De novo ASSEMBLY AND ANALYSIS OF SUMAC (*Toxicodendron vernicifluum* (STOKES) F. A. BARKLEY) TRANSCRIPTOMES PROVIDES INSIGHTS INTO THE BIOSYNTHESIS OF URUSHIOL

G.-Q. Bai, Y. Jia, W.-M. Li, H. Chen, B. Li, S.-F. Li

Supplemental materials

Table S1 Number of unigenes expressed involved in KEGG pathway

Table S2 Candidate genes encoding type III polyketide synthases in Urushiol biosynthesis

Fig. S1 Length distributions of *Toxicodendron vernicifluum* unigenes.

Fig S2 The results of unigenes Compared with other species

Fig. S3 Clusters of orthologous groups (COG) classification. In total, 22,032 sequences were grouped into 24 COG classifications.

Fig S4 Distribution of the number of differential gene expression between three varieties (*T. vernicifluum* (A), Dahongpao(B) and Huoyanzi(C)).

Fig S5 GO enrichment analysis for DEGs between *T. vernicifluum* (A) and Dahongpao (B).

Fig S6 GO enrichment analysis for DEGs between *T. vernicifluum* (A) and Huoyanzi (C).

Fig S7 GO enrichment analysis for DEGs between Dahongpao (B) and Huoyanzi (C).

Fig. S8 A type III polyketide synthase utilizing malonyl-CoA and a long chain fatty acid-CoA substrate has been put forward as the first step in urushiol biosynthetic pathways (Weisberg 2014). Urushiol is comprised of a catechol ring with a 15 or 17-carbon alkyl side chain. *Toxicodendron vernicifluum* contain a mixture of these urushiols with varying amounts, represented by the 15-carbon forms in this pathway.

Tbble S1 Number of unigenes expressed involved in KEGG pathway

	Pathway	**All genes with pathway annotation (16353)**	**Pathway ID**
1	Metabolic pathways	5189 (31.73%)	ko01100
2	Biosynthesis of secondary metabolites	2706 (16.55%)	ko01110
3	Ribosome	1547 (9.46%)	ko03010
4	Glycolysis / Gluconeogenesis	659 (4.03%)	ko00010
5	Oxidative phosphorylation	623 (3.81%)	ko00190
6	Protein processing in endoplasmic reticulum	618 (3.78%)	ko04141
7	Spliceosome	502 (3.07%)	ko03040
8	Pyruvate metabolism	467 (2.86%)	ko00620
9	RNA transport	444 (2.72%)	ko03013
10	Purine metabolism	432 (2.64%)	ko00230
11	Glyoxylate and dicarboxylate metabolism	408 (2.49%)	ko00630
12	Peroxisome	363 (2.22%)	ko04146
13	Citrate cycle (TCA cycle)	351 (2.15%)	ko00020
14	Carbon fixation in photosynthetic organisms	350 (2.14%)	ko00710
15	Fatty acid metabolism	336 (2.05%)	ko00071
16	Arginine and proline metabolism	310 (1.9%)	ko00330
17	Propanoate metabolism	305 (1.87%)	ko00640
18	Starch and sucrose metabolism	304 (1.86%)	ko00500
19	Phagosome	303 (1.85%)	ko04145
20	Valine, leucine and isoleucine degradation	297 (1.82%)	ko00280
21	Pyrimidine metabolism	295 (1.8%)	ko00240
22	Plant-pathogen interaction	284 (1.74%)	ko04626
23	Phenylpropanoid biosynthesis	278 (1.7%)	ko00940
24	RNA degradation	273 (1.67%)	ko03018
25	Ribosome biogenesis in eukaryotes	273 (1.67%)	ko03008
26	Tryptophan metabolism	267 (1.63%)	ko00380
27	Pentose phosphate pathway	267 (1.63%)	ko00030
28	Glutathione metabolism	265 (1.62%)	ko00480

