**Supplementary materials**

fig S1

**Fig. S1.** Analysis of fruit metabolomics. (A) Cluster results of 581 metabolites. (B-D) KEGG enrichment of differential abundant metabolites among the 3 tissues. The differential metabolites were identified with a threshold of fold change (FC) ≥ 2 or FC ≤ 0.5 and a variable importance in project (VIP) value ≥1. The square next to the volcano plot represent the number of metabolites enriched in each pathway, only top 10 differential metabolites which ordered by P value from top to bottom were shown.

fig S2

**Fig. S2.** Model of an integrated AAA and flavonoid biosynthesis pathways with coexpressed genes. Black arrows are known interactions, gene names are shown in purple, AAAs are shown in blue, and two classes of flavonoids are shown in green and orange.

图表, 图示

描述已自动生成

**Fig. S3.** Statistics of genes derived from whole genome duplication events involved in AAA biosynthesis. (A) Synteny analysis of the *A. trifoliata* genome. Red lines represent the collinearity blocks, and 8329 out of 24138 genes were in these synteny blocks. (B) Percentage of genes in AAA-related pathways derived from WGD events at different levels. Highly correlated c-expression genes were identified by WGCNA.

**fig S4**

**Fig. S4.** Heatmap of key gene expression levels and metabolite contents. The flavonoids were clustered into two class groups by TBtools.

图片包含 应用程序

描述已自动生成

**Fig. S5.** Functional identification of enzymes with prephenate dehydratase (PDT) activity in *A. trifoliata.* (A) Maximum likelihood tree of arogenate dehydratase (*ADT*) genes in *A. trifoliata*, *P. hybrida*, *P. pinaster* and *A. thaliana*. (B) Multiple sequence alignment of the PDT activity conferring domain.