_03561	153.921	-3.279	0.344	-6.632	3.31E-11	1.03E-08	uncharacterized protein LOC104597501
NNU_21736	582.992	-3.153	0.498	-4.324	1.53E-05	0.00175281	putative glutamine amidotransferase YLR126C
NNU_21337	109.819	-3.101	0.338	-6.218	5.04E-10	1.37E-07	T-complex protein 1 subunit gamma-like isoform X2
NNU_02264	522.935	-2.980	0.282	-7.022	2.18E-12	8.07E-10	dual specificity protein phosphatase 1-like
NNU_03293	180.459	-2.963	0.425	-4.618	3.87E-06	0.00050675	pre-mRNA-splicing factor SYF1
NNU_18536	233.984	-2.946	0.525	-3.708	0.000209	0.01827129	indole-3-acetic acid-induced protein ARG7-like
NNU_12769	410.309	-2.918	0.474	-4.048	5.16E-05	0.00533236	glutamate receptor 2.7-like
NNU_16734	367.001	-2.858	0.253	-7.356	1.90E-13	8.48E-11	probable voltage-gated potassium channel subunit beta

NNU_25214	118.010	-2.724	0.291	-5.929	3.05E-09	6.94E-07	GTP cyclohydrolase 1-like
NNU_21373	11245.353	-2.672	0.216	-7.758	8.60E-15	5.09E-12	probable (S)-N-methylcoclaurine 3;-hydroxylase isozyme 2
NNU_18275	271.471	-2.653	0.264	-6.270	3.61E-10	9.94E-08	uncharacterized protein LOC104606014
NNU_21566	253.846	-2.611	0.221	-7.284	3.23E-13	1.37E-10	WAT1-related protein At5g07050-like isoform X2
NNU_15257	289.119	-2.608	0.416	-3.867	0.00011	0.01025252	uncharacterized protein LOC104589957
NNU_17478	4900.884	-2.566	0.226	-6.937	4.00E-12	1.39E-09	probable RNA-dependent RNA polymerase 1
NNU_25483	218.127	-2.507	0.257	-5.867	4.43E-09	9.89E-07	secoisolariciresinol dehydrogenase-like isoform X2
NNU_15435	198.897	-2.392	0.314	-4.437	9.11E-06	0.00108412	cation/H(+) antiporter 15-like
NNU_04090	494.957	-2.344	0.317	-4.242	2.22E-05	0.00250252	18.2 kDa class I heat shock protein-like
NNU_06754	178.281	-2.335	0.222	-6.013	1.83E-09	4.37E-07	2-hydroxyisoflavanone dehydratase-like isoform X3
NNU_03226	1050.269	-2.275	0.198	-6.431	1.27E-10	3.75E-08	uncharacterized protein LOC104606951 isoform X1
NNU_04181	576.991	-2.184	0.122	-9.680	3.66E-22	5.77E-19	anoctamin-like protein At1g73020
NNU_03121	2925.223	-2.124	0.289	-3.887	0.0001017	0.00959005	"magnesium-chelatase subunit ChlH, chloroplastic"
NNU_00972	1855.648	-1.844	0.195	-4.326	1.52E-05	0.00174665	uncharacterized protein LOC104611602 isoform X1
NNU_24310	2276.093	-1.803	0.233	-3.442	0.0005775	0.04423769	esterase-like isoform X1
NNU_05540	11273.342	-1.648	0.169	-3.831	0.0001277	0.01162967	flavonoids synthase/flavanone 3-hydroxylase-like
NNU_01203	1550.803	1.554	0.135	4.096	4.20E-05	0.00444365	WD repeat-containing protein 75
NNU_04072	623.300	1.590	0.163	3.619	0.0002963	0.02487213	uncharacterized protein LOC104585815
NNU_06963	1278.874	1.594	0.134	4.429	9.47E-06	0.00112033	subtilisin-like protease SBT3.5
NNU_04073	1139.123	1.897	0.190	4.723	2.33E-06	0.00031632	uncharacterized protein LOC104585815
NNU_16744	28578.280	1.968	0.197	4.917	8.78E-07	0.00013066	chalcone synthase 1-like