29	Amino sugar and nucleotide sugar metabolism	264 (1.61%)	ko00520
30	Endocytosis	260 (1.59%)	ko04144
31	Alanine, aspartate and glutamate metabolism	258 (1.58%)	ko00250
32	Ubiquitin mediated proteolysis	258 (1.58%)	ko04120
33	Biosynthesis of unsaturated fatty acids	232 (1.42%)	ko01040
34	mRNA surveillance pathway	230 (1.41%)	ko03015
35	beta-Alanine metabolism	221 (1.35%)	ko00410
36	Tyrosine metabolism	221 (1.35%)	ko00350
37	Fructose and mannose metabolism	220 (1.35%)	ko00051
38	Glycerolipid metabolism	215 (1.31%)	ko00561
39	Pentose and glucuronate interconversions	213 (1.3%)	ko00040
40	Cysteine and methionine metabolism	207 (1.27%)	ko00270
41	Ascorbate and aldarate metabolism	206 (1.26%)	ko00053
42	Proteasome	200 (1.22%)	ko03050
43	Glycine, serine and threonine metabolism	197 (1.2%)	ko00260
44	Plant hormone signal transduction	188 (1.15%)	ko04075
45	Aminoacyl-tRNA biosynthesis	186 (1.14%)	ko00970
46	Phenylalanine metabolism	168 (1.03%)	ko00360
47	Nitrogen metabolism	168 (1.03%)	ko00910
48	Lysine degradation	165 (1.01%)	ko00310
49	Galactose metabolism	165 (1.01%)	ko00052
50	Glycerophospholipid metabolism	159 (0.97%)	ko00564
51	Photosynthesis	158 (0.97%)	ko00195
52	Histidine metabolism	149 (0.91%)	ko00340
53	Limonene and pinene degradation	149 (0.91%)	ko00903
54	Fatty acid biosynthesis	149 (0.91%)	ko00061
55	Butanoate metabolism	143 (0.87%)	ko00650
56	Nucleotide excision repair	134 (0.82%)	ko03420
57	Protein export	129 (0.79%)	ko03060
58	DNA replication	124 (0.76%)	ko03030
59	Inositol phosphate metabolism	121 (0.74%)	ko00562
60	alpha-Linolenic acid metabolism	115 (0.7%)	ko00592
61	N-Glycan biosynthesis	114 (0.7%)	ko00510
62	Valine, leucine and isoleucine biosynthesis	114 (0.7%)	ko00290
63	RNA polymerase	107 (0.65%)	ko03020
64	Basal transcription factors	105 (0.64%)	ko03022
65	Cyanoamino acid metabolism	100 (0.61%)	ko00460

66	Phenylalanine, tyrosine and tryptophan biosynthesis	99 (0.61%)	ko00400
67	Tropane, piperidine and pyridine alkaloid biosynthesis	98 (0.6%)	ko00960
68	Ubiquinone and other terpenoid-quinone biosynthesis	91 (0.56%)	ko00130
69	Phosphatidylinositol signaling system	91 (0.56%)	ko04070
70	Terpenoid backbone biosynthesis	90 (0.55%)	ko00900
71	Photosynthesis - antenna proteins	89 (0.54%)	ko00196
72	Porphyrin and chlorophyll metabolism	89 (0.54%)	ko00860
73	Sulfur metabolism	86 (0.53%)	ko00920
74	Pantothenate and CoA biosynthesis	86 (0.53%)	ko00770
75	Base excision repair	78 (0.48%)	ko03410
76	Natural killer cell mediated cytotoxicity	76 (0.46%)	ko04650
77	One carbon pool by folate	69 (0.42%)	ko00670
78	Sphingolipid metabolism	68 (0.42%)	ko00600
79	Homologous recombination	67 (0.41%)	ko03440
80	Selenocompound metabolism	66 (0.4%)	ko00450
81	SNARE interactions in vesicular transport	64 (0.39%)	ko04130
82	Mismatch repair	64 (0.39%)	ko03430
83	Steroid biosynthesis	55 (0.34%)	ko00100
84	Regulation of autophagy	54 (0.33%)	ko04140
85	Flavonoid biosynthesis	53 (0.32%)	ko00941
86	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	52 (0.32%)	ko00563
87	Ether lipid metabolism	51 (0.31%)	ko00565
88	Isoquinoline alkaloid biosynthesis	51 (0.31%)	ko00950
89	Carotenoid biosynthesis	46 (0.28%)	ko00906
90	Circadian rhythm - plant	45 (0.28%)	ko04712
91	Other glycan degradation	43 (0.26%)	ko00511
92	Arachidonic acid metabolism	38 (0.23%)	ko00590
93	Stilbenoid, diarylheptanoid and gingerol biosynthesis	34 (0.21%)	ko00945
94	Folate biosynthesis	34 (0.21%)	ko00790
95	Riboflavin metabolism	30 (0.18%)	ko00740
96	Nicotinate and nicotinamide metabolism	30 (0.18%)	ko00760
97	Lysine biosynthesis	30 (0.18%)	ko00300
98	Taurine and hypotaurine metabolism	29 (0.18%)	ko00430
99	Vitamin B6 metabolism	28 (0.17%)	ko00750
100	Synthesis and degradation of ketone bodies	27 (0.17%)	ko00072

