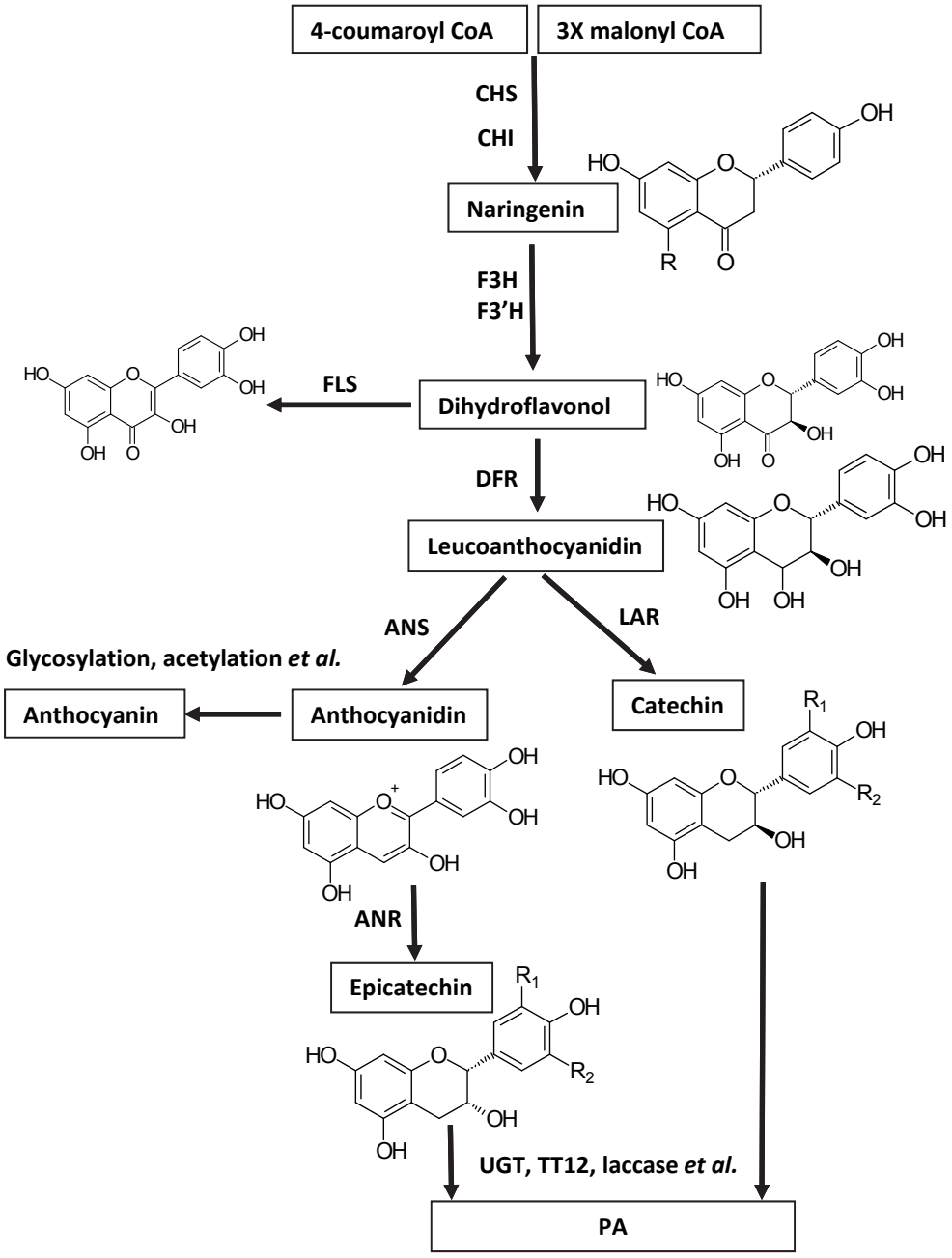

Figure S1. Biosynthesis pathway of PA.
CHS: Chalcone synthase, CHI: Chalcone isomerase, F3H: Flavanone 3 hydroxylase, F3'H: Flavonoid 3' hydroxylase, DFR: dihydroflavonol 4-reductase, FLS: Flavonol synthase, ANS: anthocyanidin synthase, LAR: Leucoanthocyanidin reductase, ANR: anthocyanidin reductase, UGT: UDP glucosyltransferase, TT12: transparent testa 12

