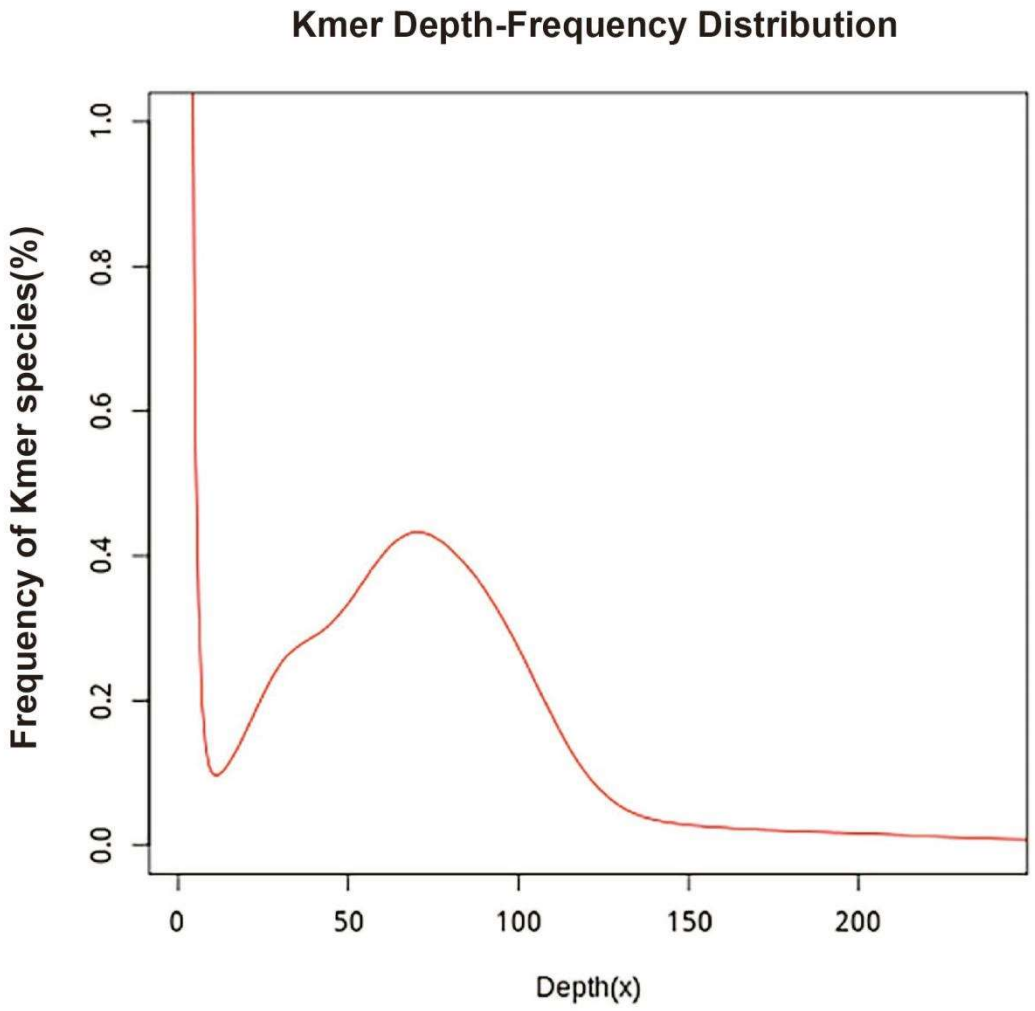

Figure S1. Genomic survey of *P. vulgaris*.
The horizontal axis represents the depth at Kmer=19, and the vertical axis represents the frequency at different depths.