101	Glycosphingolipid biosynthesis - globo series	27 (0.17%)	ko00603
102	Sulfur relay system	27 (0.17%)	ko04122
103	C5-Branched dibasic acid metabolism	26 (0.16%)	ko00660
104	Fatty acid elongation in mitochondria	26 (0.16%)	ko00062
105	Flavone and flavonol biosynthesis	26 (0.16%)	ko00944
106	Zeatin biosynthesis	25 (0.15%)	ko00908
107	Circadian rhythm - mammal	25 (0.15%)	ko04710
108	Thiamine metabolism	23 (0.14%)	ko00730
109	Non-homologous end-joining	23 (0.14%)	ko03450
110	Glucosinolate biosynthesis	21 (0.13%)	ko00966
111	ABC transporters	21 (0.13%)	ko02010
112	Diterpenoid biosynthesis	20 (0.12%)	ko00904
113	Linoleic acid metabolism	20 (0.12%)	ko00591
114	Sesquiterpenoid biosynthesis	18 (0.11%)	ko00909
115	Glycosaminoglycan degradation	16 (0.1%)	ko00531
116	Glycosphingolipid biosynthesis - ganglio series	13 (0.08%)	ko00604
117	Lipoic acid metabolism	12 (0.07%)	ko00785
118	Biotin metabolism	10 (0.06%)	ko00780
119	Caffeine metabolism	10 (0.06%)	ko00232
120	Brassinosteroid biosynthesis	9 (0.06%)	ko00905
121	Monoterpenoid biosynthesis	7 (0.04%)	ko00902
122	Betalain biosynthesis	5 (0.03%)	ko00965
123	Other types of O-glycan biosynthesis	5 (0.03%)	ko00514
124	Anthocyanin biosynthesis	3 (0.02%)	ko00942
125	D-Arginine and D-ornithine metabolism	1 (0.01%)	

Table S2 Candidate genes encoding type III polyketide synthases in Urushiol biosynthesis

GeneID	Length	RPKM			Description
		T. vernicifluum	Dahongpao	Huoyanzi	
Unigene0035593	1963	695.72	300.14	508.69	chalcone synthase [Dictamnus albus]
Unigene0050227	878	1.5	1.82	1.83	chalcone synthase A [Vitis vinifera]
Unigene0050228	873	1.84	2.02	1.01	chalcone synthase 2-like [Solanum lycopersicum]
Unigene0051769	668	1.81	5.85	1.08	chalcone synthase [Populus trichocarpa]
Unigene0051770	250	64.66	3.46	67.54	chalcone synthase [Pueraria candollei var. mirifica]
Unigene0051771	823	381.97	32.61	258.38	chalcone synthase [Rhus chinensis]
Unigene0051772	289	17.07	1.11	125.24	chalcone synthase 2 [Astragalus membranaceus]
Unigene0051773	618	148.57	38.09	124.838	chalcone synthase [Rhus chinensis]
Unigene0051774	250	220.35	32.28	21.38	chalcone synthase 3 [Vitis vinifera]

