

Supplementary Data

**Alternative analyses of compensatory base changes in an ITS2 phylogeny of
Corydalis (Papaveraceae)**

**Meihui Li^{1,†}, Hong Zhao^{1,†}, Fengxi Zhao¹, Lu Jiang², Huasheng Peng², Wei Zhang^{1,3,*}
and Mark P. Simmons³**

¹Marine College, Shandong University, Weihai, 264209, China, ²College of Pharmacy, Anhui University of Chinese Medicine, Hefei, 230012, China and ³Department of Biology, Colorado State University, Fort Collins, Co 80523, USA

†These authors contributed equally to this work.

*For correspondence. E-mail wzhang@sdu.edu.cn

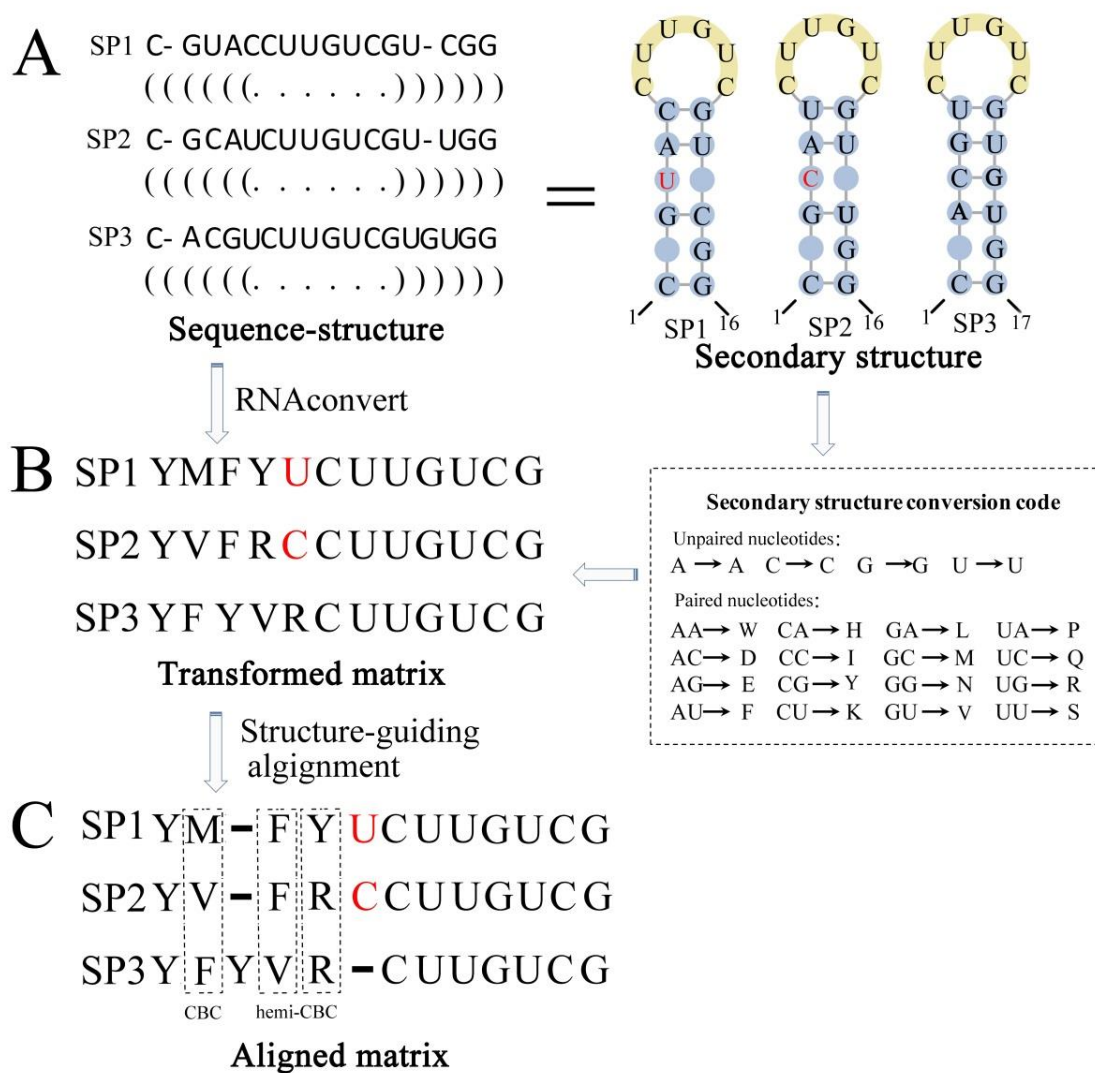


Fig. S1. A workflow illustrating how base-pair information is transformed into an alignment. A. Sequences of three species and their corresponding secondary structures. B. The