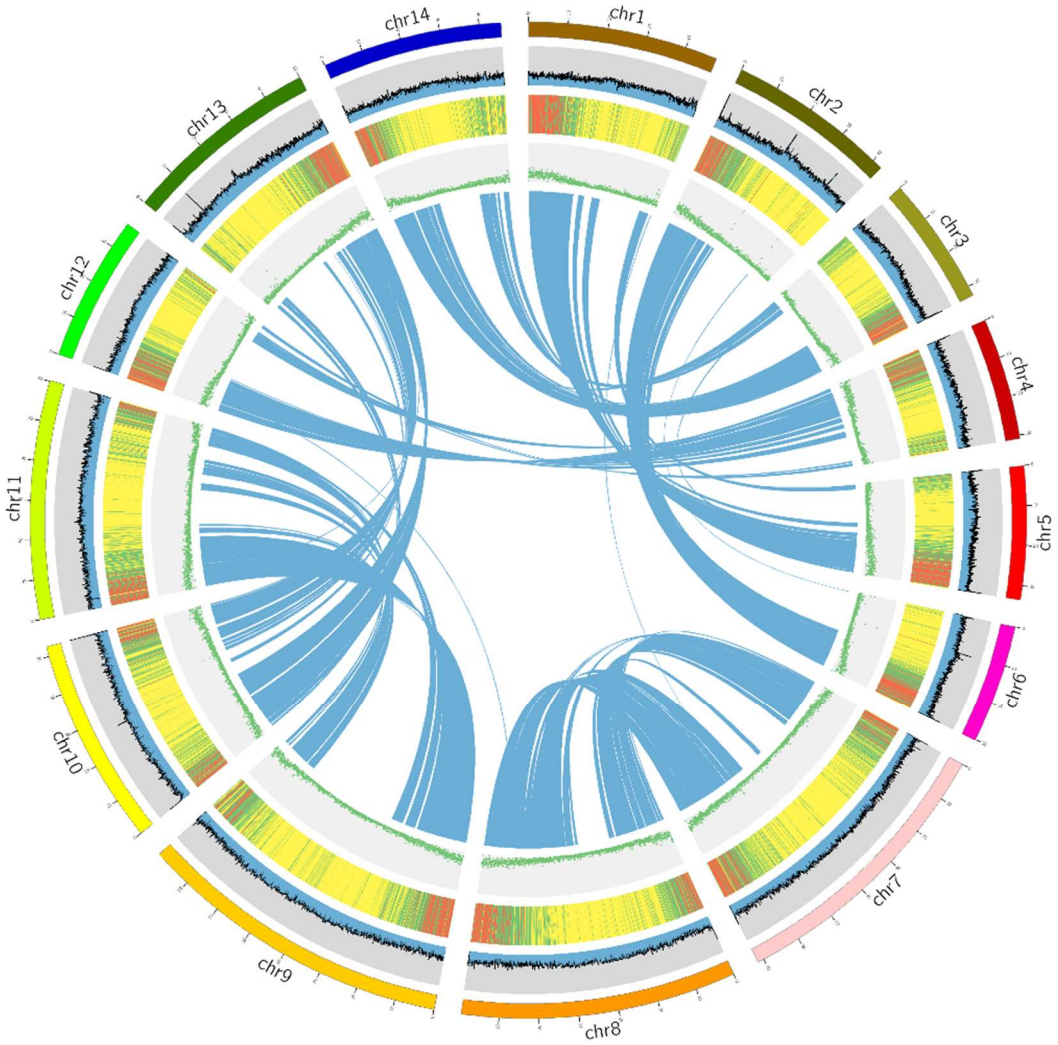

Figure S2. Circos diagram of the *P. vulgaris* genome.
Outside to inside: Pseudochromosome length, gene density, repeat density, GC content, and intragenomic collinearity.

Figure S3. Cluster diagram of overall metabolites.

Horizontal axis represents sample names, vertical axis represents metabolite information, and "Group" denotes the grouping. Different colors are used to fill cells based on different values obtained after relative content standardization (red indicates high content, green indicates low content). MS represents stem of *P. vulgaris*, MZ represents seed of *P. vulgaris*, MG represents ear of *P. vulgaris*, MR represents root of *P. vulgaris*, ML represents leaf of *P. vulgaris*.

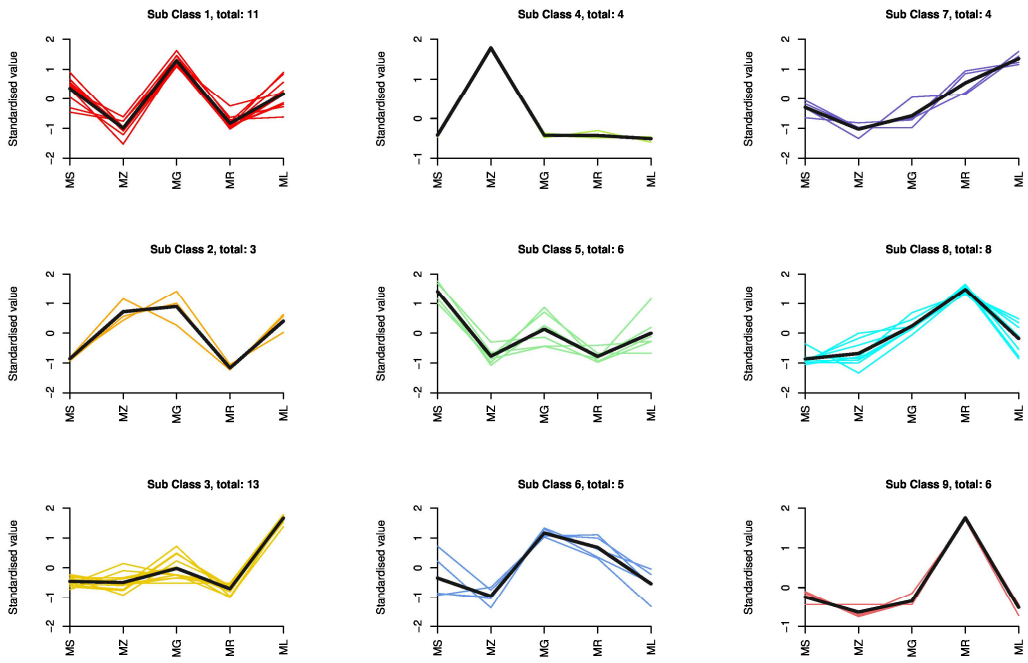

Figure S4. K-Means diagram of differential metabolites.