NNU_08477	1086.923	1.974	0.171	5.709	1.14E-08	2.41E-06	Pyruvate decarboxylase isozyme 2
NNU_10466	459.424	1.974	0.193	5.057	4.26E-07	6.82E-05	vacuolar protein sorting-associated protein 2 homolog 3-like
NNU_01884	328.596	2.039	0.271	3.827	0.0001298	0.01177377	probable magnesium transporter NIPA8 isoform X3
NNU_23269	211.059	2.058	0.290	3.654	0.0002587	0.02202665	uncharacterized protein LOC104604536

NNU_02567	1674.260	2.222	0.271	4.506	6.61E-06	0.00082364	uncharacterized protein LOC104604050 isoform X1
NNU_02186	2323.776	2.289	0.183	7.039	1.93E-12	7.26E-10	very-long-chain 3-oxoacyl-CoA reductase 1
NNU_12937	806.917	2.311	0.248	5.289	1.23E-07	2.14E-05	uncharacterized protein LOC104590170 isoform X1
NNU_02288	1647.803	2.312	0.186	7.053	1.75E-12	6.78E-10	ribose-phosphate pyrophosphokinase 4
NNU_12943	227.285	2.599	0.313	5.114	3.16E-07	5.12E-05	COBRA-like protein 6
NNU_21105	331.709	2.640	0.220	7.439	1.02E-13	4.82E-11	putative pentatricopeptide repeat-containing protein At1g68930
NNU_03204	148.891	2.706	0.279	6.123	9.16E-10	2.31E-07	"37S ribosomal protein S16, mitochondrial-like"
NNU_22952	9596.996	2.745	0.282	6.198	5.72E-10	1.53E-07	putative lipid-transfer protein DIR1
NNU_02812	94.617	2.780	0.456	3.903	9.50E-05	0.00899347	uncharacterized protein LOC104610418
NNU_10401	175.150	2.834	0.272	6.733	1.66E-11	5.32E-09	dirigent protein 7-like
NNU_15595	556.367	2.935	0.254	7.611	2.73E-14	1.47E-11	disease resistance protein RPM1-like isoform X1
NNU_19796	42.089	2.976	0.566	3.492	0.0004794	0.03770347	--
NNU_02799	83.097	2.978	0.401	4.935	8.03E-07	0.00012179	uncharacterized protein LOC104597433
NNU_06656	818.987	3.040	0.173	11.769	5.66E-32	1.68E-28	uncharacterized protein At4g22758-like
NNU_01692	4870.403	3.255	0.301	7.490	6.88E-14	3.32E-11	anthocyanidin 3-O-glucosyltransferase 2-like
NNU_07927	369.052	3.374	0.480	4.943	7.69E-07	0.00011747	transcription factor bHLH25-like
NNU_09299	1204.251	3.524	0.156	16.213	4.08E-59	9.67E-55	hypothetical protein POPTR_0005s08840g [Populus trichocarpa]
NNU_22627	169.956	3.622	0.545	4.807	1.53E-06	0.00021562	"protein CutA 1, chloroplastic-like isoform X2"
NNU_17479	1020.024	3.690	0.203	13.263	3.80E-40	1.80E-36	sec14 cytosolic factor-like isoform X1
NNU_06751	650.116	3.827	0.327	8.658	4.81E-18	4.22E-15	2-hydroxyisoflavanone dehydratase-like
NNU_02860	212.770	4.008	0.350	8.594	8.39E-18	7.10E-15	ras-related protein RABA4d
NNU_09866	925.564	4.021	0.268	11.275	1.74E-29	4.57E-26	protein TRANSPARENT TESTA 12 isoform X2
NNU_20906	566.053	4.074	0.497	6.186	6.16E-10	1.62E-07	GATA transcription factor 21-like
NNU_17649	70.135	4.082	0.741	4.159	3.19E-05	0.00345266	F-box protein At4g09920
NNU_05623	138.114	4.299	0.607	5.436	5.44E-08	9.91E-06	uncharacterized protein LOC104590014