secondary-structure base-pair information is coded by RNAconvert and transformed into character sequences. The base pairs are listed first and then followed by the unpaired bases in the transformed sequence. The first unpaired nucleotides in sp1 and sp2 are highlighted in red to show their corresponding positions in the transformed matrix. C. Alignment of the coded sequences according to their homologous positions in secondary structure. CBC and hemi-CBC can be identified from this alignment.

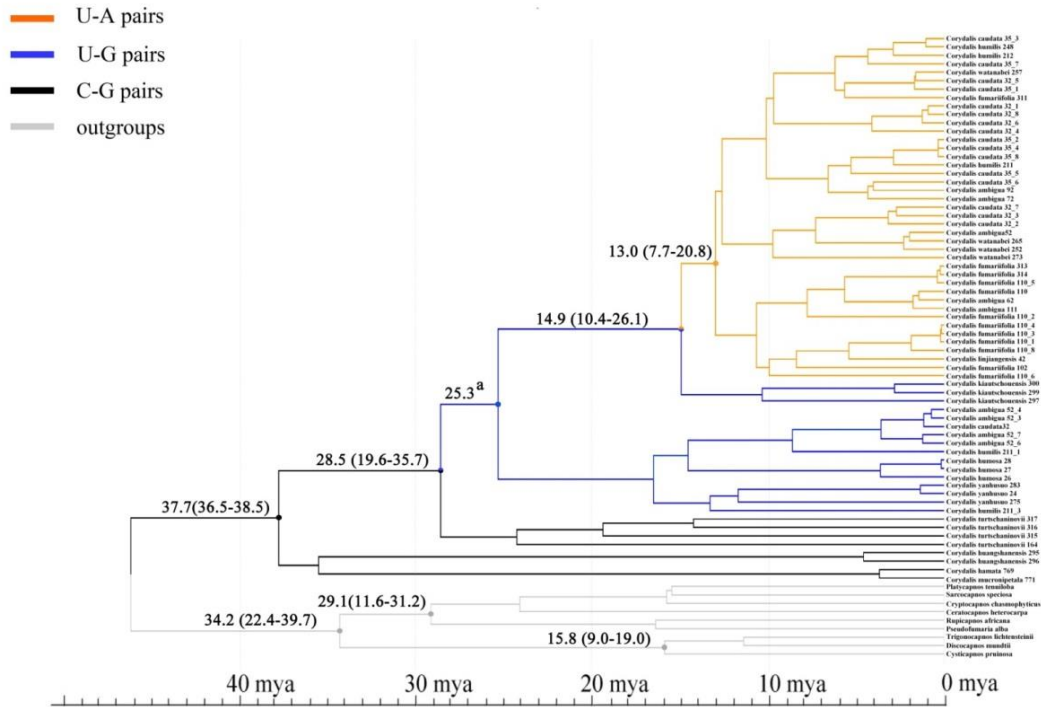


Fig. S2. A Molecular chronogram showing the divergence times of each substitution in the CBC process. Numbers in brackets are 95% highest posterior density (HPD) of this node age. ^a the 95% HPD in this node age was unavailable from BEAST.

