

***De novo* transcriptome assembly of *Zanthoxylum bungeanum* using Illumina sequencing for evolutionary analysis and simple sequence repeat marker development**

Shijing Feng¹, Lili Zhao¹, Zhenshan Liu², Yulin Liu¹, Tuxi Yang¹ & Anzhi Wei¹

¹College of Forestry, Northwest A&F University, Yangling, Shaanxi 712100, China. ²College of Life Science, Northwest A&F University, Yangling, Shaanxi 712100, China. Correspondence and requests for materials should be addressed to A.W. (email: weianzhi@126.com). [†]Co-first author: this author contributed equally to this paper

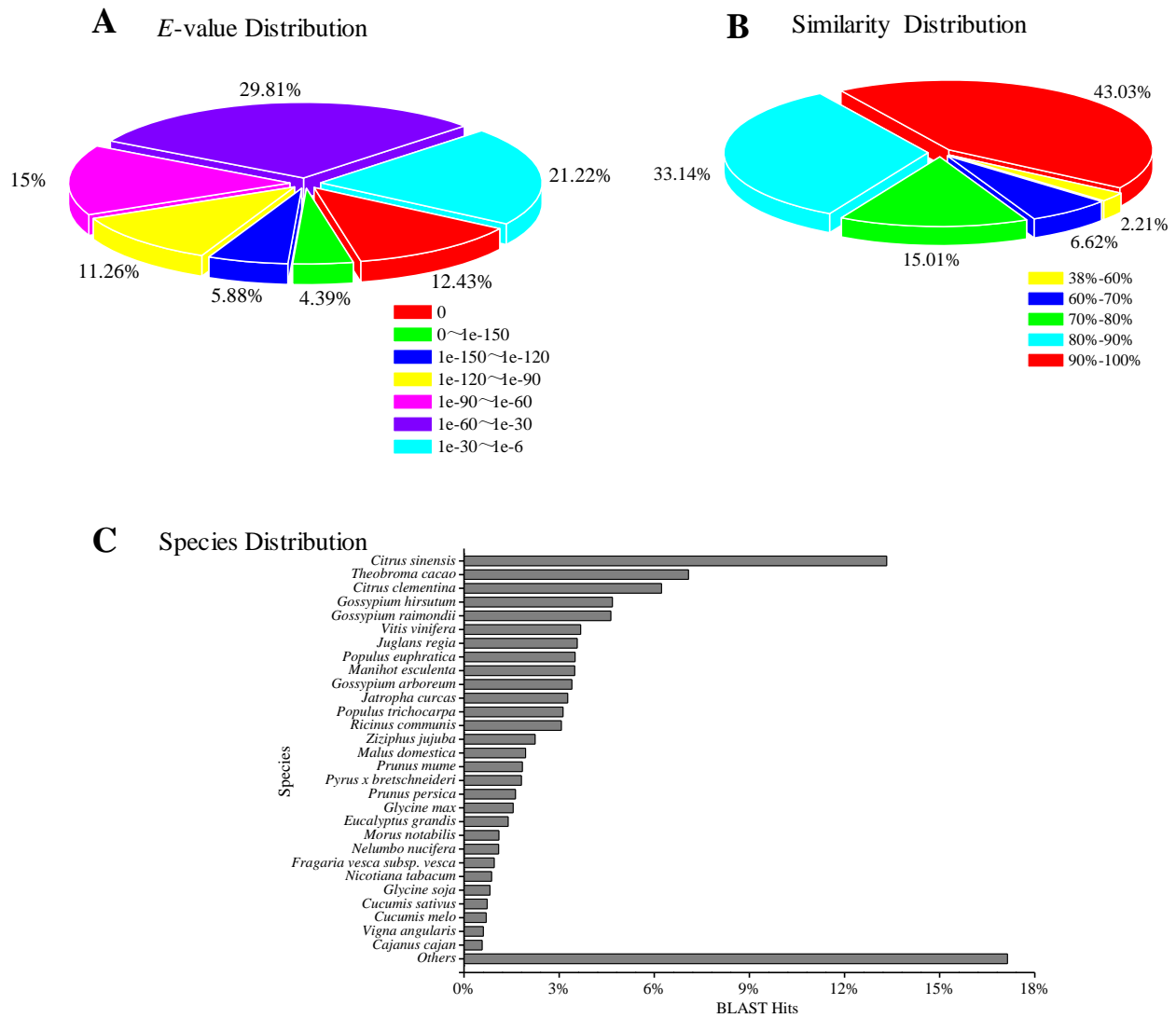
Supplement Table S1. The 31 cultivars of *Zanthoxylum bungeanum* and *Z. armatum* used in the present study.