The horizontal axis represents sample grouping, while the vertical axis represents standardized relative metabolite content. "Sub Class" indicates the category number of metabolites with the same changing trend, and "total" represents the total number of metabolites in that category. MS represents stem of *P. vulgaris*, MZ represents seed of *P. vulgaris*, MG represents ear of *P. vulgaris*, MR represents root of *P. vulgaris*, ML represents leaf of *P. vulgaris*.

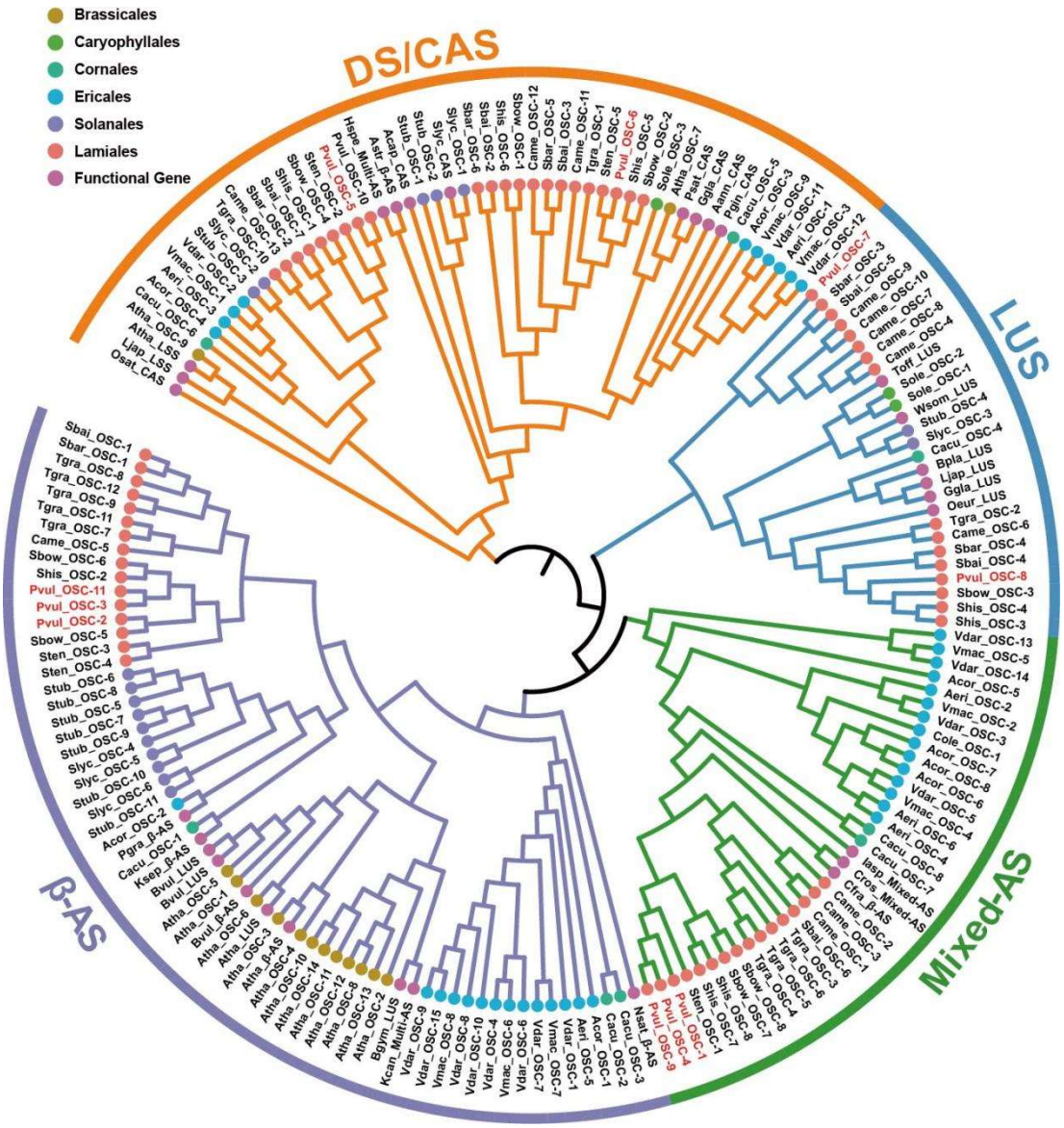

Figure S5. Phylogeny of OSC genes extracted from 18 plant genomes and functionally characterized enzyme genes.

Phylogenetic tree constructed from all OSC genes in 18 species and characterized enzymes, with the DS/CAS subfamily being the outgroup. Different colored circles preceding each gene represent the family to which the gene's species belongs. Each subfamily branch is distinguished by unique colors and shaded bands. *P. vulgaris* β-AS and the mixed-AS candidates are labeled in red. The IDs for the candidate genes are provided in Supplementary Table S1.