Table S1. List of sample information used in this study

Taxa	Provenance	Voucher	GenBank Accession
<i>Corydalis ambigua</i> 52	Ji'an City, China	JA140421017	MH730587
<i>Corydalis ambigua</i> 52_3	Ji'an City, China	JA140421017	MH730612
<i>Corydalis ambigua</i> 52_4	Ji'an City, China	JA140421017	MH730613
<i>Corydalis ambigua</i> 52_6	Ji'an City, China	JA140421017	MH730611
<i>Corydalis ambigua</i> 52_7	Ji'an City, China	JA140421017	MH730610
<i>Corydalis ambigua</i> 62	Benxi City,China	BX140428001	MH730565
<i>Corydalis ambigua</i> 72	TonghuaCity,China	TH140422001	MH730596
<i>Corydalis ambigua</i> 92	Linjiang City,China	LJ140430016	MH730597
<i>Corydalis caudata</i> 32_1	Luoyang City, China	LY140415001	MH730576
<i>Corydalis caudata</i> 32_2	Luoyang City, China	LY140415001	MH730584
<i>Corydalis caudata</i> 32_3	Luoyang City, China	LY140415001	MH730585
<i>Corydalis caudata</i> 32_4	Luoyang City, China	LY140415001	MH730579
<i>Corydalis caudata</i> 32_5	Luoyang City, China	LY140415001	MH730580
<i>Corydalis caudata</i> 32_6	Luoyang City, China	LY140415001	MH730577
<i>Corydalis caudata</i> 32_7	Luoyang City, China	LY140415001	MH730586
<i>Corydalis caudata</i> 32_8	Luoyang City, China	LY140415001	MH730578
<i>Corydalis caudata</i> 32	Luoyang City, China	LY140415001	MH730609
<i>Corydalis caudata</i> 35_1	Luoyang City, China	LY140415017	MH730601
<i>Corydalis caudata</i> 35_2	Luoyang City, China	LY140415017	MH730592
<i>Corydalis caudata</i> 35_3	Luoyang City, China	LY140415017	MH730599
<i>Corydalis caudata</i> 35_4	Luoyang City, China	LY140415017	MH730593
<i>Corydalis caudata</i> 35_5	Luoyang City, China	LY140415017	MH730594
<i>Corydalis caudata</i> 35_6	Luoyang City, China	LY140415017	MH730598
<i>Corydalis caudata</i> 35_7	Luoyang City, China	LY140415017	MH730581
<i>Corydalis caudata</i> 35_8	Luoyang City, China	LY140415017	MH730595
<i>Corydalis fumariifolia</i> 102	Jilin City, China	JL140419001	MH730568
<i>Corydalis fumariifolia</i> 110	Shangzhi City, China	SZ140428002	MH730567
<i>Corydalis fumariifolia</i> 110_1	Shangzhi City, China	SZ140428002	MH730569
<i>Corydalis fumariifolia</i> 110_2	Shangzhi City, China	SZ140428002	MH730574
<i>Corydalis fumariifolia</i> 110_3	Shangzhi City, China	SZ140428002	MH730571
<i>Corydalis fumariifolia</i> 110_4	Shangzhi City, China	SZ140428002	MH730572
<i>Corydalis fumariifolia</i> 110_5	Shangzhi City, China	SZ140428002	MH730563
<i>Corydalis fumariifolia</i> 110_6	Shangzhi City, China	SZ140428002	MH730573
<i>Corydalis fumariifolia</i> 110_8	Shangzhi City, China	SZ140428002	MH730570
<i>Corydalis fumariifolia</i> 111	Mudanjiang City, China	MD140429001	MH730566
<i>Corydalis fumariifolia</i> 311	Huinan County, China	HN140419001	MH730575
<i>Corydalis fumariifolia</i> 313	Huinan County, China	HN140423001	MH730564
<i>Corydalis fumariifolia</i> 314	Tonghua City, China	TH140502001	MH730562
<i>Corydalis huangshanensis</i> 295	Huangshan District, China	HS160323001	MH730623
<i>Corydalis huangshanensis</i> 296	Huangshan District, China	HS160323007	MH730624
<i>Corydalis humilis</i> 211	Dongchang District, China	DC140420005	MH730591
<i>Corydalis humilis</i> 211_1	Dongchang District, China	DC140420005	MH730614
<i>Corydalis humilis</i> 211_3	Dongchang District, China	DC140420005	MH730615
<i>Corydalis humilis</i> 212	Ji'an City, China	JA140421001	MH730582
<i>Corydalis humilis</i> 248	Liuhe County, China	LH140423005	MH730583
<i>Corydalis humosa</i> 26	Song County, China	S140415001	MH730606

