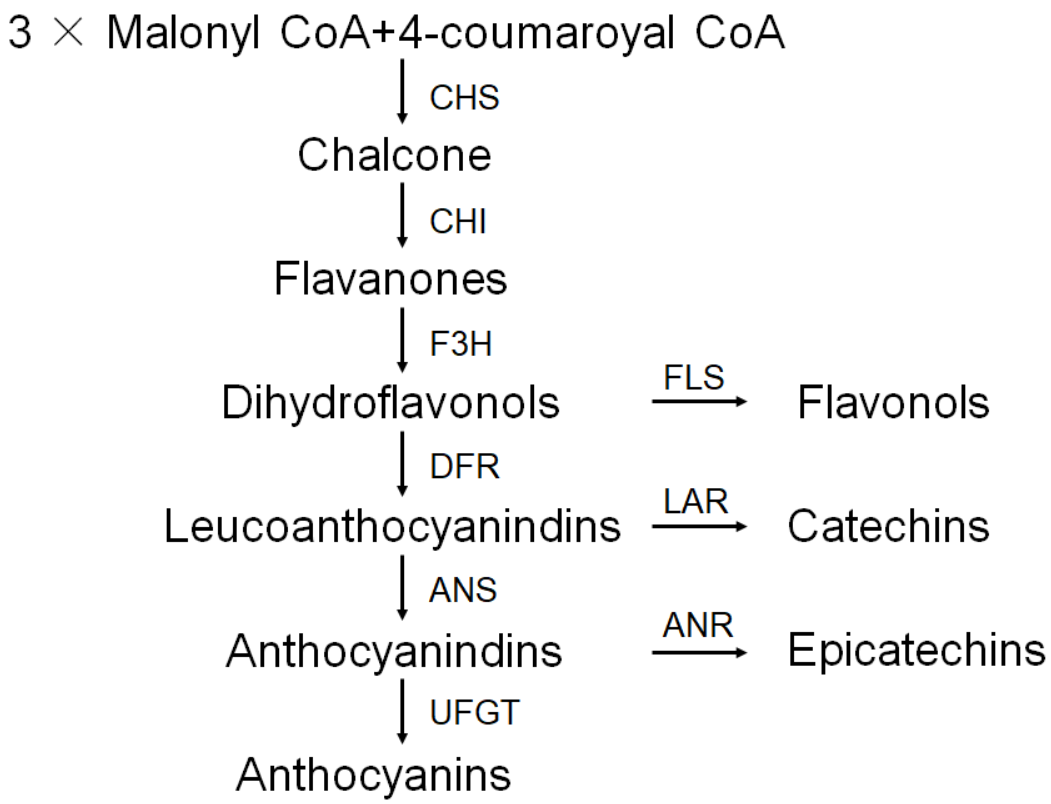

Figure S1. Simplified diagram of genes involved in flavonoid biosynthesis.

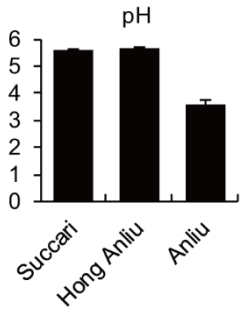

Figure S2. pH in juice from breaker stage ‘Anliu’, ‘Hong Anliu’ and ‘Succari’ fruits. Error bars represent the mean ± SD of four biological replicates.

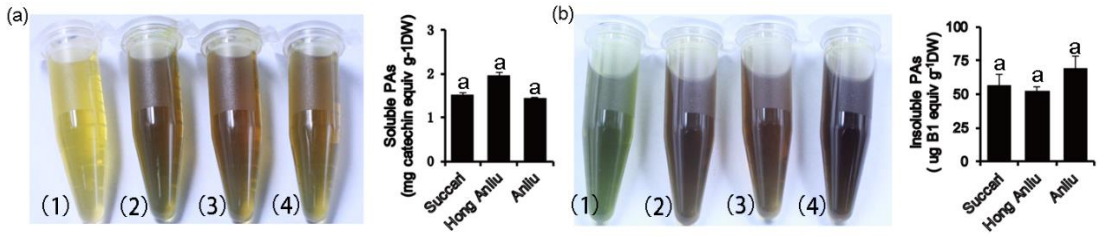

Figure S3. Characterization of PA levels of flavedos in ‘Anliu’, ‘Hong Anliu’ and ‘Succari’. (a) The contents and extracts after DMACA staining of soluble PA in flavedos. (b) The contents and extracts after DMACA staining of insoluble PA in flavedos. (1) Blank control, (2) ‘Succari’, (3) ‘Hong Anliu’, (4) ‘Anliu’. DW, dry weight. Error bars represent the mean ± SD of three biological replicates.