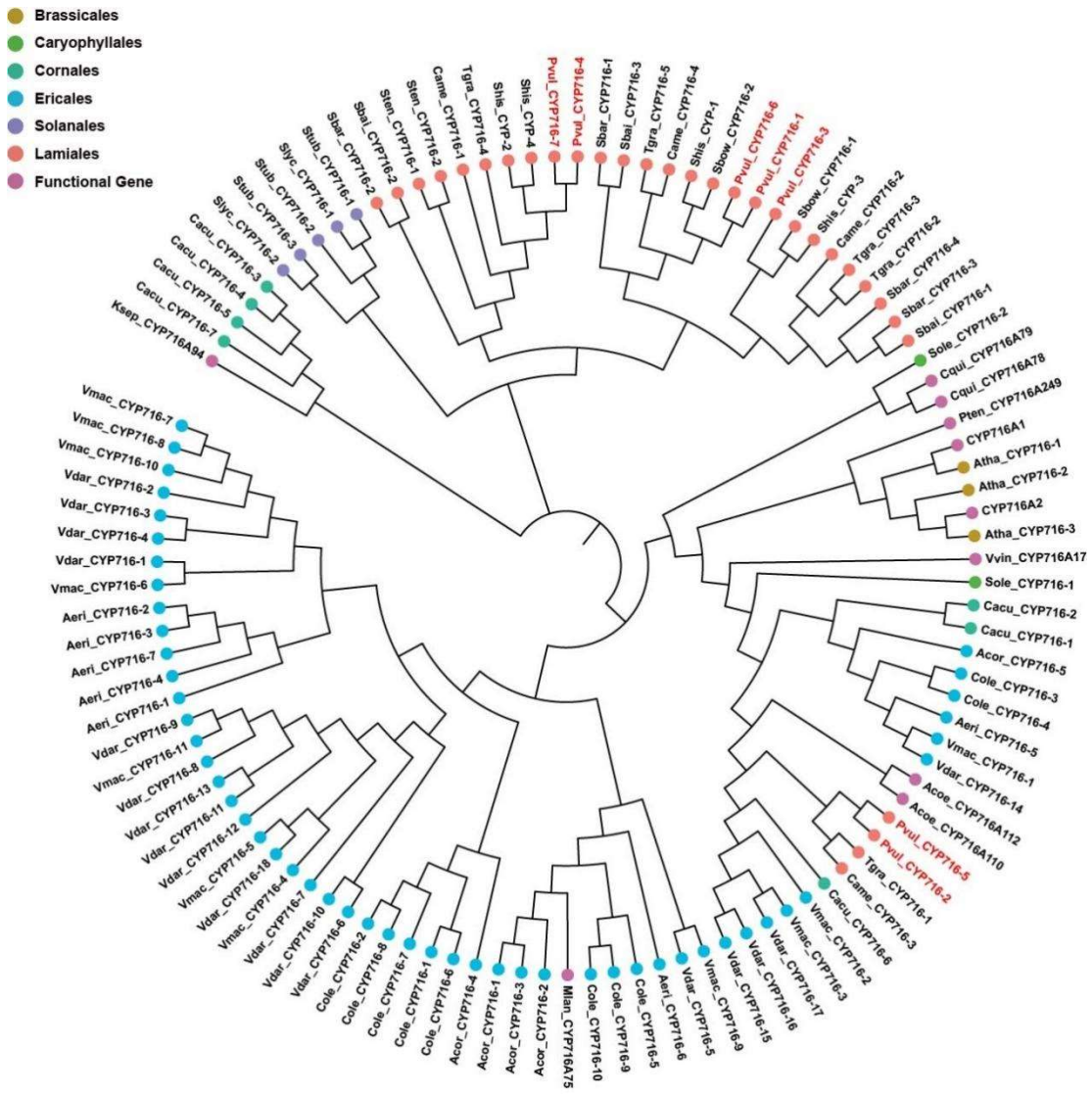

Figure S6. Detailed Phylogeny of CYP716 genes.

Detailed phylogenetic tree of CYP716 genes, corresponding to the red branches in CYP85 from Figure 2E. Different colored circles preceding each gene represent the botanical family to which the gene's species belongs. *P. vulgaris* CYP716s are labeled in red.