<i>Corydalis humosa</i> 27	Song County, China	S140415010	MH730607
<i>Corydalis humosa</i> 28	Song County, China	S140415012	MH730608
<i>Corydalis kiautschouensis</i> 297	Weihai City, China	WH170409001	MH730602
<i>Corydalis kiautschouensis</i> 299	Muping district, China	MP140407001	MH730603
<i>Corydalis kiautschouensis</i> 300	Yantai City, China	YT170407013	MH730604
<i>Corydalis linjiangensis</i> 42	Linjiang City, China	LJ170430001	MH730605
<i>Corydalis turtschaninovii</i> 164	Dongchang District, China	DC140420007	MH730619
<i>Corydalis turtschaninovii</i> 315	Huinan County, China	HN170422012	MH730620
<i>Corydalis turtschaninovii</i> 316	Baishan City, China	BS140424001	MH730621
<i>Corydalis turtschaninovii</i> 317	Liaoyuan City, China	LY140427001	MH730622
<i>Corydalis watanabei</i> 252	Tonghua City, China	TH140422021	MH730588
<i>Corydalis watanabei</i> 257	Benxi City, China	BX140429001	MH730600
<i>Corydalis watanabei</i> 265	Linjiang City, China	LJ140430023	MH730589
<i>Corydalis watanabei</i> 273	Harbin City, China	HRB140418001	MH730590
<i>Corydalis yanhusuo</i> 24	Hefei City, China	HF140309001	MH730616
<i>Corydalis yanhusuo</i> 275	Chuzhou City, China	CZ170315001	MH730618
<i>Corydalis yanhusuo</i> 283	Jinhua City, China	JH170316001	MH730617
<i>Corydalis hamata</i>	GenBank	--	LN610769
<i>Corydalis mucronipetala</i>	GenBank	--	LN610771
<i>Cryptocapnos chasmophyticus</i>	GenBank	--	HE603303
<i>Cysticapnos pruinosa</i>	GenBank	--	HE603318
<i>Discocapnos mundtii</i>	GenBank	--	HE603316
<i>Ceratocapnos heterocarpa</i>	GenBank	--	HE603310
<i>Platycapnos tenuiloba</i>	GenBank	--	HE603312
<i>Pseudofumaria alba</i>	GenBank	--	HE603315
<i>Rupicapnos africana</i>	GenBank	--	HE603304
<i>Sarcocapnos speciosa</i>	GenBank	--	AJ250626
<i>Trigonocapnos lichtensteinii</i>	GenBank	--	HE603317

Table S2. Substitution saturation test for different ITS2 partitions

Partition	Substitution saturation test ^a				
	Iss	Iss.c (Sym)	P	Iss.c (Asym)	P
Paired regions	0.060	0.761	0.000	0.509	0.000
Unpaired regions	0.149	0.897	0.000	0.748	0.000
All regions	0.096	0.691	0.000	0.376	0.000

^a Iss: estimated index of substitution saturation; Iss.c (Sym): critical values for the index of substitution saturation, if the true tree is symmetrical; Iss.c (Asym): critical values for the index of substitution saturation, if the true tree is asymmetrical. Iss < Iss.c significantly indicates no saturation.

Table S3. Comparison of likelihood scores between DNA- and RNA-specific models applied to the ITS2 alignments

Rank	DNA model			Mixed model (DNA model_ RNA model)			
	Model Name	-lnL	AIC	Model Name	-lnL	lnL Adjustment ^a	AIC
1	GTR+G	1462.9765	3199.953	REV+G_RNA16D+G	1370.3015		3020.603
2	GTR+I+G	1462.9833	3201.9666	REV+G_RNA16E+G	1372.6267		3023.2534
3	TIM3+G	1467.3607	3204.7213	REV+G_RNA16A+G	1359.8264		3025.6528
4	TIM3+I+G	1467.3834	3206.7669	REV+G_RNA16A	1361.3247		3026.6494
5	TVM+G	1467.5081	3207.0163	REV+G_RNA16J+G	1364.4634		3030.9268
6	TIM1+G	1469.4458	3208.8916	REV+G_RNA16D	1377.9739		3033.9478
7	TVM+I+G	1467.5125	3209.025	REV+G_RNA16F+G	1378.8042		3035.6084
8	TIM1+I+G	1469.45	3210.8999	REV+G_RNA16C+G	1374.2618		3036.5236
9	TPM3uf+G	1472.023	3212.0459	REV+G_RNA16I+G	1364.2858		3036.5716
10	TPM1uf+G	1472.6944	3213.3888	REV+G_RNA16C	1375.6239		3037.2478
11	TPM3uf+I+G	1472.0272	3214.0544	REV+G_RNA16K+G	1369.9687		3039.9374
12	TPM1uf+I+G	1472.6989	3215.3978	REV+G_RNA16J	1370.2972		3040.5944
13	TIM2+G	1474.9474	3219.8948	REV+G_RNA16B+G	1371.921		3041.842
14	TrN+G	1475.9901	3219.9802	REV+G_RNA16F	1383.6011		3043.2022
15	TIM2+I+G	1474.9514	3221.9028	HKY85+G_RNA16D+G	1386.4661		3044.9322
16	TrN+I+G	1475.9949	3221.9898	REV+G_RNA16I	1370.2123		3046.4246
17	TPM2uf+G	1478.151	3224.302	HKY85+G_RNA16E+G	1388.8091		3047.6182
18	HKY+G	1479.3695	3224.739	REV+G_RNA16E	1385.9934		3047.9868
19	TPM2uf+I+G	1478.1562	3226.3123	HKY85+G_RNA16A+G	1375.9903		3049.9806
20	HKY+I+G	1479.377	3226.754	HKY85+G_RNA16A	1377.5322		3051.0644
21	GTR+I	1490.6966	3255.3931	HKY85+G_RNA16J+G	1380.6521		3055.3042
22	TPM1+G	1497.1791	3256.3582	REV+G_RNA16K	1379.0422		3056.0844
23	TIM1ef+G	1496.3689	3256.7377	HKY85+G_RNA16D	1394.1016		3058.2032
24	TPM1+I+G	1497.1838	3258.3677	HKY85+G_RNA16F+G	1394.918		3059.836
25	TIM1ef+I+G	1496.3736	3258.7472	HKY85+G_RNA16C+G	1390.4142		3060.8284