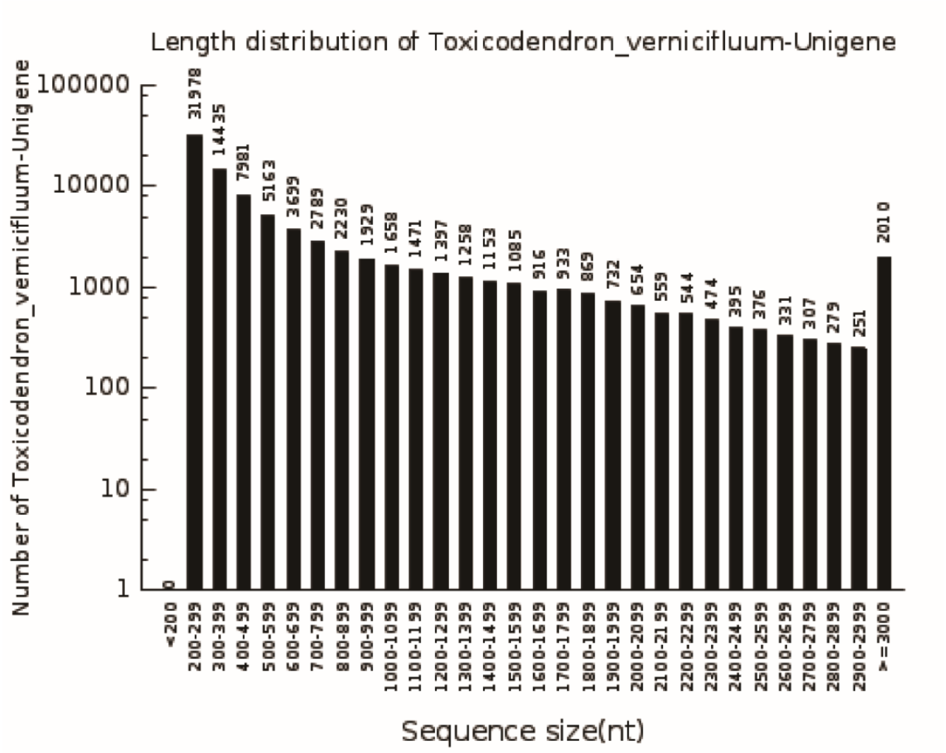

Fig. S1 Length distributions of *Toxicodendron vernicifluum* unigenes.