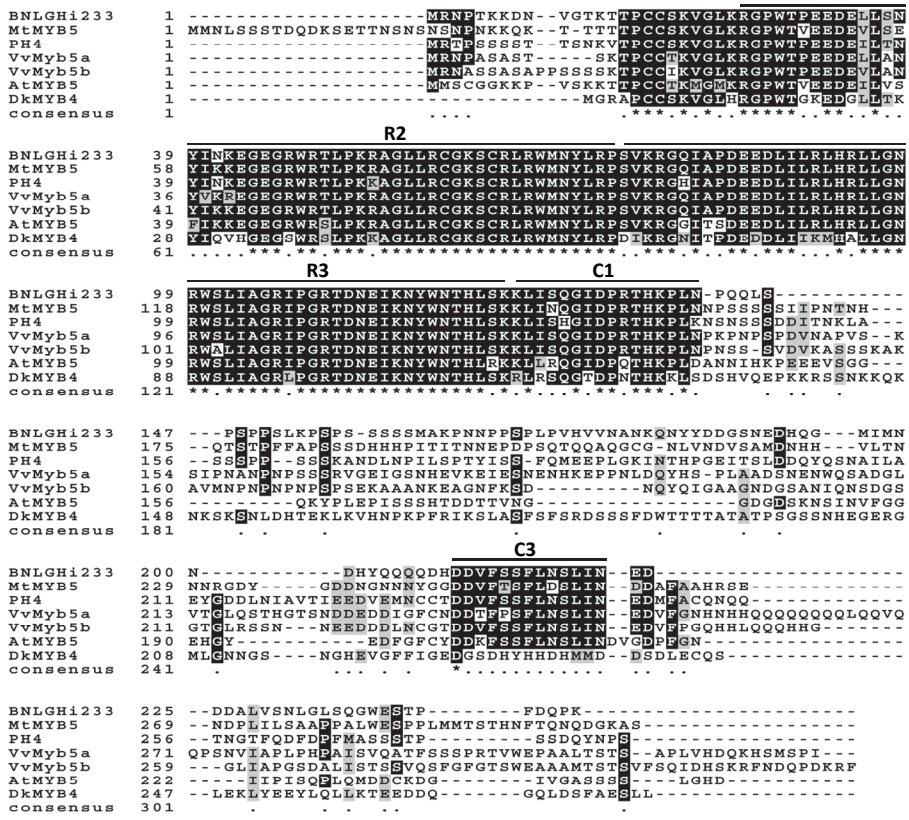

Figure S2. Mt MYB5 is the homolog of At MYB5.
Protein sequences of BNLGHi233, PH4 (AY973324), Vv MYB5a, Vv MYB5b, At MYB5, DkMYB4 and Mt MYB5 were aligned by the Muscle program (http://www.ebi.ac.uk/Tools/msa/muscle/). Multiple sequence alignment image was produced by Boxshade (http://www.ch.embnet.org/software/BOX_form.html. Genbank accession numbers: At MYB5 (NP_187963.1), Dk MYB4 (BAI49721.1), Vv MYB5a (NP_001268108.1), Vv MYB5b (NP_001267854.1), Mt MYB5 (XP_003601609.1), PH4 (AAY51377.1), BNLGHi233 (AAK19611.1).