26	K80+G	1499.5283	3259.0566	HKY85+G_RNA16I+G	1380.4762		3060.9524
27	TVMef+G	1496.5959	3259.1918	HKY85+G_RNA16C	1391.8151		3061.6302
28	TrNef+G	1498.7125	3259.425	HKY85+G_RNA16K+G	1386.2277		3064.4554
29	TPM3+G	1498.7133	3259.4266	HKY85+G_RNA16J	1386.4711		3064.9422
30	TIM3+I	1494.8468	3259.6935	HKY85+G_RNA16B+G	1388.1894		3066.3788
31	TIM1+I	1494.9052	3259.8103	HKY85+G_RNA16F	1399.7225		3067.445
32	SYM+G	1495.9445	3259.8891	REV+G_RNA16B	1385.6902		3067.3804
33	TIM3ef+G	1498.0523	3260.1046	HKY85+G_RNA16I	1386.3946		3070.7892
34	TPM2+G	1499.5185	3261.0371	HKY85+G_RNA16E	1402.1011		3072.2022
35	K80+I+G	1499.5331	3261.0663	HKY85+G_RNA16K	1396.0562		3082.1124
36	TVMef+I+G	1496.6006	3261.2011	HKY85+G_RNA16B	1401.9035		3091.807
37	TIM2ef+G	1498.6893	3261.3786	REV+G_RNA7C+G.empirical	1433.4843	154.7539	3182.9686
38	TrNef+I+G	1498.7168	3261.4336	REV+G_RNA7C.empirical	1436.9497	154.7539	3187.8994
39	TPM3+I+G	1498.7181	3261.4362	REV+G_RNA7F+G.empirical	1446.9	154.7539	3191.8
40	SYM+I+G	1495.9488	3261.8976	REV+G_RNA7G+G.empirical	1449.5637	154.7539	3193.1274
41	TIM3ef+I+G	1498.0567	3262.1134	REV+G_RNA7F.empirical	1449.5859	154.7539	3195.1718
42	TPM2+I+G	1499.5229	3263.0458	REV+G_RNA7A+G.empirical	1430.4465	154.7539	3198.893
43	TIM2ef+I+G	1498.6937	3263.3873	REV+G_RNA7G.empirical	1454.4753	154.7539	3200.9506
44	GTR	1495.6954	3263.3908	REV+G_RNA7A.empirical	1432.1082	154.7539	3200.2164
45	TVM+I	1496.8109	3265.6217	REV+G_RNA7E+G.empirical	1452.7082	154.7539	3205.4164
46	TIM1	1499.1185	3266.237	REV+G_RNA7D+G.empirical	1450.8772	154.7539	3205.7544
47	TPM3uf+I	1499.2883	3266.5766	HKY85+G_RNA7C+G.empirical	1449.7442	154.7539	3207.4884
48	TrN+I	1499.3564	3266.7128	REV+G_RNA7B+G.empirical	1437.8138	154.7539	3207.6276
49	TIM3	1499.8525	3267.7049	REV+G_RNA7D.empirical	1453.596	154.7539	3209.192
50	TPM1uf+I	1499.9509	3267.9019	REV+G_RNA7B.empirical	1439.5691	154.7539	3209.1382
51	TIM2+I	1500.3099	3270.6198	REV+G_RNA7E.empirical	1457.6013	154.7539	3213.2026
52	HKY+I	1504.2704	3274.5408	HKY85+G_RNA7C.empirical	1453.8049	154.7539	3213.6098
53	TrN	1504.7091	3275.4182	HKY85+G_RNA7F+G.empirical	1463.1517	154.7539	3216.3034
54	TIM2	1504.1002	3276.2004	HKY85+G_RNA7G+G.empirical	1465.8433	154.7539	3217.6866
55	TPM2uf+I	1504.9887	3277.9773	HKY85+G_RNA7F.empirical	1465.8306	154.7539	3219.6612