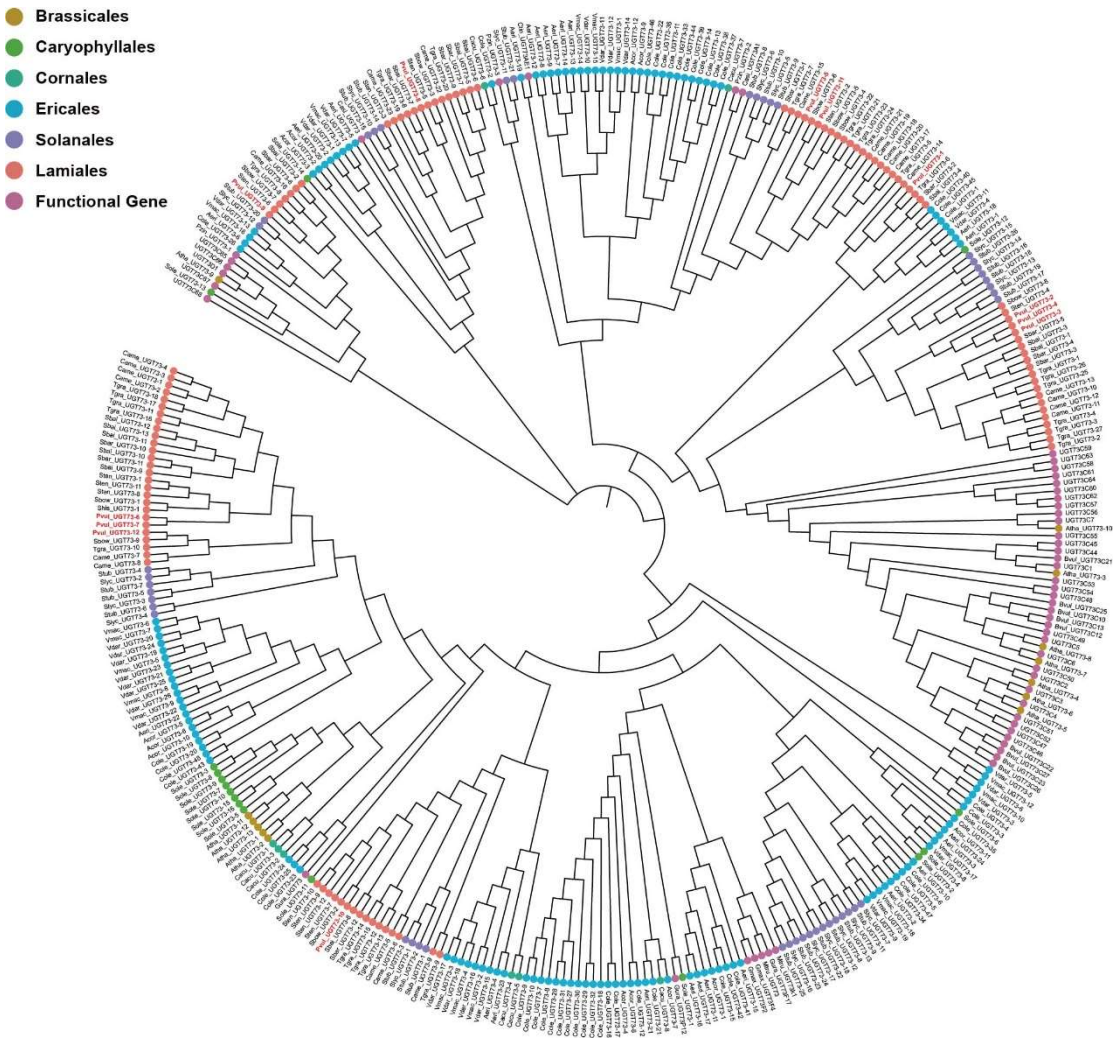

Figure S7. Detailed phylogeny of UGT73 genes.

Detailed phylogenetic tree of UGT73 genes, corresponding to the red branches within Cluster B in Figure 2F. Different colored circles preceding each gene signify the botanical family to which the gene's species belongs. *P. vulgaris* UGT73s are labeled in red.