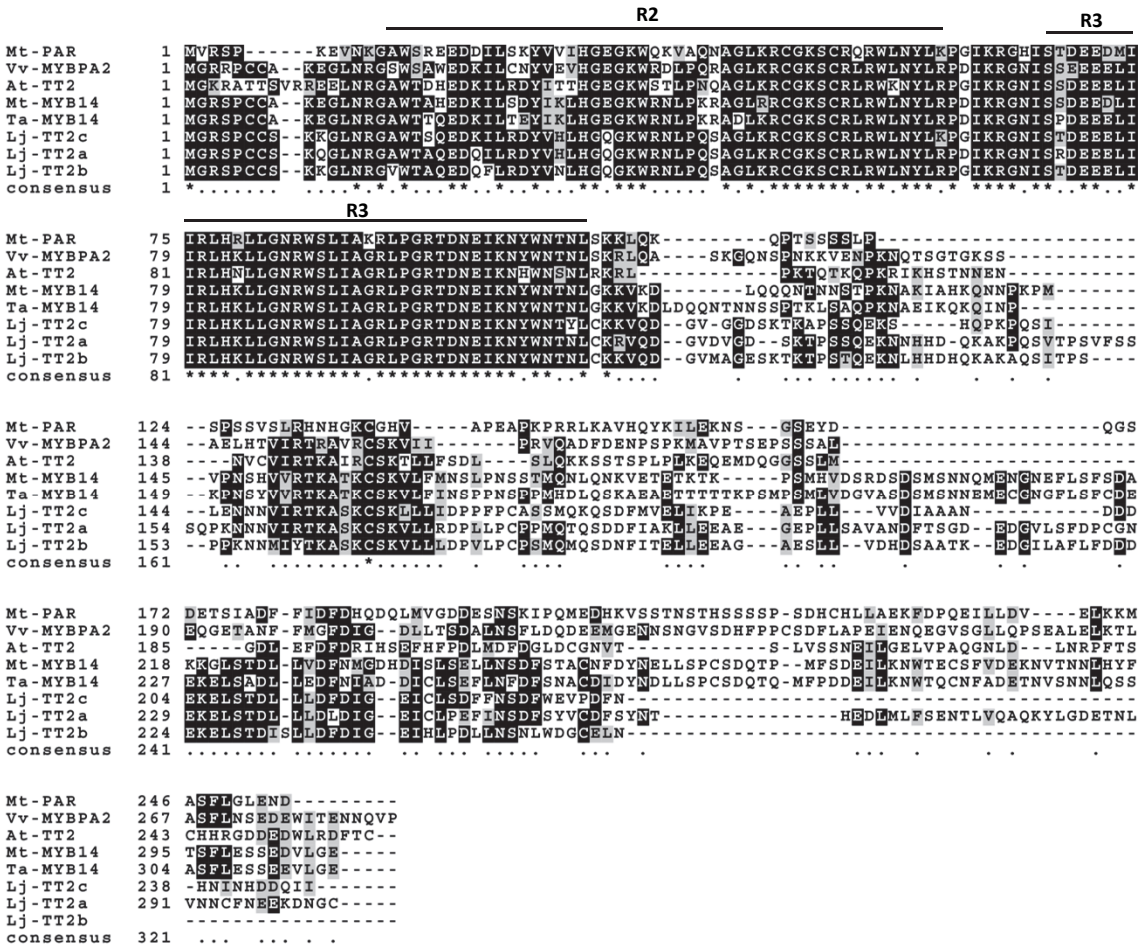

Figure S3. Multiple alignment of the homologs of Mt MYB14.
Protein sequences of Mt PAR, VV MYBPA2, At TT2, Mt MYB14, Ta MYB14, Lj TT2a, Lj TT2b, Lj TT2c were aligned by the Muscle program (http://www.ebi.ac.uk/Tools/msa/muscle/). Multiple alignment image was produced by Boxshade (http://www.ch.embnet.org/software/BOX_form.html. Genbank accession numbers: At TT2 (NP_198405.1), Lj TT2a (BAG12893.1), Lj TT2b (BAG12894.1), Lj TT2c (BAG12895.1), Ta MYB14-2(AFJ53054), Mt MYB14-2 (Mt3.5v4 contig_238935_1, Mt4.0v1 Medtr4g125520.1), Mt PAR (XP_003627264.1), Vv PA2 (NP_001267953.1).