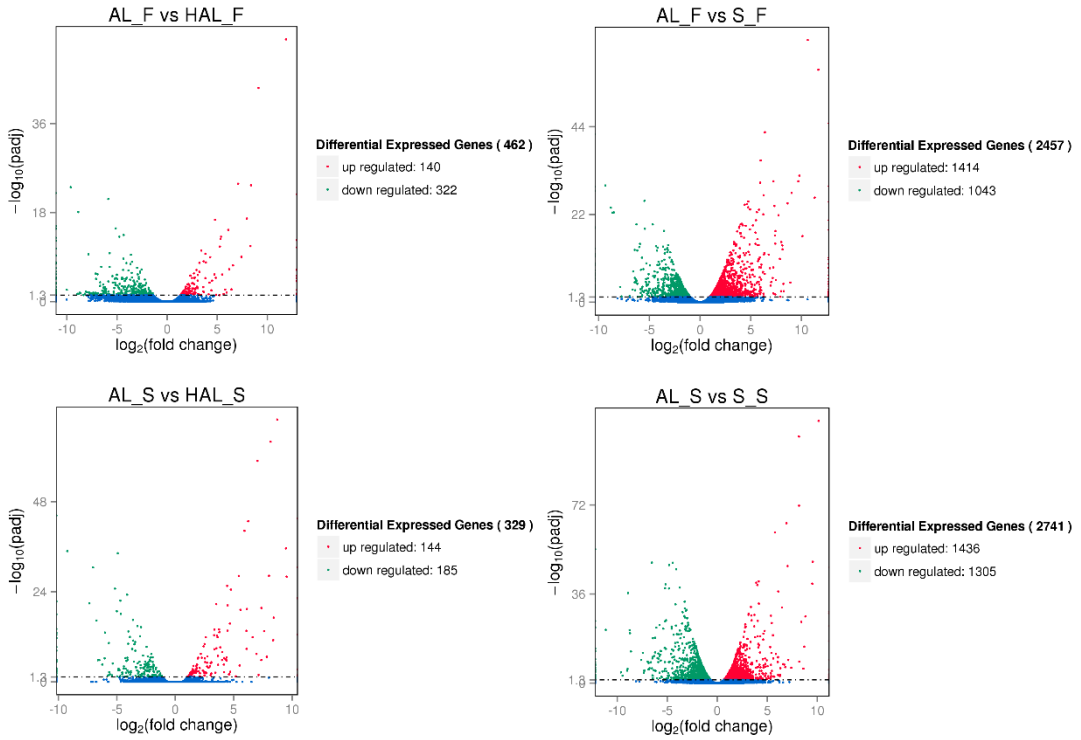

Figure S4. Volcano plots showing the whole distribution of DEGs in the AL_F vs HAL_F, AL_F vs S_F, AL_S vs HAL_S and AL_S vs S_S comparisons. Transcriptome of the 'Anliu' pulps (AL_F) was compared with that of (a) 'Hong Anliu' pulps (HAL_F) and (b) 'Succari' pulps (S_F). Transcriptome of the 'Anliu' seeds (AL_S) was compared with that of (c) 'Hong Anliu' seeds (HAL_S) and (d) 'Succari' seeds (S_S). The x-axis indicated log2 of estimated fold change and the y-axis indicated negative log10 of the adjusted *P*-value. Red and green dots represent significantly up-regulated and down-regulated genes, respectively, while blue dots represent non-significantly expressed genes.