NNU_08538	195.296	4.362	0.311	10.808	3.17E-27	7.50E-24	uncharacterized protein LOC104589039 isoform X1
NNU_05730	65.723	4.363	0.585	5.746	9.15E-09	1.95E-06	unnamed protein product
NNU_16593	46.492	4.568	0.876	4.073	4.64E-05	0.00481552	aldo-keto reductase
NNU_02910	123.923	4.610	0.539	6.702	2.06E-11	6.50E-09	uncharacterized protein LOC104597541
NNU_19679	54.923	4.705	1.076	3.444	0.0005727	0.04401723	uncharacterized protein LOC104588063
NNU_09103	1083.933	4.831	0.685	5.593	2.23E-08	4.46E-06	extracellular ribonuclease LE-like
NNU_25125	362.600	4.960	0.391	10.123	4.37E-24	8.70E-21	flotillin-like protein 4
NNU_07677	121.087	5.167	0.839	4.965	6.87E-07	0.00010564	"LL-diaminopimelate aminotransferase, chloroplastic-like isoform X3"
NNU_22805	78.697	5.525	1.028	4.403	1.07E-05	0.00125773	zeatin O-glucosyltransferase-like
NNU_05068	1884.713	5.714	0.299	15.742	7.82E-56	6.17E-52	peroxisomal (S)-2-hydroxy-acid oxidase GLO4-like
NNU_06612	450.450	5.725	0.503	9.395	5.72E-21	7.13E-18	"lycopene epsilon cyclase, chloroplastic isoform X2"
NNU_25678	74.312	5.951	0.948	5.222	1.77E-07	3.03E-05	F-box protein At3g07870-like
NNU_23143	9.523	6.143	1.427	3.604	0.0003134	0.02584899	protein ROOT PRIMORDIUM DEFECTIVE 1-like
NNU_09116	7.664	6.298	1.522	3.480	0.0005011	0.03927902	flotillin-like protein 4
NNU_04984	9.404	6.341	1.464	3.648	0.0002644	0.02243526	"calcium-transporting ATPase 12, plasma membrane-type-like"
NNU_06947	10.367	6.382	1.466	3.671	0.0002416	0.02079801	--
NNU_05732	52.292	6.498	0.924	5.947	2.73E-09	6.28E-07	hypothetical protein ZEAMMB73_659156 [Zea mays]
NNU_10247	17.124	6.887	1.363	4.318	1.58E-05	0.00179524	uncharacterized protein LOC104602732 isoform X3
NNU_06922	12.793	7.032	1.348	4.476	7.60E-06	0.0009268	--
NNU_02784	80.856	7.171	0.786	7.849	4.19E-15	2.61E-12	pentatricopeptide repeat-containing protein At3g12770-like
NNU_10084	14.728	7.271	1.364	4.598	4.27E-06	0.00055533	uncharacterized protein LOC104593377
NNU_03146	21.416	7.277	1.332	4.713	2.44E-06	0.00032847	uncharacterized protein LOC104585899
NNU_09212	17.819	7.501	1.416	4.590	4.44E-06	0.00057288	CTD nuclear envelope phosphatase 1 homolog

NNU_22778	32.520	7.545	1.279	5.119	3.08E-07	5.02E-05	probable LRR receptor-like serine/threonine-protein kinase At3g47570 isoform X1
NNU_23034	19.476	7.589	1.379	4.779	1.76E-06	0.00024517	--
NNU_04979	37.407	7.618	1.478	4.478	7.55E-06	0.0009258	"beta-Amyrin Synthase 1-like, partial"
NNU_19678	73.190	7.750	1.388	4.863	1.16E-06	0.0001678	--
NNU_02207	18.410	7.899	1.424	4.846	1.26E-06	0.00018205	RNA polymerase II C-terminal domain phosphatase-like 1
NNU_03887	30.516	7.919	1.307	5.294	1.19E-07	2.09E-05	hypothetical protein L484_010893
NNU_01592	25.588	7.981	1.359	5.135	2.82E-07	4.63E-05	uncharacterized protein LOC104606841
NNU_24729	19.489	8.000	1.495	4.681	2.85E-06	0.00038097	uncharacterized protein LOC104609700