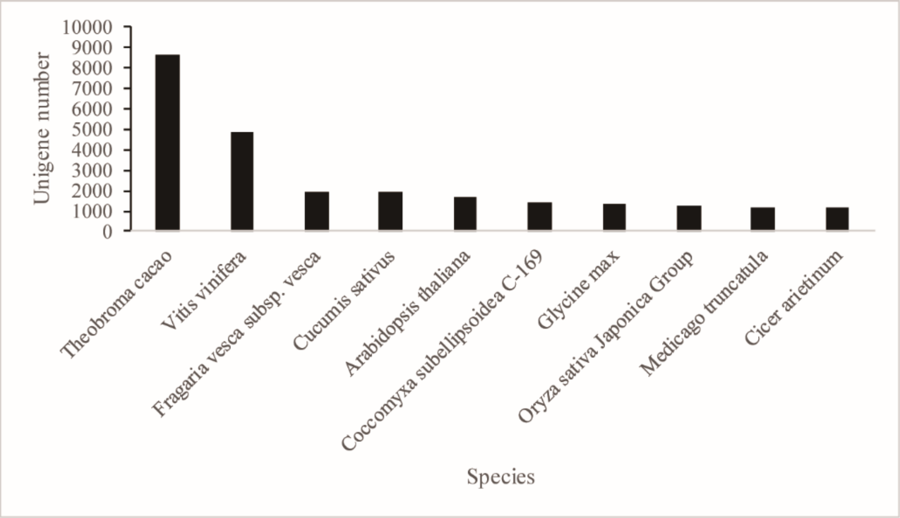

Fig S2 The results of unigenes compared with other species

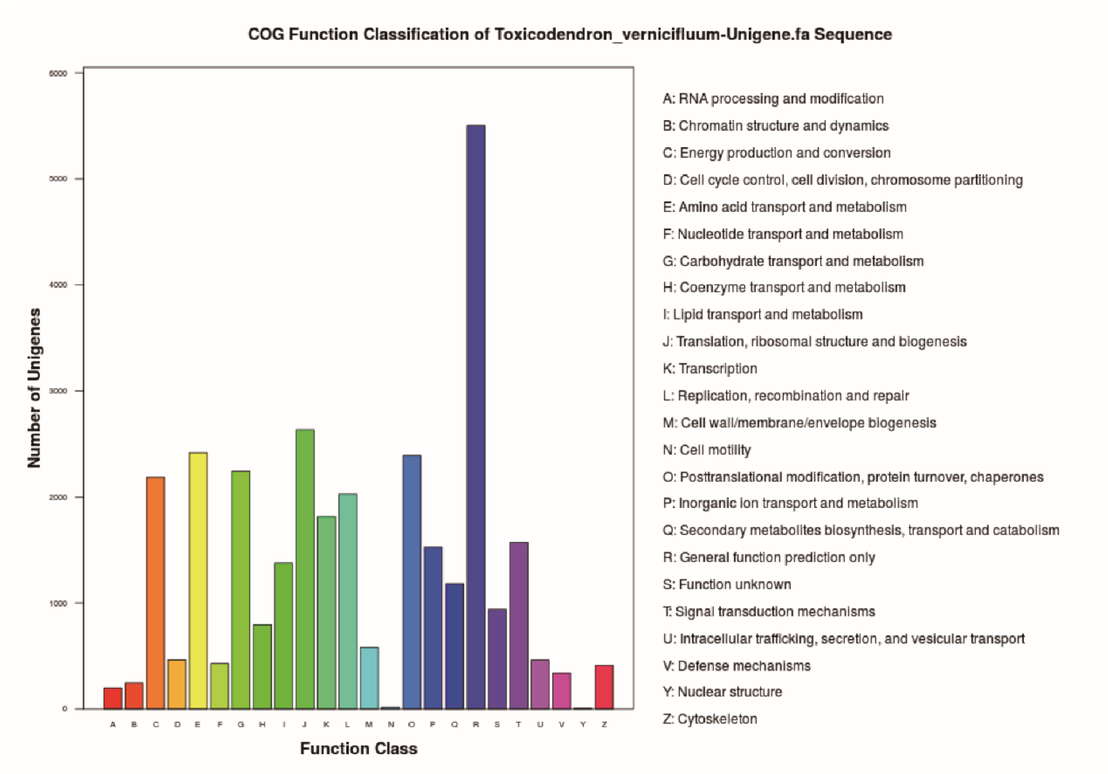

Fig. S3 Clusters of orthologous groups (COG) classification. In total, 22,032 sequences were grouped into 24 COG classifications.

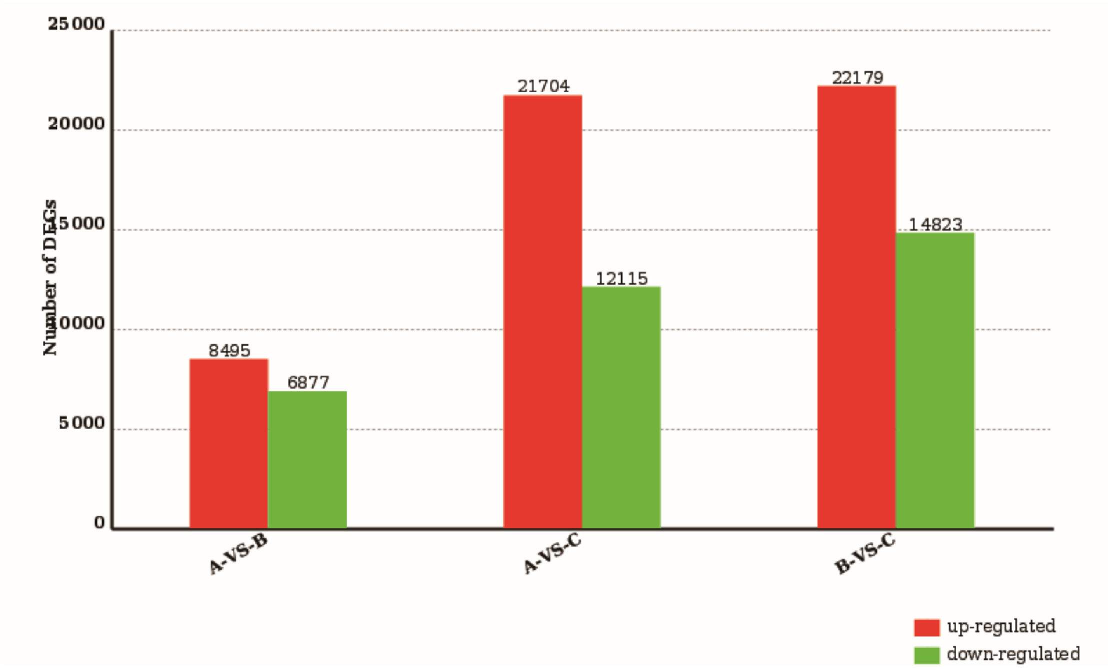

Fig S4 Distribution of the number of differential gene expression between three varieties (*T. vernicifluum* (A), Dahongpao(B) and Huoyanzi(C)).