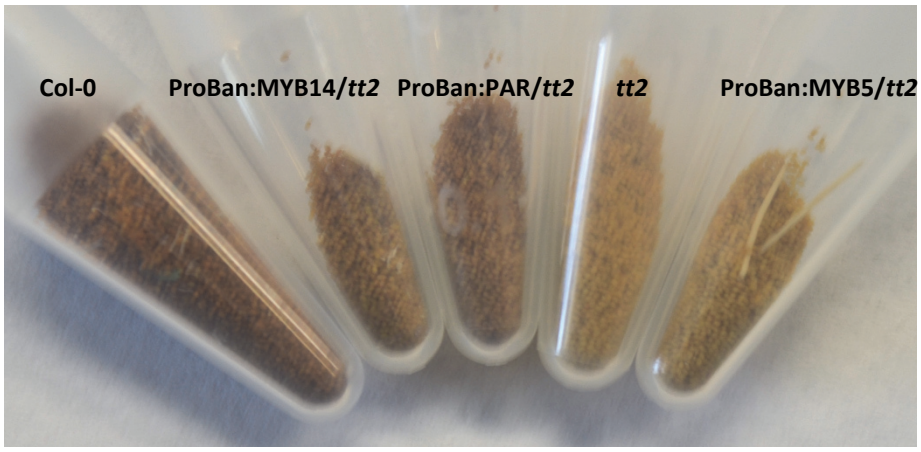

Figure S4. Genetic complementation of Arabidopsis *tt2* with Mt *Myb14*, Mt *Par* and Mt *Myb5* Medicago *Myb14, Par* and *Myb5* were driven by Arabidopsis *Anr* (Banylus) promoter and used to complement Arabidopsis *tt2* mutant (Salk_005260). Fifteen to twenty independent lines were observed and representative lines were presented. Mt MYB14 and Mt PAR clearly rescue the *tt2* phenotype, while Mt MYB5 does not.