NNU_18918	92.931	8.109	1.008	7.051	1.78E-12	6.80E-10	40S ribosomal protein S13-like
NNU_02485	30.564	8.126	1.450	4.913	8.95E-07	0.00013238	anthocyanidin reductase-like isoform X1 [Vitis vinifera]
NNU_21730	70.704	8.215	1.255	5.747	9.06E-09	1.95E-06	S-norcoclaurine synthase-like
NNU_10383	60.656	8.424	1.282	5.792	6.97E-09	1.51E-06	cellulose synthase, putative
NNU_19155	42.305	8.648	1.283	5.959	2.54E-09	5.89E-07	probable LRR receptor-like serine/threonine-protein kinase At1g56140 isoform X2
NNU_16404	87.457	8.781	1.230	6.326	2.52E-10	7.11E-08	pectinesterase
NNU_20167	42.291	8.905	1.303	6.069	1.29E-09	3.18E-07	alanine--glyoxylate aminotransferase 2 homolog 1, mitochondrial
NNU_18966	35.772	8.940	1.281	6.198	5.73E-10	1.53E-07	casparian strip membrane protein 2-like
NNU_00440	70.483	9.342	1.282	6.507	7.67E-11	2.30E-08	probable receptor-like protein kinase At5g39020
NNU_24113	96.386	9.516	1.215	7.008	2.42E-12	8.83E-10	acetylornithine deacetylase-like
NNU_25356	50.157	9.529	1.422	5.999	1.99E-09	4.71E-07	uncharacterized protein LOC104591318
NNU_10386	104.527	9.566	1.251	6.849	7.46E-12	2.49E-09	dammarenediol 12-hydroxylase-like
NNU_10885	83.395	9.599	1.214	7.081	1.43E-12	5.63E-10	"pentatricopeptide repeat-containing protein At4g16835, mitochondrial-like"

NNU_15937	130.128	10.191	1.212	7.584	3.36E-14	1.77E-11	vacuolar protein sorting-associated protein VTA1 homolog isoform X1
NNU_06620	141.535	10.632	1.200	8.026	1.01E-15	7.24E-13	F-box/FBD/LRR-repeat protein At5g56420-like
NNU_14867	223.628	10.718	1.307	7.435	1.05E-13	4.86E-11	uncharacterized protein LOC104607328
NNU_19156	171.699	10.822	1.207	8.140	3.96E-16	3.02E-13	probable LRR receptor-like serine/threonine-protein kinase At1g56140 isoform X2
NNU_09114	202.821	10.838	1.213	8.109	5.10E-16	3.77E-13	calcium-binding mitochondrial carrier protein SCaMC-1-like
NNU_12337	281.338	11.854	1.231	8.816	1.19E-18	1.17E-15	uncharacterized protein LOC104595426
NNU_03972	425.435	12.272	1.254	8.986	2.56E-19	2.88E-16	myb-related protein 330-like
NNU_03253	2906.416	12.423	0.780	14.654	1.28E-48	7.57E-45	glutathione S-transferase F11-like
NNU_24494	218.107	22.389	4.247	5.036	4.76E-07	7.46E-05	uncharacterized protein LOC104594007

Supplemental Table S5. Gene ID and accession numbers of *GSTs* used for phylogenetic analysis.

Gene name	Accession	Gene name	Accession
AtDHAR1	AT1G19570.1	NnGST1	XP_010270817.1
AtDHAR2	AT1G75270.1	NnGST2	XP_010270745.1
AtDHAR3	AT5G16710.1	NnGST3	XP_019054394.1
AtDHAR4	Q9FG59.1	NnGST4	DAD19041.1
AtDHAR5	AT1G19550.1	NnGST5	XP_010272271.1
AtEF1B1	AT1G09640.1	NnGST6	XP_010269470.1
AtEF1B2	AT1G57720.1	NnGST7	XP_010269471.1
AtGHR1	AT4G20940.1	NnGST8	XP_010259661.1
AtGHR2	NP_001119375.1	NnGST9	XP_010259659.1
AtGHR3	NP_001031672.1	NnGST10	XP_010259660.1
AtGHR4	POPTR_0001S34980G	NnGST11	XP_010242290.1
AtGSTF1	P42769.1	NnGST12	XP_010272551.1