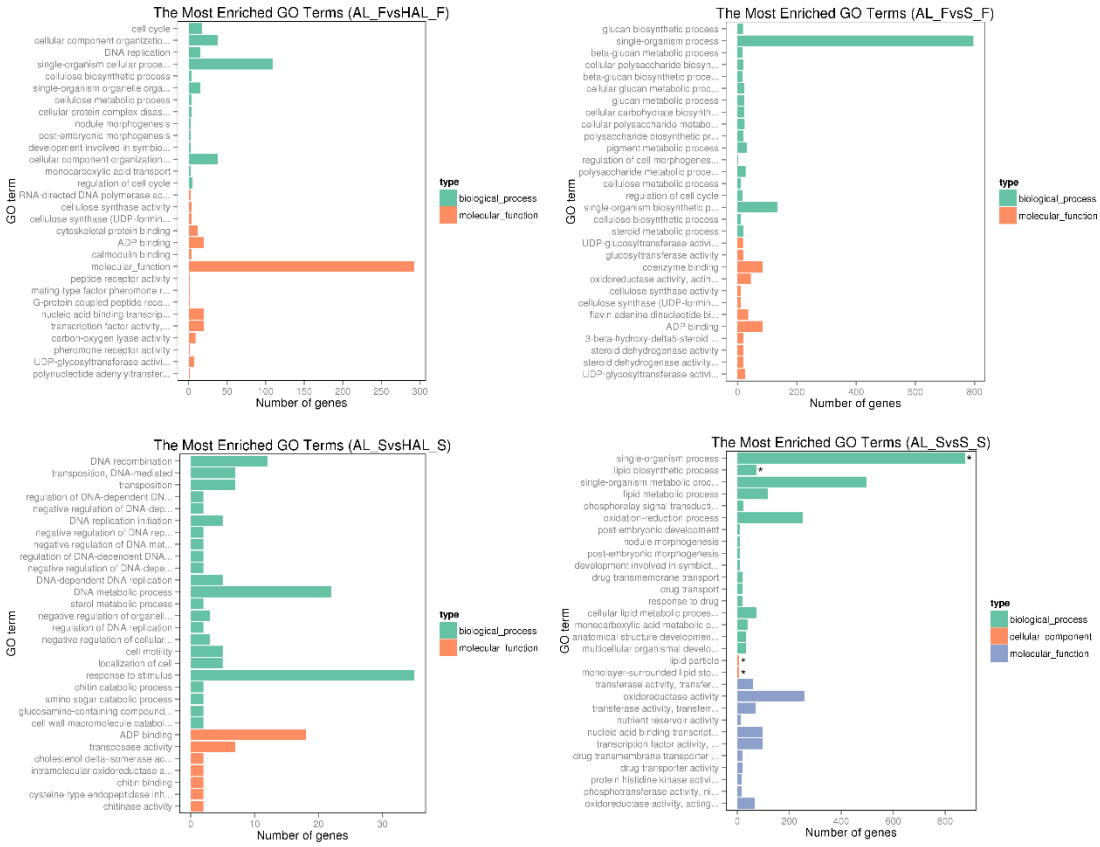

Figure S5. GO terms of DEGs in the AL_F vs HAL_F, AL_F vs S_F, AL_S vs HAL_S and AL_S vs S_S comparisons. AL_F, 'Anliu' pulps; HAL_F, 'Hong Anliu' pulps; S_F, 'Succari' pulps; AL_S, 'Anliu' seeds; HAL_S, 'Hong Anliu' seeds and S_S, 'Succari' seeds. The 30 GO terms with the most significant enrichment were selected and shown in the figure; if there are less than 30, all of them were shown. Asterisks indicate significantly enriched GO terms.