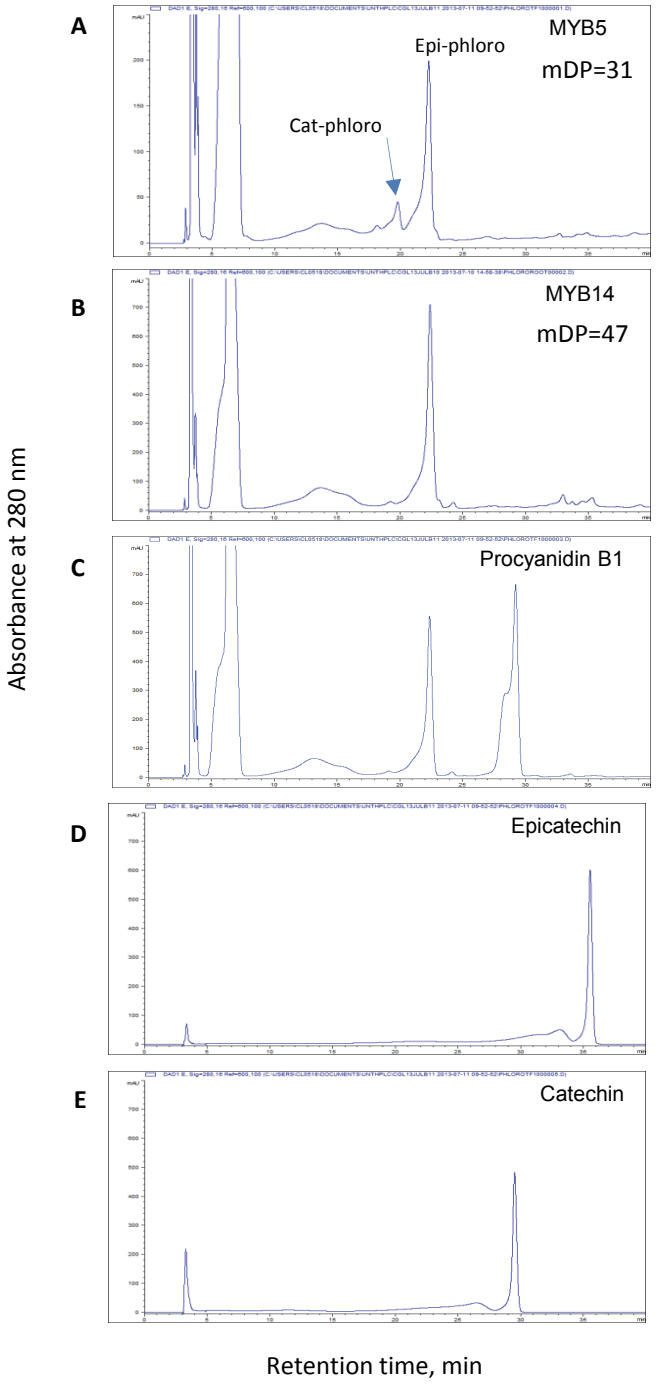

Figure S5. Phloroglucinolysis analysis of PAs in MYB5 and MYB14 over-expressing hairy roots. PAs purified by Sephadex LH20 resins were hydrolyzed by acidic phloroglucinol/mehanol solution. Hydrolyzed PA were analyzed by reverse phase HPLC. A, PAs purified from MtMYB5 over-expressing hairy roots. B, PAs purified from MtMYB14 over-expressing hairy roots. C, Procyanin B1 standard analyzed by phloroglucinolysis. D, Epicatechin standard. E, Catechin standard. mDP: mean degree of polymerization.

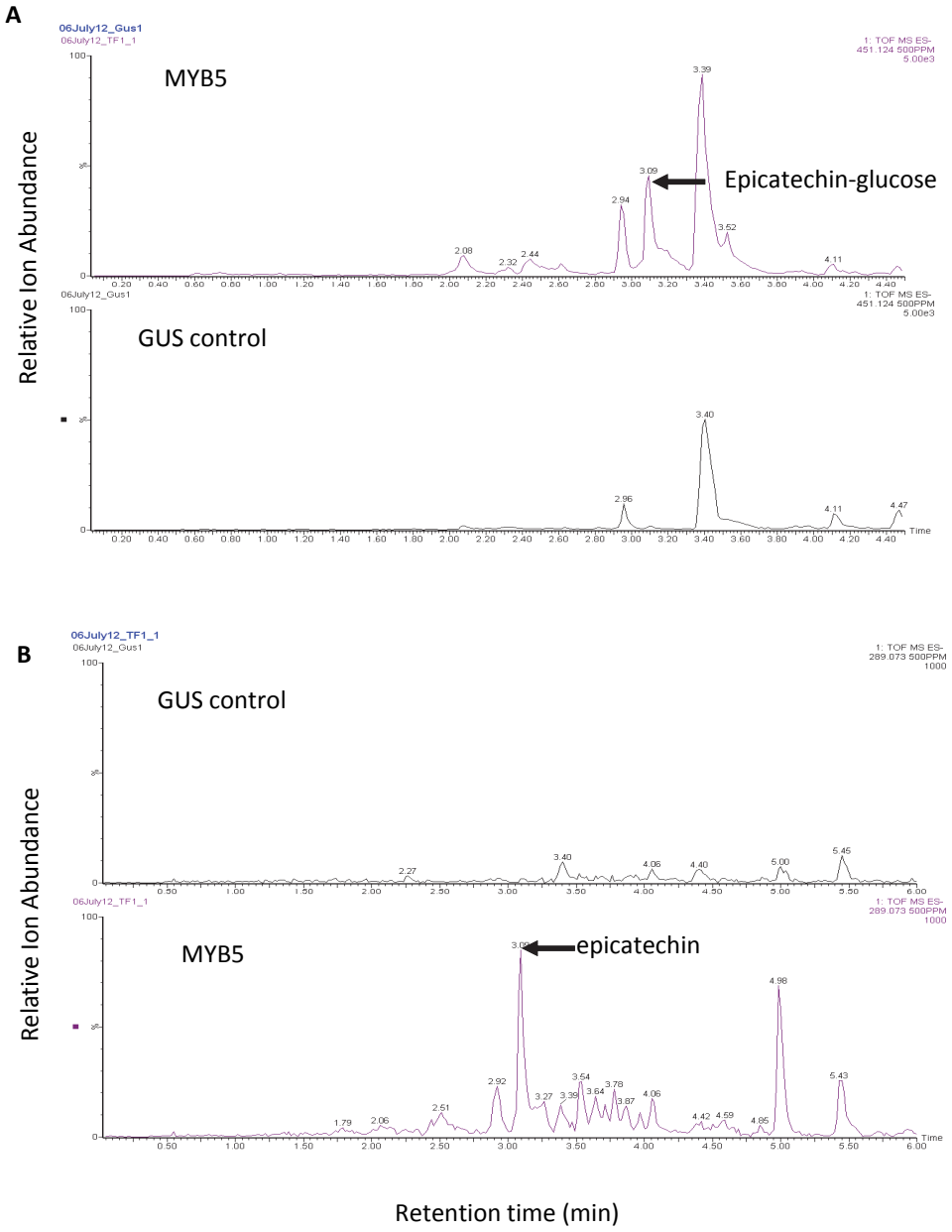

Figure S6. UPLC-MS detection of epicatechin and epicatechin-3'-*O*- glucoside ions in hairy roots over-expressing MYB5.
A, epicatechin-3'-O- glucoside (m/z 451.124). B, epicatechin (m/z 289.073).