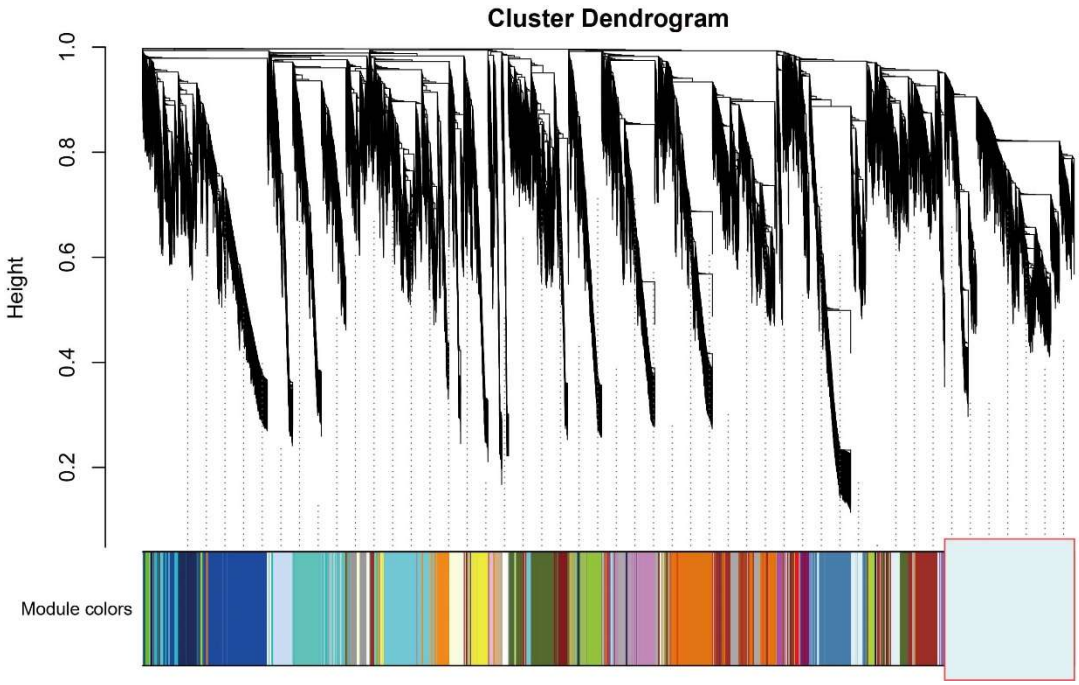

Figure S8. Gene co-expression module classification by WGCNA using trasncriptomic data of roots, stems, leaves, seeds and ears (three replicates for each organ) of *P. vulgaris*.

Each branch of the tree represents a gene, and genes with similar expression patterns (1-cor=0.25) cluster together to form modules of various colors. The module of interest is the turquoise module, which is zoomed in and highlighted with a red border.

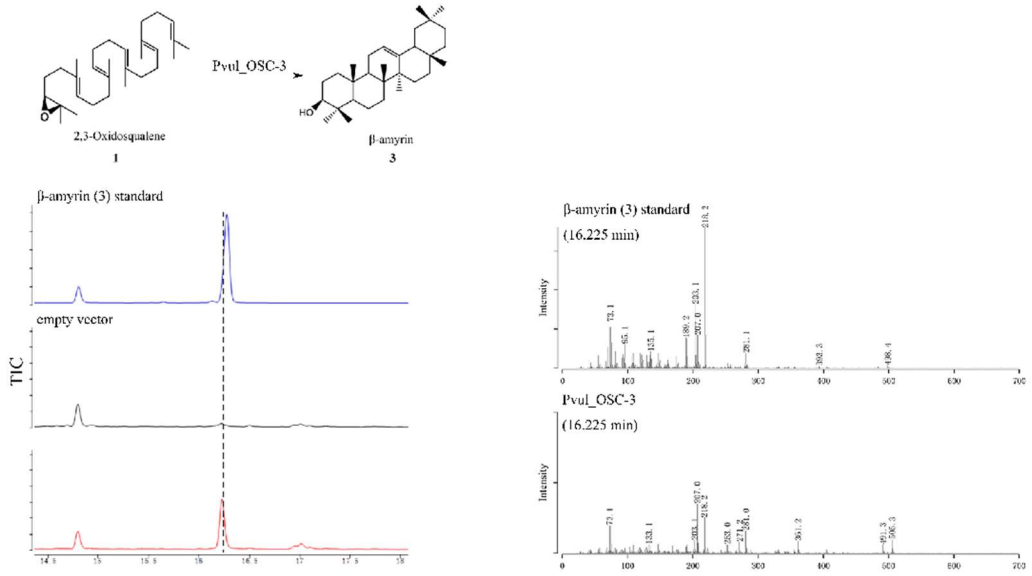

Figure S9. Functional characterization of Pvul_OSC-3.

Pvul_OSC-3 was transiently expressed in *N. benthamiana* leaves with substrate. GC-MS analysis of leaf extracts of *N. benthamiana* following Agrobacterium-mediated transient expression. Leaves were agro-infiltrated with control vector or Pvul_OSC-3. Total ion chromatograms (TIC) are shown on the left, and mass spectra on the right. The retention time and mass spectrum for the Pvul_OSC-3 product are identical to that of a β-amyrin (3) standard.