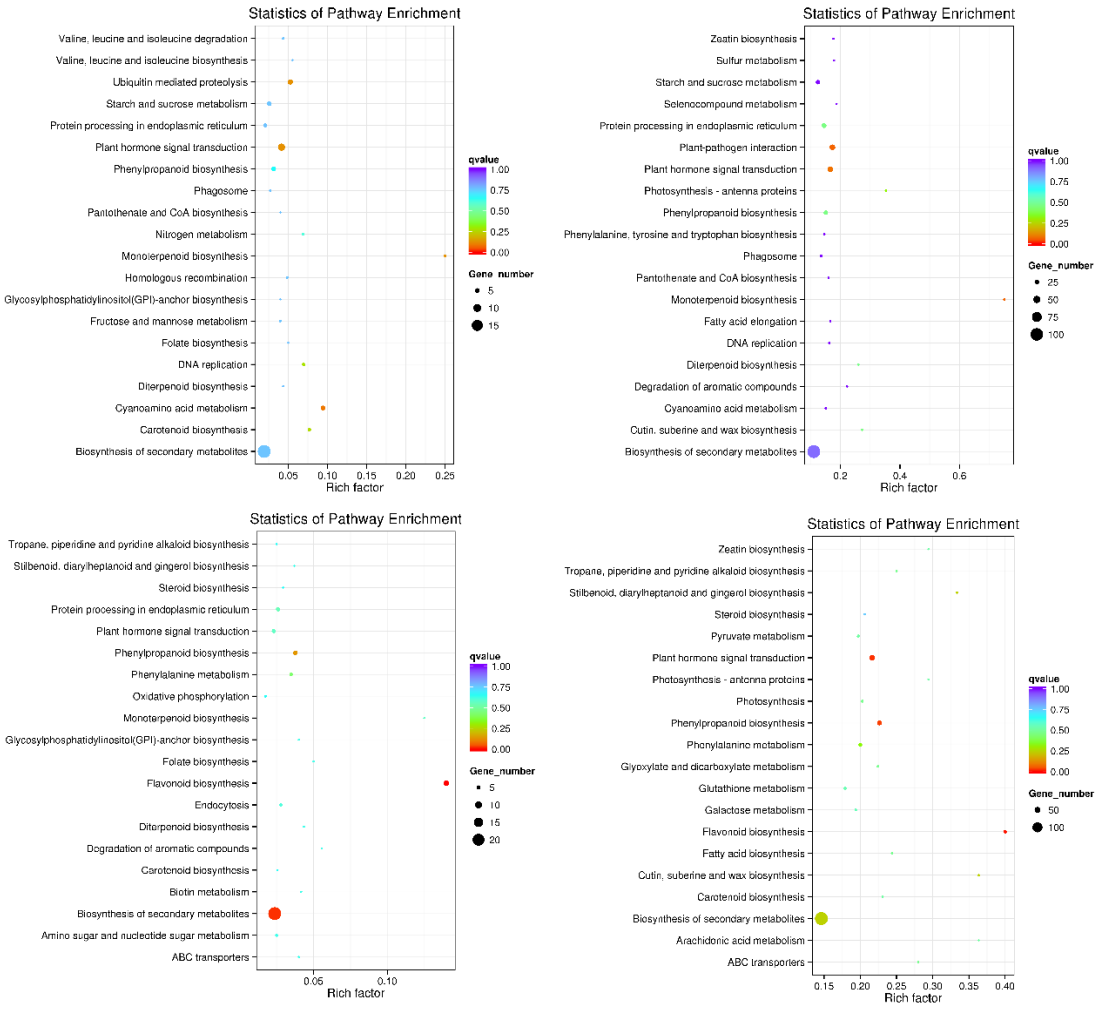

Figure S6. KEGG classifications of DEGs in the AL_F vs HAL_F, AL_F vs S_F, AL_S vs HAL_S and AL_S vs S_S comparisons. AL_F, 'Anliu' pulps; HAL_F, 'Hong Anliu' pulps; S_F, 'Succari' pulps; AL_S, 'Anliu' seeds; HAL_S, 'Hong Anliu' seeds and S_S, 'Succari' seeds. The y-axis represents the pathway name, the x-axis represents the rich factor, the size of the dots indicates the number of differentially expressed genes in this pathway, and the colour of the dots corresponds to different Q-value ranges.

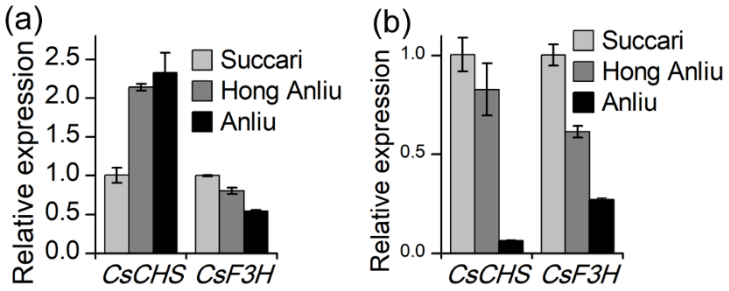

Figure S7. (a) Relative expression patterns of *CsCHS* and *CsF3H* in the seeds of the three citrus varieties. (b) Relative expression patterns of *CsCHS* and *CsF3H* in the pulps of the three citrus varieties. Error bars represent the mean ± SD of three biological.