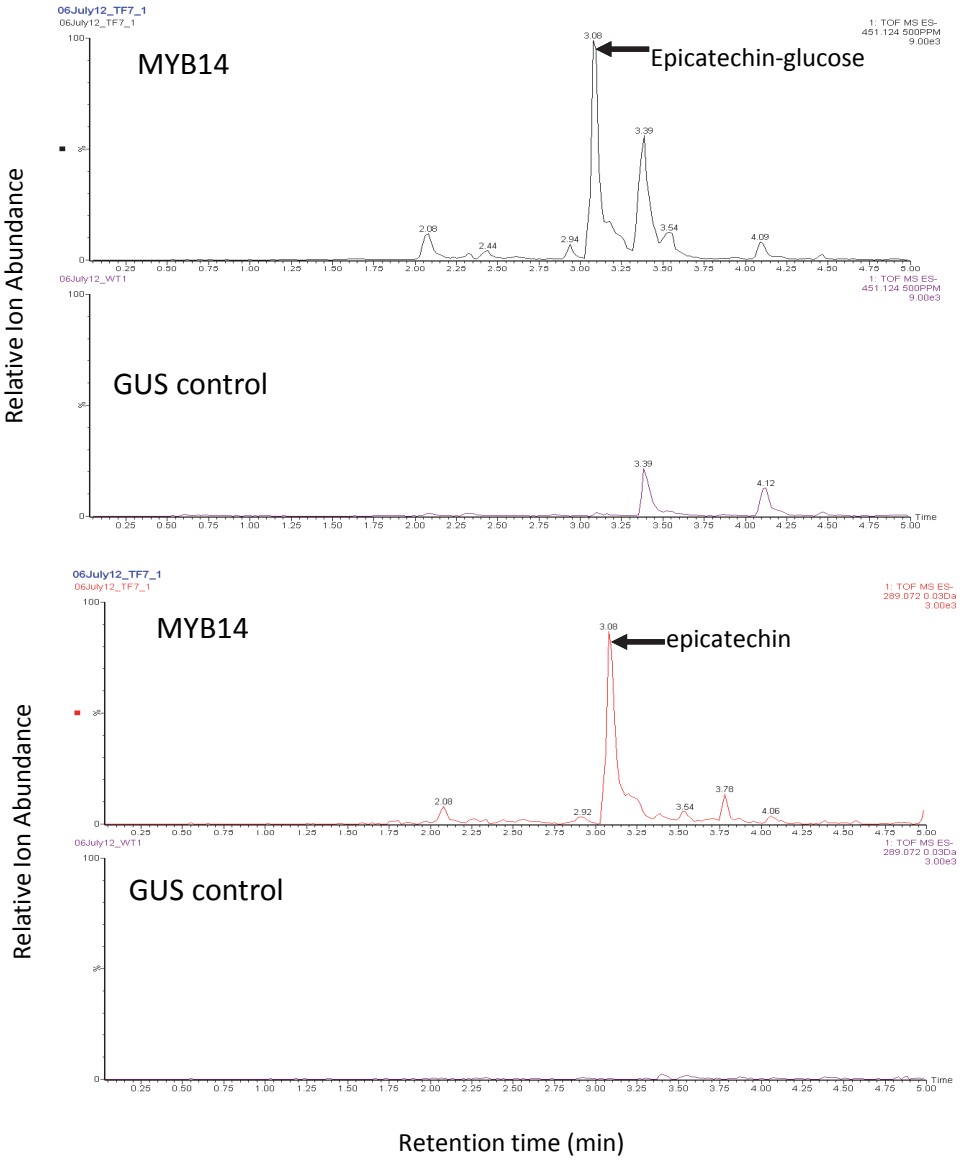

Figure S7. UPLC-MS detection of epicatechin and epicatechin-3'-*O*- glucoside ions in hairy roots over-expressing MYB14.
A, epicatechin-3'-O- glucoside (m/z 451.124). B, epicatechin (m/z 289.073).