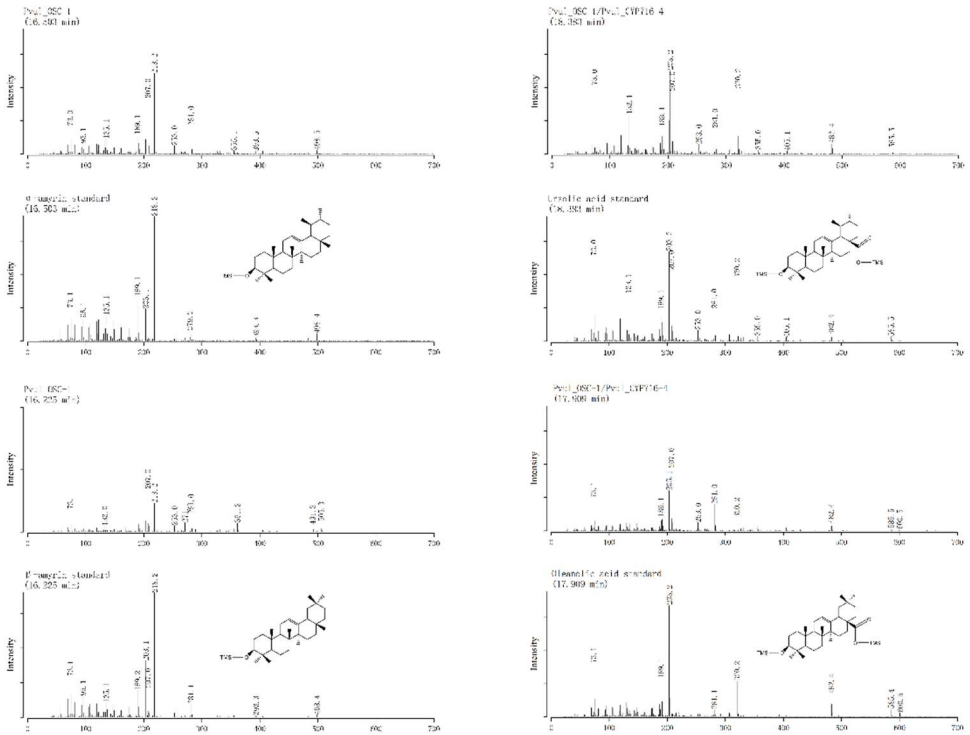

Figure S10. Functional characterization of Pvul_OSC_1 and Pvul_CYP716s.
Pvul_CYP716s and/or Pvul_OSC-1 were transiently expressed in *N. benthamiana* leaves with substrate. GC-MS analysis of leaf extracts of *Nicotiana benthamiana* following Agrobacterium-mediated transient expression. Leaves were agro-infiltrated with expression constructs for vector control, Pvul_OSC-1 or Pvul_OSC-1/Pvul_CYP716-4. Mass spectra of the products are shown and compared to each authentic standard.

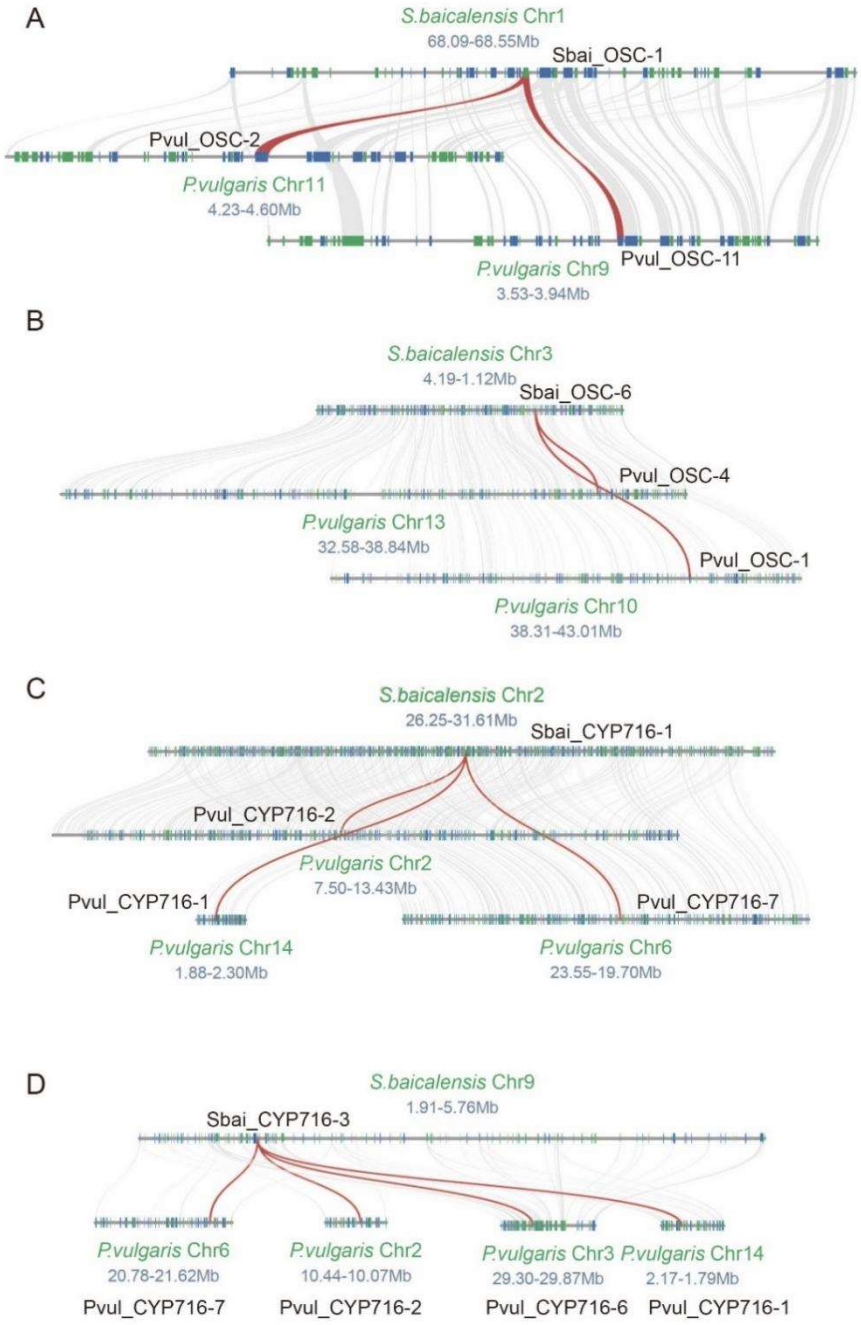

Figure S11. Synteny of characterized OSCs and CYP716s in *P. vulgaris* with their orthologs in *S. baicalensis*.