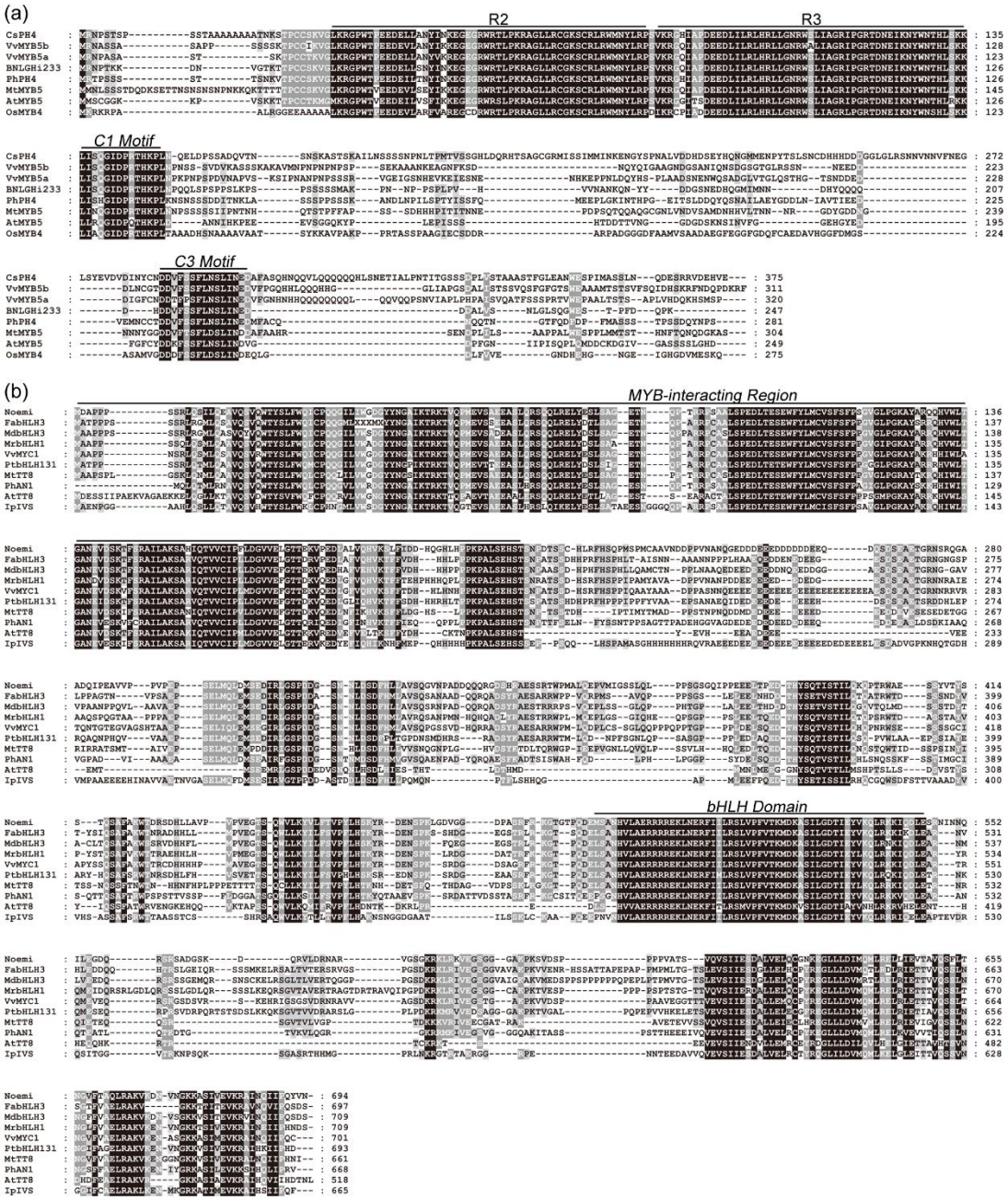

Figure S8. **Protein sequence alignment of CsPH4 and Noemi.** (a) Multiple sequence alignment of CsPH4 and related proteins from other plants. (b) Multiple sequence alignment of Noemi and related proteins from other plants. Identical residues are shown in black, conserved residues in dark gray, and similar residues in light gray. The horizontal lines mark the conserved domains and motifs.