AtGSTF10	AT2G30870.1	NnGST13	XP_010270300.1
AtGSTF11	AT3G03190.1	NnGST14	XP_010269081.1
AtGSTF12	AT5G17220.1	NnGST15	XP_010253346.1
AtGSTF13	AT3G62760.1	NnGST16	XP_010260035.1
AtGSTF14	AT1G49860.1	NnGST17	XP_010274730.1
AtGSTF2	AT4G02520.1	NnGST18	XP_010255416.1
AtGSTF3	AT2G02930.1	NnGST19	DAD25226.1
AtGSTF4	AT1G02950.1	NnGST20	XP_010253425.1
AtGSTF5	AT1G02940.1	NnGST21	XP_010259964.1
AtGSTF6	AT1G02930.1	NnGST22	XP_010270861.1
AtGSTF7	AT1G02920.1	NnGST23	XP_010253430.1
AtGSTF8	AT2G47730.1	NnGST24	XP_010264895.1
AtGSTF9	AT2G30860.1	NnGST25	XP_010276184.1
AtGSTL1	AT5G02780.1	NnGST26	XP_010252225.1
AtGSTL2	AT3G55040.1	NnGST27	DAD18685.1
AtGSTL3	AT5G02790.1	NnGST28	XP_010265991.1
AtGSTT1	AT5G41210.1	NnGST29	XP_010255950.1
AtGSTT2	AT5G41240.1	NnGST30	DAD31247.1
AtGSTT3	AT5G41220.1	NnGST31	XP_010258434.1
AtGSTU1	AT2G29490.1	NnGST32	XP_010271921.1
AtGSTU10	AT1G74590.1	NnGST33	XP_010279423.1
AtGSTU11	AT1G69930.1	NnGST34	XP_010271919.1
AtGSTU12	AT1G69920.1	NnGST35	XP_010279540.1
AtGSTU13	AT1G27130.1	NnGST36	XP_010278876.1
AtGSTU14	AT1G27140.1	NnGST37	XP_010257729.1
AtGSTU15	AT1G59670.1	NnGST38	XP_010276441.1
AtGSTU16	AT1G59700.1	NnGST39	XP_010243633.1

AtGSTU17	AT1G10370.1	NnGST40	XP_010243635.1
AtGSTU18	AT1G10360.1	NnGST41	DAD20129.1
AtGSTU19	AT1G78380.1	NnGST42	XP_010253818.1
AtGSTU2	AT2G29480.1	NnGST43	XP_019054125.1
AtGSTU20	AT1G78370.1	NnGST44	XP_019054126.1
AtGSTU21	AT1G78360.1	NnGST45	XP_010264128.1
AtGSTU22	AT1G78340.1	NnGST46	XP_010243633.1
AtGSTU23	AT1G78320.1	NnGST47	XP_010269443.1
AtGSTU24	AT1G17170.1	NnGST48	XP_010243636.1
AtGSTU25	AT1G17180.1	NnGST49	XP_010263987.1
AtGSTU26	AT1G17190.1	NnGST50	XP_010243637.1
AtGSTU27	AT3G43800.1	NnGST51	DAD20138.1
AtGSTU28	AT1G53680.1	NnGST52	DAD20132.1
AtGSTU3	AT2G29470.1	NnGST53	DAD33380.1
AtGSTU4	AT2G29460.1	NnGST54	DAD33379.1
AtGSTU5	AT2G29450.1	NnGST55	DAD20134.1
AtGSTU6	AT2G29440.1	NnGST56	XP_043696296.1
AtGSTU7	AT2G29420.1	NnGST57	XP_010269621.1
AtGSTU8	AT3G09270.1	NnGST58	DAD20131.1
AtGSTU9	AT5G62480.1	NnGST59	XP_010243638.1
AtGSTZ1	AT2G02390.1	NnGST60	XP_010248269.1
AtGSTZ2	AT2G02380.1	NnGST61	DAD20133.1
AtTCHQD1	NP_001031292.1	NnGST62	DAD20130.1
ZmGSTF1	NP_001105626.1	NnGST63	DAD18886.1
ZmGSTF2	NP_001105366.1	NnGST64	DAD38436.1
ZmGSTF3	NP_001105720.2	NnGST65	XP_010245128.1
ZmGSTF4	ACG25283.1	NnGST66	XP_010245130.1
ZmGSTF5	NP_001105111.2	NnGST67	XP_010245129.1
ZmGSTF7	ACF80172.1	PhAN9	CAA68993.1
ZmGSTU1	NP_001104989.1	VvGSTF12	XP_034683898.1
ZmGSTU10	ONM39050.1	FvRAP	XP_004288578.1
ZmGSTU11	AAG34828.1	PpRiant	ALE31200.1
ZmGSTU14	XP_008674526.1	MdGSTF6	NP_001315851.1
ZmGSTU2	AQK85998.1		
ZmGSTU24	ONM39050.1		
ZmGSTU25	ONM39033.1		
ZmGSTU26	NP_001266752.2		
ZmGSTU27	ACF84945.1		
ZmGSTU3	AQK85978.1		
ZmGSTU6	NP_001152168.1		
ZmGSTU7	NP_001153168.1		
ZmGSTU8	NP_001152486.1		
ZmGSTZ1	XP_008661994.1		