Code	Cultivar	Species	Location
1	Fengxianqiyuejiao	<i>Z. bungeanum</i>	
2	Fengxiandahongpao	<i>Z. bungeanum</i>	Fengxian, Shaanxi
3	Fengxianzhuyejiao	<i>Z. armatum</i>	
4	Yehujiao	<i>Z. bungeanum</i>	
5	Hanchengdahongpao	<i>Z. bungeanum</i>	
6	Shizitou	<i>Z. bungeanum</i>	Hancheng, Shaanxi
7	Choujiao	<i>Z. bungeanum</i>	
8	Fuguhujiao	<i>Z. bungeanum</i>	Fugu, Shaanxi
9	Xunyanghuajiao	<i>Z. bungeanum</i>	
10	Goujiao	<i>Z. armatum</i>	Xunyang, Shaanxi
11	Wududahongpao	<i>Z. bungeanum</i>	
12	Wudubayuejiao	<i>Z. bungeanum</i>	Wudu, Gansu
13	Baijiao	<i>Z. bungeanum</i>	
14	Youjiao	<i>Z. bungeanum</i>	Qin'an, Gansu
15	Qin'anyihao	<i>Z. bungeanum</i>	
16	Hanyuanhuajiao	<i>Z. bungeanum</i>	
17	Maowenhuajiao	<i>Z. bungeanum</i>	Hanyuan, Sichuan
18	Hanyuanqinghuajiao	<i>Z. armatum</i>	
19	Zizhongzhuyejiao	<i>Z. armatum</i>	
20	Tengjiao	<i>Z. armatum</i>	Zizhong, Sichuan
21	Guizhouhuajiao	<i>Z. bungeanum</i>	
22	Guizhouzhuyejiao	<i>Z. armatum</i>	Liupanshui, Guizhou
23	Dingtanhujiao	<i>Z. armatum</i>	Guanling, Guizhou
24	Yunnanhuajiao	<i>Z. bungeanum</i>	
25	Jiuyeqinghuajiao	<i>Z. armatum</i>	Kunming, Yunnan
26	Yunnanzhuyejiao	<i>Z. armatum</i>	
27	Shandongdahongpao	<i>Z. bungeanum</i>	
28	Shandongxiaohongpao	<i>Z. bungeanum</i>	Laiwu, Shandong
29	Huangjinjiao	<i>Z. bungeanum</i>	
30	Shanxidahongpao	<i>Z. bungeanum</i>	Wutai, Xinzhou
31	Germanhuajiao	<i>Z. bungeanum</i>	Germa

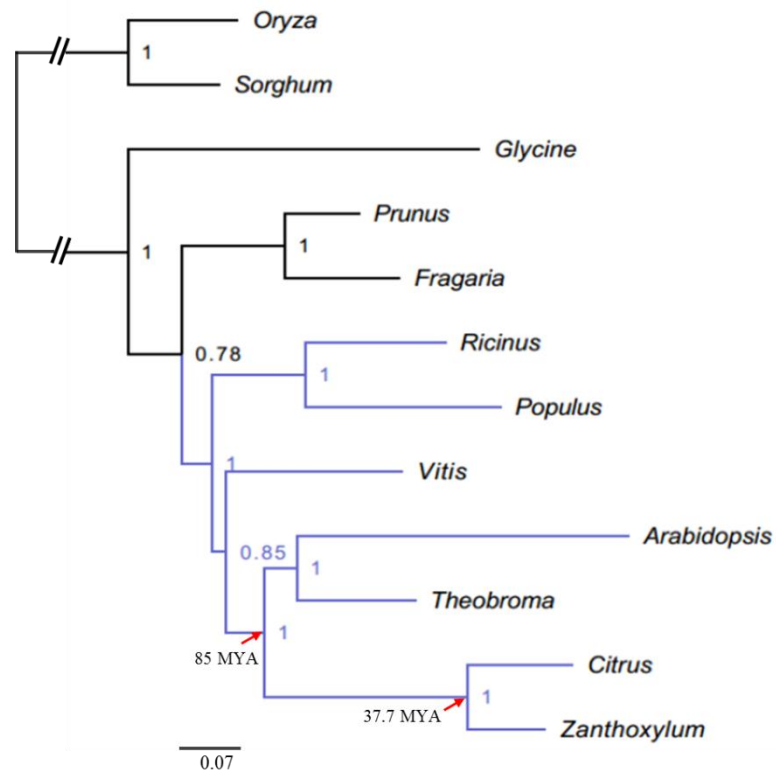
Supplement Table S2. The genetic diversity parameters of 11 SSR primer pairs.