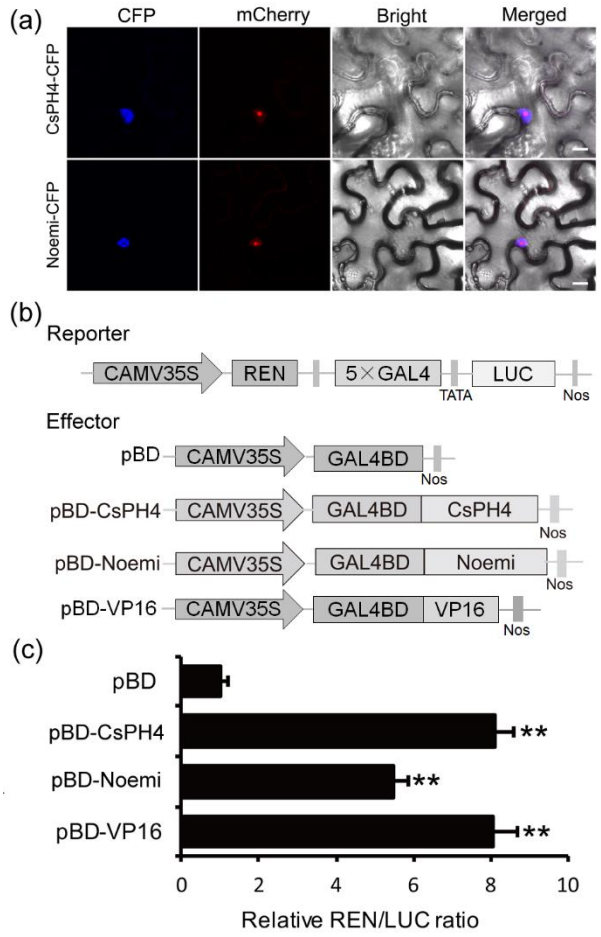

Figure S9**. Subcellular localization and transcriptional activity of *CsPH4* and *Noemi*.** (a) Subcellular localization of *CsPH4* and *Noemi* in tobacco epidermal cells. CsPH4-CFP and Noemi-CFP were co-transformed with mCherry, which was used as a nuclear marker. CsPH4-CFP and Noemi-CFP, cyan fluorescent signal; mCherry, red fluorescent signal; Bars = 10 μm. (b) Schematic diagrams of vectors used for the transcriptional activity assay. pBD-Noemi and pBD-Noemi, vectors containing *CsPH4* and *Noemi*, respectively; pBD and pBD-VP16 served as the negative and positive controls, respectively. (c) The transcriptional activity of *CsPH4* and *Noemi*. Transcriptional activity of *CsPH4* and *Noemi* was quantified using a dual-luciferase assay. Error bars represent the mean ± SD of eight biological replicates. Asterisks indicate significant differences relative to the empty vector control (**, $P < 0.01$).