56	TVM	1504.3027	3278.6055	HKY85+G_RNA7A+G.empirical	1446.1848	154.7539	3222.3696
57	TPM1uf	1506.6756	3279.3512	HKY85+G_RNA7G.empirical	1470.6894	154.7539	3225.3788
58	TPM3uf	1508.5862	3283.1724	HKY85+G_RNA7A.empirical	1448.3157	154.7539	3224.6314
59	HKY	1509.779	3283.558	HKY85+G_RNA7E+G.empirical	1468.9788	154.7539	3229.9576
60	TPM2uf	1511.3191	3288.6381	HKY85+G_RNA7D+G.empirical	1467.1223	154.7539	3230.2446
61	F81+G	1514.8886	3293.7772	HKY85+G_RNA7B+G.empirical	1454.0143	154.7539	3232.0286
62	F81+I+G	1515.1861	3296.3722	HKY85+G_RNA7D.empirical	1469.8283	154.7539	3233.6566
63	TIM1ef+I	1518.5697	3301.1395	HKY85+G_RNA7B.empirical	1455.7662	154.7539	3233.5324
64	TPM1+I	1520.2679	3302.5357	HKY85+G_RNA7E.empirical	1473.7931	154.7539	3237.5862
65	TrNef+I	1520.6333	3303.2666	REV+G_RNA7C+G.equal	1472.1475	193.4171	3242.295
66	TIM1ef	1521.2244	3304.4489	REV+G_RNA7C.equal	1475.6129	193.4171	3247.2258
67	TIM3ef+I	1520.2462	3304.4924	REV+G_RNA7F+G.equal	1485.5632	193.4171	3251.1264
68	K80+I	1522.3233	3304.6467	REV+G_RNA7G+G.equal	1488.2269	193.4171	3252.4538
69	SYM+I	1518.3805	3304.7611	REV+G_RNA7F.equal	1488.2491	193.4171	3254.4982
70	TIM2ef+I	1520.5744	3305.1487	REV+G_RNA7A+G.equal	1469.1097	193.4171	3258.2194
71	TPM3+I	1521.8728	3305.7456	REV+G_RNA7G.equal	1493.1385	193.4171	3260.277
72	TVMef+I	1520.0417	3306.0834	REV+G_RNA7A.equal	1470.7714	193.4171	3259.5428
73	TrNef	1523.2345	3306.4691	REV+G_RNA7E+G.equal	1491.3714	193.4171	3264.7428
74	TPM2+I	1522.2613	3306.5226	REV+G_RNA7D+G.equal	1489.5404	193.4171	3265.0808
75	TPM1	1523.7237	3307.4474	HKY85+G_RNA7C+G.equal	1488.4074	193.4171	3266.8148
76	TIM3ef	1522.87	3307.7401	REV+G_RNA7B+G.equal	1476.477	193.4171	3266.954
77	SYM	1521.0103	3308.0206	REV+G_RNA7D.equal	1492.2592	193.4171	3268.5184
78	TIM2ef	1523.2106	3308.4213	REV+G_RNA7B.equal	1478.2323	193.4171	3268.4646
79	K80	1525.7239	3309.4478	REV+G_RNA7E.equal	1496.2645	193.4171	3272.529
80	TPM3	1525.3741	3310.7482	HKY85+G_RNA7C.equal	1492.4681	193.4171	3272.9362
81	TVMef	1523.5213	3311.0426	HKY85+G_RNA7F+G.equal	1501.8149	193.4171	3275.6298
82	TPM2	1525.6989	3311.3978	HKY85+G_RNA7G+G.equal	1504.5065	193.4171	3277.013
83	JC+G	1527.2976	3312.5952	HKY85+G_RNA7F.equal	1504.4938	193.4171	3278.9876
84	JC+I+G	1527.3051	3314.6102	HKY85+G_RNA7A+G.equal	1484.848	193.4171	3281.696
85	F81+I	1539.8552	3343.7103	HKY85+G_RNA7G.equal	1509.3526	193.4171	3284.7052