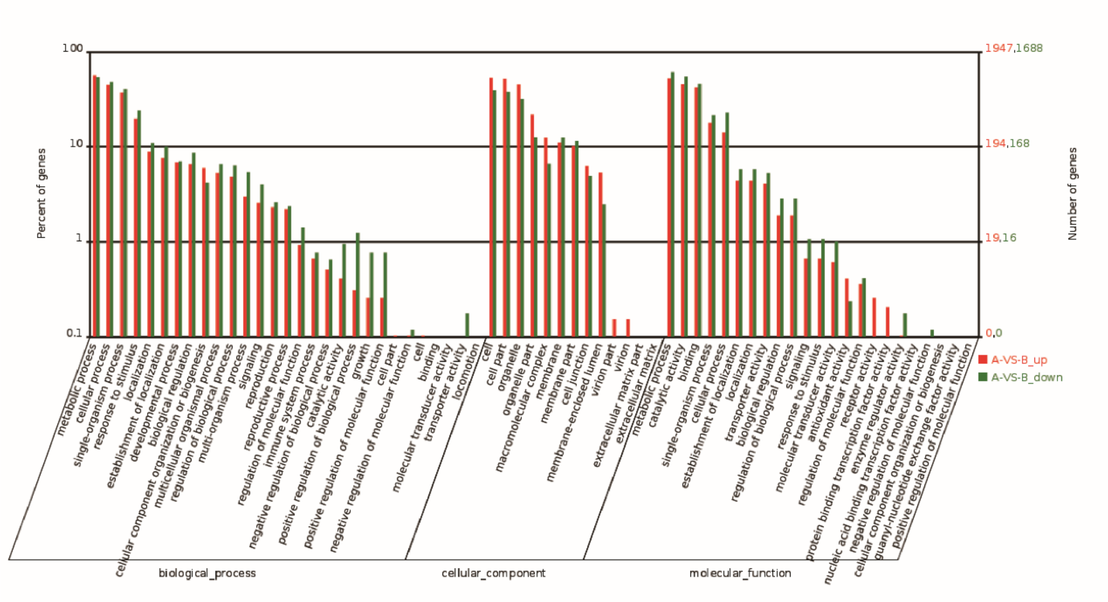

Fig S5GO enrichment analysis for DEGs between *T. vernicifluum* (A) and Dahongpao (B).

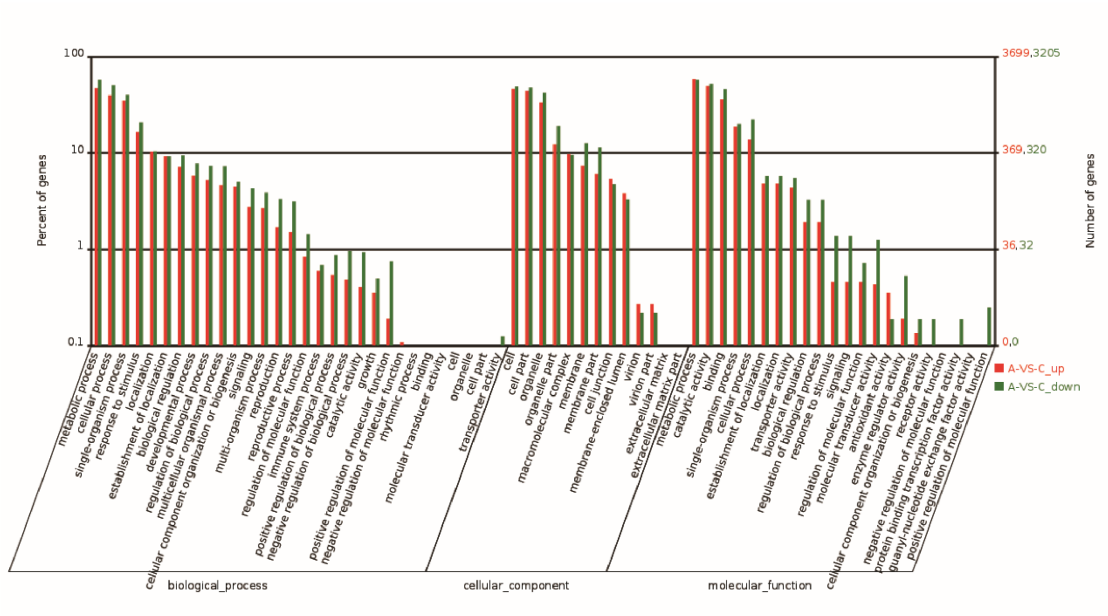

Fig S6 GO enrichment analysis for DEGs between *T. vernicifluum* (A) and Huoyanzi (C).