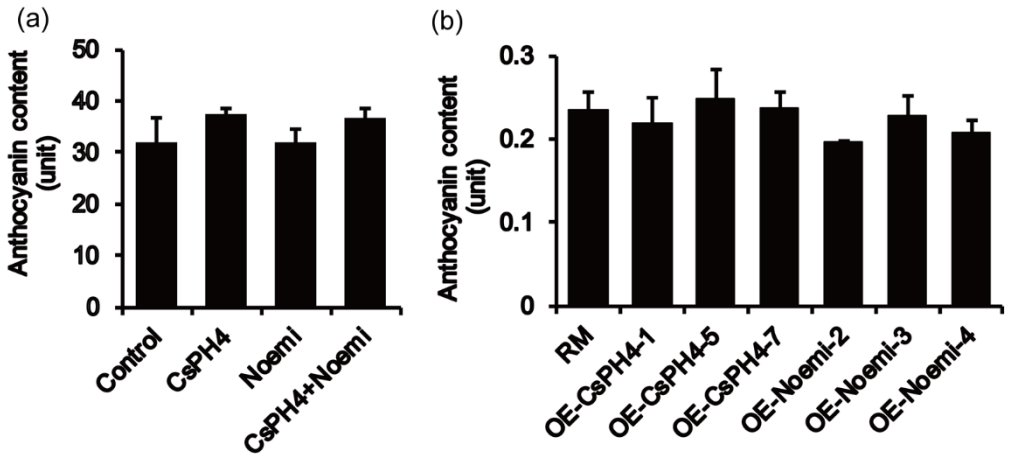

Figure S10. (a) The anthocyanin content (unit) of the transient overexpression of *CsPH4*, *Noemi*, *CsPH4* plus *Noemi* and an empty vector control in tobacco (*N. benthamiana*) leaves. (b) The anthocyanin content (unit) of transgenic citrus callus. (A530–0.25*A657)/dry weight was considered one anthocyanin unit. Error bars represent the mean ± SD of three biological replicates. The experiment was repeated independently three times with similar results obtained for each.

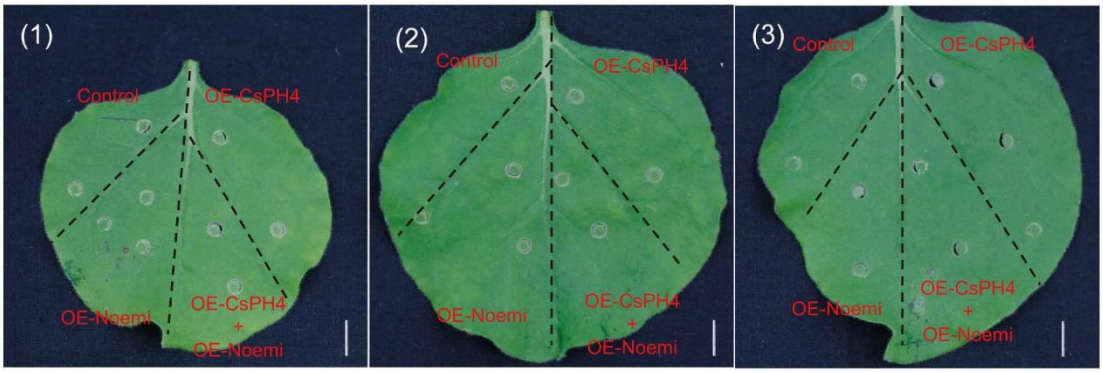

Figure S11. Phenotype of the transient overexpression of *CsPH4* (OE-CsPH4), *Noemi* (OE-Noemi), *CsPH4* plus *Noemi* (OE-CsPH4 + OE-Noemi) and an empty vector control (Control) in tobacco (*N. benthamiana*) leaves. Different regions in each tobacco leaf represent different combinations of transient overexpression. (1), (2), and (3) represent the three biological replicates. Scale bar = 1 cm.

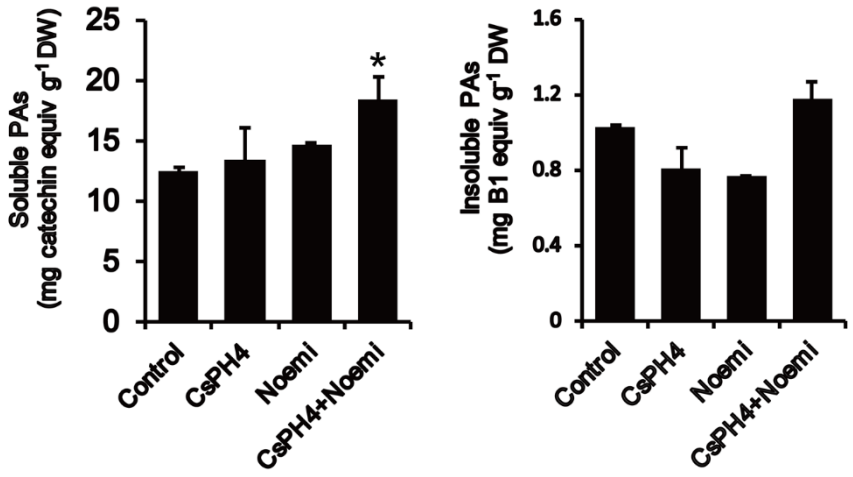

Figure S12. Characterization of PA levels of the transient overexpression of *CsPH4*, *Noemi*, *CsPH4* plus *Noemi* and an empty vector control in tobacco (*N. benthamiana*) leaves. Error bars represent the mean ± SD of three biological replicates. Asterisks indicate significant differences relative to the empty vector control (*, $P < 0.05$).