(A) Synteny between Sbai_OSC-1 of *S. baicalensis* and Pvul_OSC-2, Pvul_OSC-11 of *P. vulgaris*;

(B) Synteny between Sbai_OSC-6 of *S. baicalensis* and Pvul_OSC-1, Pvul_OSC-4 of *P. vulgaris*;

(C) Synteny between Sbai_CYP716-1 of *S. baicalensis* and Pvul_CYP716-1, Pvul_CYP716-2, Pvul_CYP716-7 of *P. vulgaris*;

(D) Synteny between Sbai_CYP716-3 of *S. baicalensis* and Pvul_CYP716-1, Pvul_CYP716-2, Pvul_CYP716-6, Pvul_CYP716-7 of *P. vulgaris*.

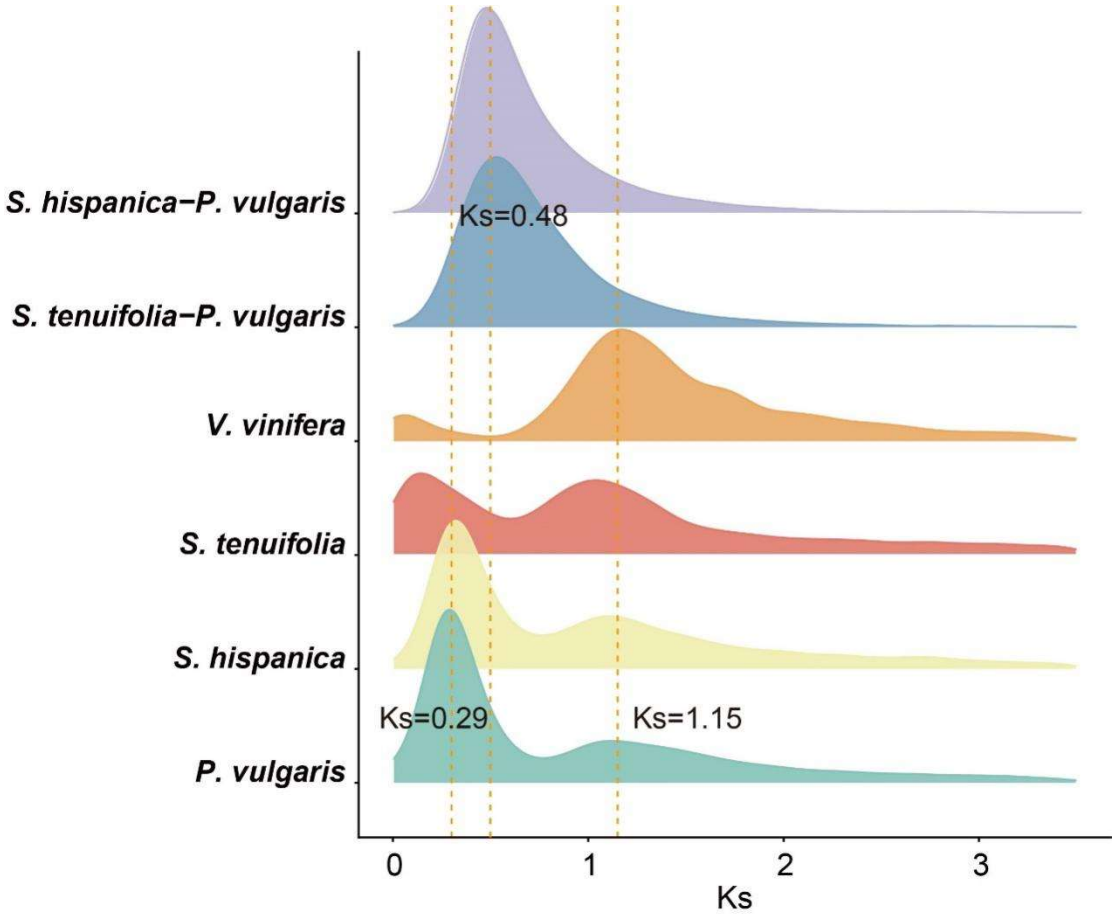

Figure S12. Whole-genome duplication (WGD) events analysis in *P. vulgaris* and three other plants.

WGD analysis of *P. vulgaris*, *Salvia hispanica*, *Schizonepeta tenuifolia* and *Vitis vinifera*. The peak corresponding to the Ks value 0.29 indicates an independent WGD event in *P. vulgaris*, and peak corresponding to the Ks value 1.15 indicates the eudicots-shared whole genome triplication event (γ-WGT).