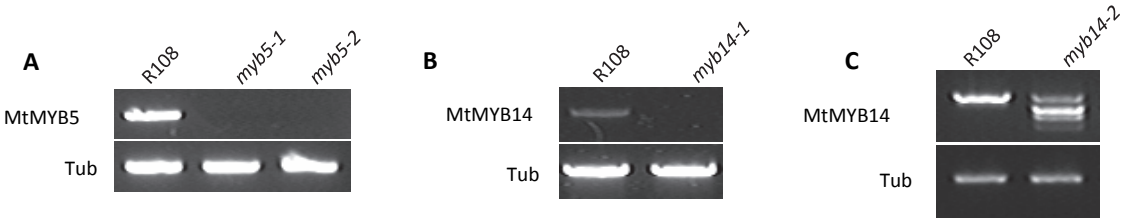

Figure S8. RT-PCR to detect MYB5 and MYB14 transcripts in *myb5* and *myb14* mutant seeds.
A, *myb5*. B, *myb14-1*. C, *myb14-2*.

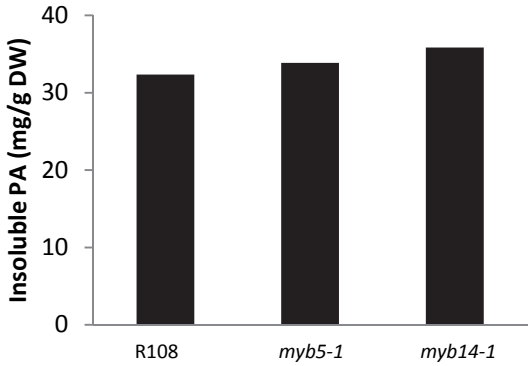

Figure S9. Butanol-HCl analysis of insoluble PA-like material in R108, *myb5* and *myb14* mutant seeds.
Insoluble PAs were hydrolyzed in butanol-HCl solution and levels expressed as procyanidin B1 equivalents.

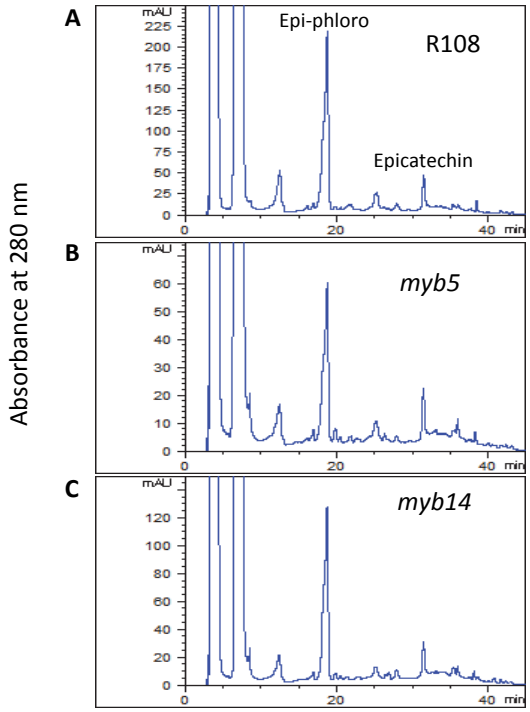

Figure S10. Phloroglucinolysis analysis of soluble PAs in the seeds of wild-type (R108), *myb5-1* and *myb14-1*.
PAs purified with Sephadex LH20 resin were hydrolyzed by acidic phloroglucinol/methanol solution. Hydrolyzed PAs were analyzed by reverse phase HPLC. A, PAs purified from R108 seeds. B, PAs purified from *mtmyb5* seeds. C, PAs purified from *myb14* seeds.