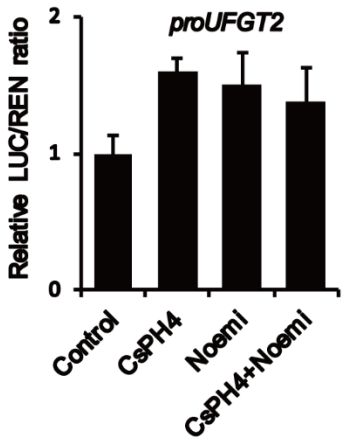

Figure S13. Transient promoter activity assays were carried out using *LUC* reporter gene under the control of the the promoters of *CsUFGT2*, along with effectors (*CsPH4* + *Noemi*) and the empty vector as an internal control. Error bars represent the mean ± SD of eight biological replicates.

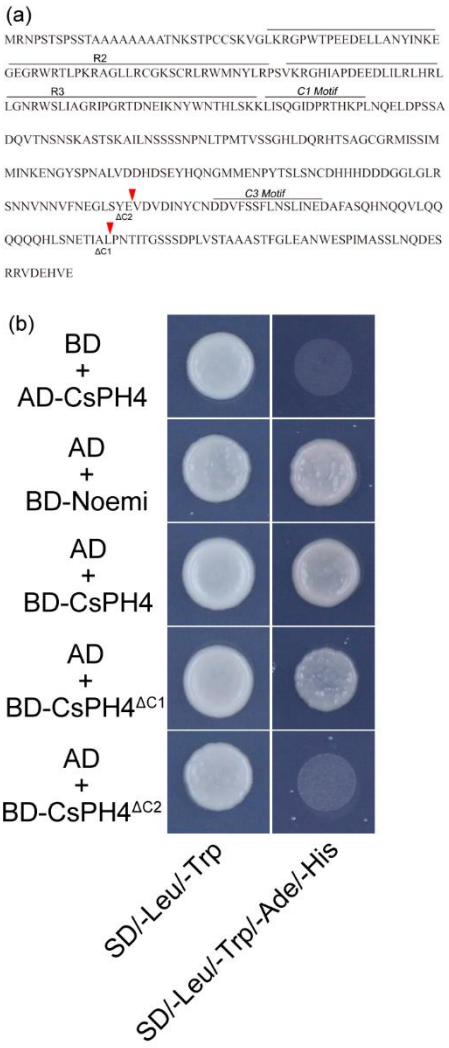

Figure S14. (a) Amino acid sequence analysis of *CsPH4*. The horizontal lines mark the conserved domains and motifs. The red triangle marks the protein truncation position. (b) The transcriptional self-activation activity assay of *CsPH4* and *Noemi* in yeast AH109. The full-length coding sequences of CsPH4, Noemi and truncated coding sequence of CsPH4$^{\Delta C1}$, CsPH4$^{\Delta C2}$ were cloned into PGADT7 (AD-CsPH4 and AD-Noemi) and PGBKT7 (BD-Noemi, BD-CsPH4, BD-CsPH4$^{\Delta C1}$, and BD-CsPH4$^{\Delta C2}$), respectively. The constructs BD-CsPH4 plus AD, BD-Noemi plus AD, BD-CsPH4$^{\Delta C1}$ plus AD, BD-CsPH4$^{\Delta C2}$ plus AD, AD-CsPH4 plus BD, and AD- Noemi plus BD were co-transformed into the yeast AH109. The transcriptional self-activation activity is

indicated by yeast growth. Yeast grown in SD/-Trp/-Leu medium and SD/-Trp/-Leu/-His/-Ade medium are indicated.