	Primers	N_a	N_e	I	H_o	H_e	F_{IS}
<i>Z. bungeanum</i>	ZB2	2.000	1.658	0.586	0.546	0.406	-0.485
	ZB4	3.000	2.390	0.978	0.864	0.595	-0.375
	ZB11	5.000	4.246	1.523	1.000	0.782	-0.375
	ZB28	6.000	3.796	1.524	0.773	0.754	-0.308
	ZB30	6.000	4.840	1.641	1.000	0.812	-0.281
	ZB32	6.000	2.830	1.266	0.409	0.662	-0.260
	ZB43	7.000	3.270	1.514	0.955	0.710	-0.202
	ZB69	4.000	2.085	0.939	0.667	0.533	-0.199
	ZB73	4.000	1.972	0.937	0.591	0.504	-0.049
	ZB89	4.000	3.133	1.213	0.818	0.697	0.156
	ZB97	4.000	3.753	1.350	0.619	0.752	0.367
Mean	4.636	3.088	1.225	0.749	0.655	-0.183	
<i>Z. armatum</i>	ZB2	1.000	1.000	0.000	0.000	0.000	NA
	ZB4	2.000	1.385	0.451	0.111	0.294	0.600
	ZB11	4.000	1.421	0.634	0.222	0.314	0.250
	ZB28	3.000	1.742	0.730	0.444	0.451	-0.044
	ZB30	2.000	1.670	0.591	0.556	0.425	-0.385
	ZB32	1.000	1.000	0.000	0.000	0.000	NA
	ZB43	1.000	1.000	0.000	0.000	0.000	NA
	ZB69	1.000	1.000	0.000	0.000	0.000	NA
	ZB73	2.000	1.528	0.530	0.444	0.366	-0.286
	ZB89	2.000	1.976	0.687	0.889	0.523	-0.800
	ZB97	3.000	2.455	0.965	0.889	0.628	-0.500
Mean	2.000	1.471	0.417	0.323	0.273	0.027	

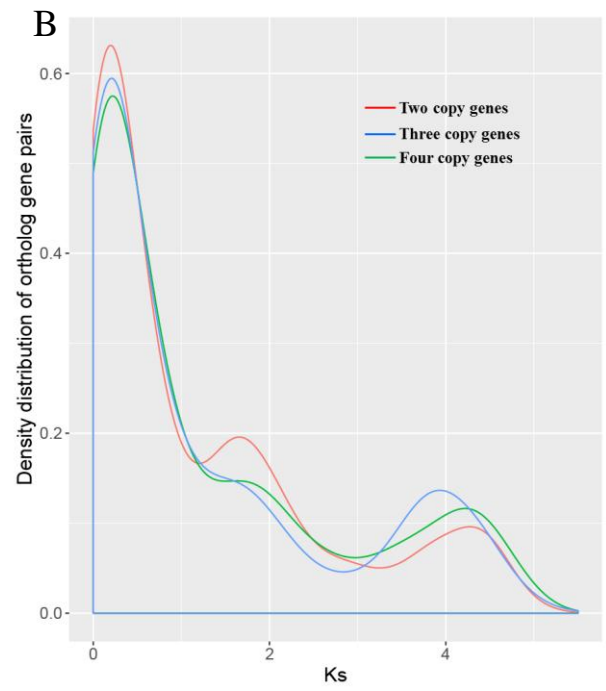
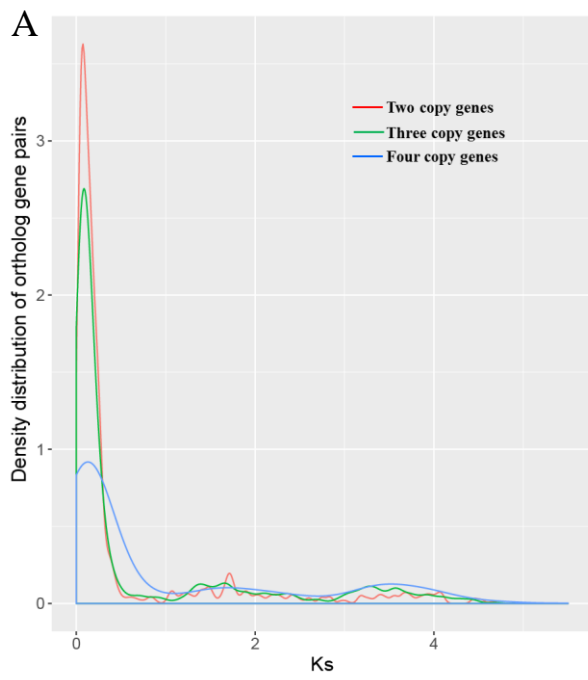
Note: N_a =Observed number of alleles, N_e =Effective number of alleles, I =Shannon's Information index, H_o =Observed heterozygosity, H_e =Expected heterozygosity, and F_{IS} =Fixation index