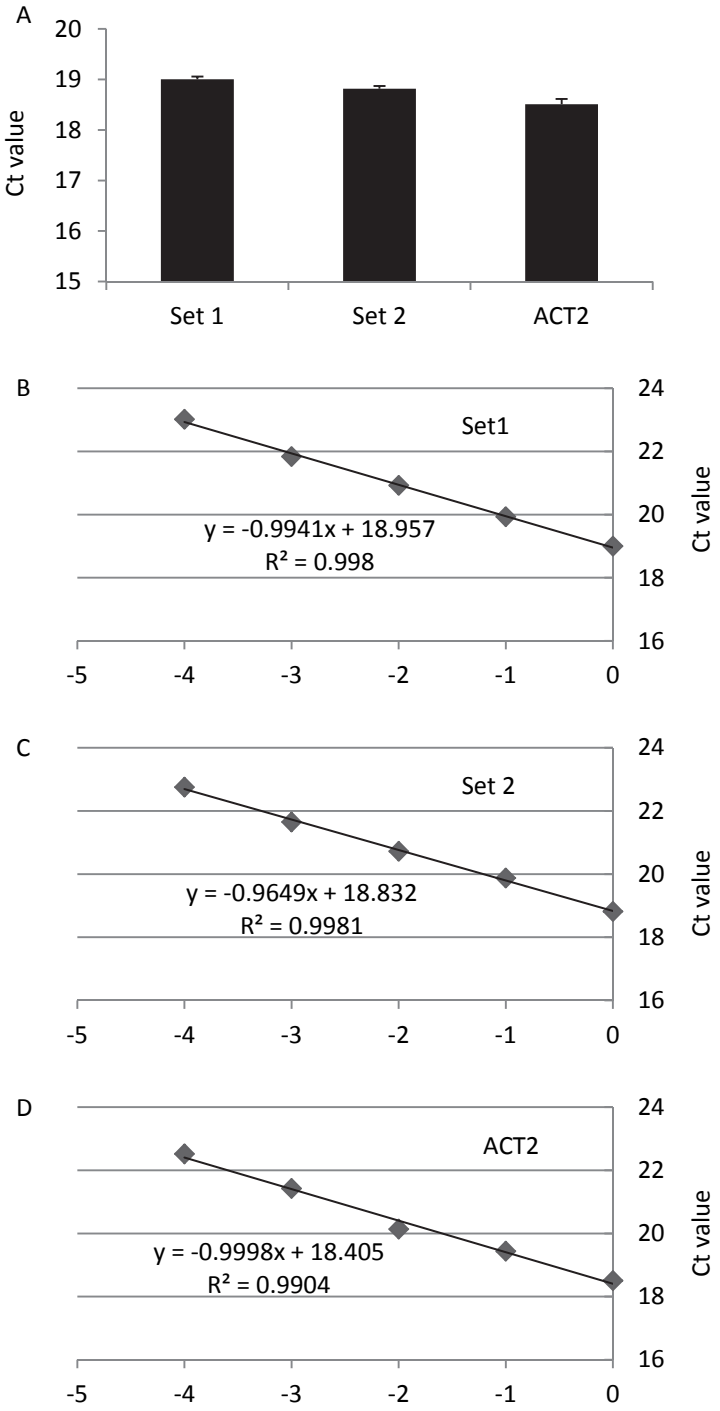

Figure S11. Measurement of MYB14 copy number in the *Medicago truncatula* genome by qPCR.
A, The Ct values of *MYB14* determined by two sets of primers (Set 1 and Set 2, Supplemental Table 4) and reference *ACT2* genes.
B, The amplification efficiency of primer set 1 determined by serial dilution of genomic DNA. Y-axis, Ct value. X-axis, log2 value of dilution factors.
C, The amplification efficiency of primer set 2 determined by serial dilution of genomic DNA. Y-axis, Ct value. X-axis, log2 value of dilution factors.
D, The amplification efficiency of primer pairs of ACT2 determined by serial dilution of genomic DNA. Y-axis, Ct value. X-axis, log2 value of dilution factors.

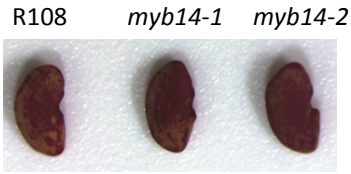

Figure S12. Ruthenium red staining showing that mucilage levels in *myb14* are similar to those in wild-type R108.