86	F81	1543.7009	3349.4019	HKY85+G_RNA7A.equal	1486.9789	193.4171	3283.9578
87	JC+I	1549.0168	3356.0336	HKY85+G_RNA7E+G.equal	1507.642	193.4171	3289.284
88	JC	1552.2796	3360.5592	HKY85+G_RNA7D+G.equal	1505.7855	193.4171	3289.571
89				HKY85+G_RNA7B+G.equal	1492.6775	193.4171	3291.355
90				HKY85+G_RNA7D.equal	1508.4915	193.4171	3292.983
91				HKY85+G_RNA7B.equal	1494.4294	193.4171	3292.8588
92				HKY85+G_RNA7E.equal	1512.4563	193.4171	3296.9126

^a In order to comparable with other models, the 7-state models require a likelihood-adjustment value to account for the mapping from 1 mismatch state to 10, and each of the frequencies needs to be defined. This problem can be addressed by using empirical frequencies or by assuming that all non-canonical di-nucleotides are equally likely (Allen & Whelan, 2014). The corresponding results are marked as “empirical” and “equal,” respectively, after their model names.

Table S4. The best-fit substitution rate matrix, mutabilities, base-pair frequencies and substitution rate parameters for the ITS2 paired region in *Corydalis*, inferred using the RNA16D+G model

Base pair ^a	Elements of the rate matrix																Mutability	Frequency
	GC	CG	AU	UA	GU	UG	AA	AG	GA	CC	CU	UC	GG	UU	AC	CA		
GC	----	0	0	0	0.14	0	0	0	0.05	0.06	0	0.03	0.07	0	0.10	0	0.45	31.69%
CG	0	----	0	0		0.14	0	0.05	0	0.06	0.03	0	0.07	0	0	0.10	0.45	31.69%
AU	0	0	----	0	0.37	0	0.05	0.07	0	0	0.06	0	0	0.03	0.13	0	0.71	9.62%
UA	0	0	0	----	0	0.37	0.05	0	0.07	0	0	0.06	0	0.03	0	0.13	0.71	9.62%
GU	1.11	0	0.89	0	----	0	0	0	0.09	0	0.11	0	0.12	0.05	0	0	2.37	4.03%
UG	0	1.11	0	0.89	0	----	0	0.09	0	0	0	0.11	0.12	0.05	0	0	2.37	4.03%
AA	0	0	0.56	0.56	0	0	----	0.68	0.68	0	0	0	0	0	0.28	0.28	3.04	0.81%
AG	0	1.32	0.56	0	0	0.31	0.49	----	0	0	0	0	0.68	0	0.28	0	3.64	1.14%
GA	1.32	0	0	0.56	0.31	0	0.49	0	----	0	0	0	0.68	0	0	0.28	3.64	1.14%
CC	1.48	1.48	0	0	0	0	0	0	0	----	0.26	0.26	0	0	0.22	0.22	3.92	1.27%
CU	0	1.48	1.05	0	0.81	0	0	0	0	0.61	----	0	0	0.26	0	0.22	4.43	0.54%
UC	1.48	0	0	1.05	0	0.81	0	0	0	0.61	0	----	0	0.26	0.22	0	4.43	0.54%
GG	1.32	1.32	0	0	0.31	0.31	0	0.49	0.49	0	0	0	----	0	0	0	4.24	1.60%
UU	0	0	1.05	1.05	0.81	0.81	0	0	0	0	0.61	0.61	0	----	0	0	4.94	0.23%
AC	3.22	0	1.22	0	0	0	0.22	0.31	0	0.28	0	0.12	0	0	----	0.0	5.37	1.02%
CA	0	3.22	0	1.22	0	0	0.22	0	0.31	0.28	0.12	0	0	0	0	----	5.37	1.02%
Substitution rates	r _d =0				r _v =0				r _f =1.03				r _b =4.01					

^a The last ten base-pairs (from AA to CA) are collectively called mismatch (MM) pairs in this study.