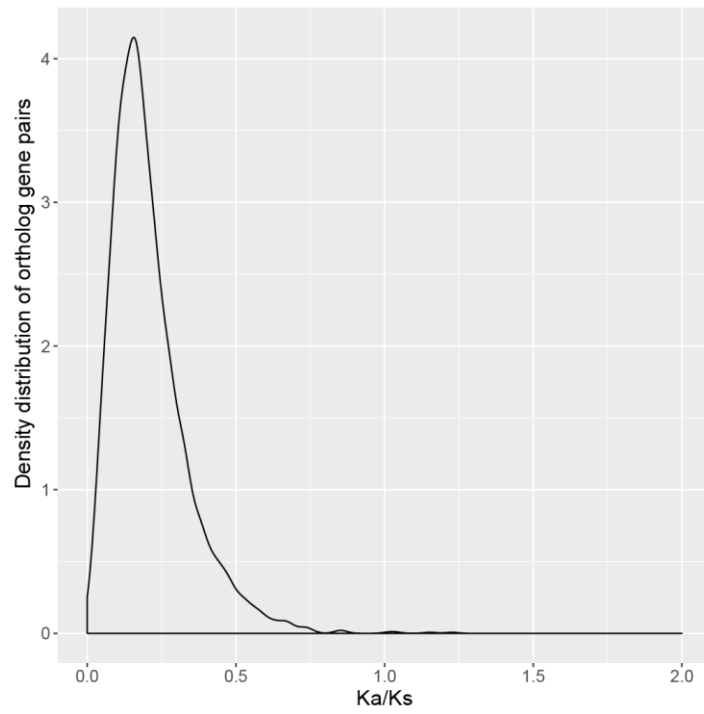
Supplement Figure S2. Characteristics of similarity of *Z. bungeanum* all-unigene against Nr database. (A) *E*-value distribution of all-unigene. (B) Similarity distribution of all-unigene. (C) Species distribution of all-unigene.



Supplement Figure S3. *Zanthoxylum* phylogeny on the basis of 68 chloroplast genes shared in *Zanthoxylum*, sweet orange (*Citrus sinensis*), poplar (*Populus euphratica*), grape (*Vitis aestivalis*), strawberry (*Fragaria × ananassa*), soybean (*Glycine max*), *Arabidopsis*, cacao (*Theobroma cacao*), peach (*Prunus persica*), castor bean (*Ricinus communis*). Rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*) serve as out-groups. MYA, million years ago.

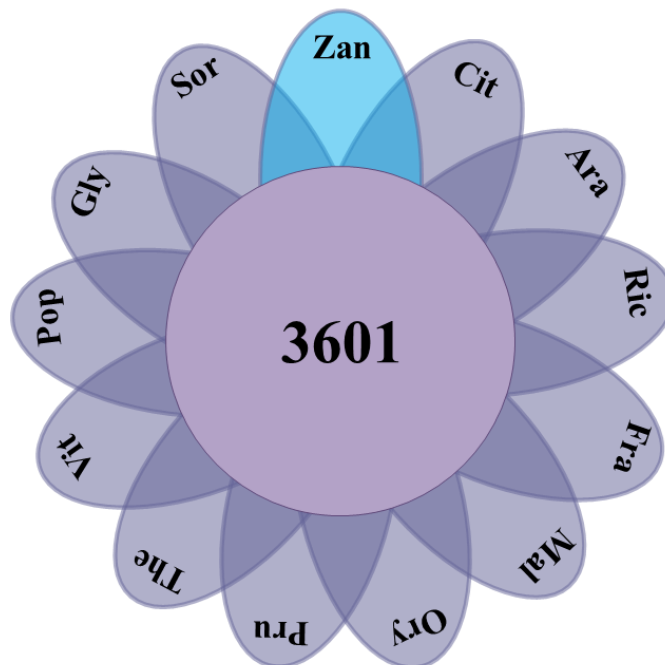


Supplement Figure S4. Distribution of synonymous substitution rates (K_s) for gene pairs of multi-copy homologous genes of (A) *Zanthoxylum* and (B) *Citrus*.



Supplement Figure S5. Ka/Ks distribution of *Zanthoxylum* coding sequences by referring to the *Citrus* genome.

Zanthoxylum
(unique genes 173)



Supplement Figure S6. *Zanthoxylum*-specific genes identified by comparative analyses of *Zanthoxylum* protein coding sequences with 12 other plant genomes. Zan-*Zanthoxylum*, Cit-*Citrus*, Ara-*Arabidopsis*, Ric-*Ricinus*, Fra-*Fragaria*, Mal-*Malus*, Ory-*Oryza*, Pru-*Prunus*, The-*Theobroma*, Vit-*Vitis*, Pop-*Populus*, Gly-*Glycine*, and Sor-*Sorghum*.