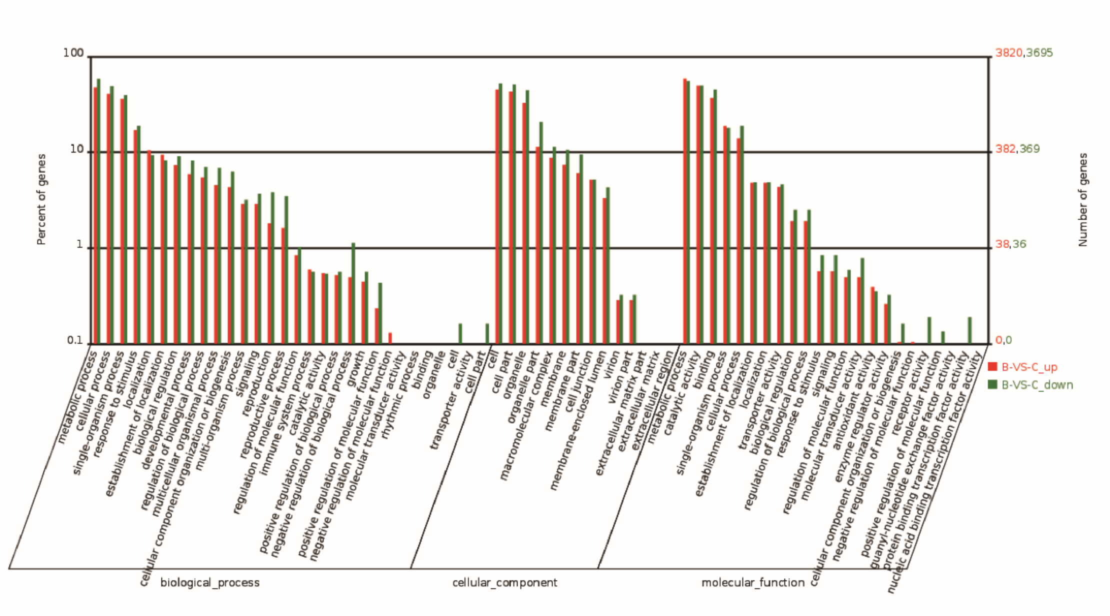

Fig S7 GO enrichment analysis for DEGs between Dahongpao (B) and Huoyanzi (C).

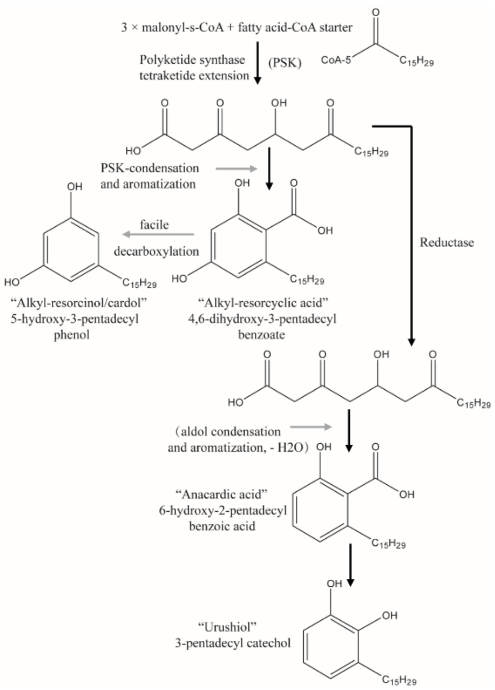

Fig. S8 A type III polyketide synthase utilizing malonyl-CoA and a long chain fatty acid-CoA substrate has been put forward as the first step in urushiol biosynthetic pathways (Weisberg 2014). Urushiol is comprised of a catechol ring with a 15 or 17-carbon alkyl side chain. *Toxicodendron vernicifluum* contain a mixture of these urushiols with varying amounts, represented by the 15-carbon forms in this pathway.