Table S1. Genbank numbers for DFR sequences used in PAML analyses. * denotes sequences collected for this study.

Species	Genbank Number	Flower color
Dunalia brachyacantha	JN593319*	Blue
D. solanacea	JN593320*	Yellow
Iochroma baumii	JN593317*	Blue
I. calycinum	JN593314*	Blue
I. cyaneum	GU595064	Blue
I. fuchsioides	Allele A: JN593316* Allele B: JN593315*	Red
I. gesnerioides	GU595063	Red
I. loxense	Allele A: JN593312* Allele B: JN593313*	White
I. parvifolium	JN593318*	Blue
I. umbellatum	Allele A: JN593321* Allele B: JN593322*	Polymorphic (blue, white, or green)
Capsicum annuum	FJ705846	White
Solanum tuberosum	AF449422	Polymorphic (blue or white)
S. lycopersicum	Z18277	Yellow
Nicotiana tabaccum	EF421430	Pink
Petunia hybrida	EU189078	Purple
Nierembergia caerulea	AB078510	Blue
Evolvulus glomeratus	EU189082	Blue
Gentiana triflora	D85185	Blue

Table S2. Data from DFR assays. The name for each sequence gives a basal sequence plus the additional mutations present. The nodes corresponding to the Blue Ancestor and Red Ancestor are shown in fig. 2. The amino acid present at each of the five mutation sites (M1 – M5) is indicated by standard amino acid code. The header gives the amino acid position for each of these amino acids (e.g. mutation M1 is at position 218). Rep = replicate. The amount of product produced in µmoles is indicated for each of the substrates.

	Amino Acid						µmoles product produced		
Sequence	M1: 218	M2: 237	M3: 267	M4: 146	M5: 350	Rep	DHK	DHQ	DHM
Blue Ancestor	L	Q	H	V	A	1	0.002640	0.003718	0.003638
						2	0.002586	0.003655	0.003821
						3	0.002559	0.003655	0.003979
Blue Ancestor + M1	F	Q	H	V	A	1	0.022324	0.008008	0.001544
						2	0.016097	0.005785	0.001675
						3	0.016909	0.006348	0.001413
Blue Ancestor + M2	L	K	H	V	A	1	0.017423	0.006474	0.004345
						2	0.018696	0.006223	0.004790
						3	0.019319	0.006505	0.004005
Blue Ancestor + M3	L	Q	Y	V	A	1	0.006810	0.007257	0.003272
						2	0.005185	0.006317	0.003141
						3	0.006349	0.006881	0.004581
Blue Ancestor + M1 + M2	F	K	H	V	A	1	0.016070	0.005190	0.004371
						2	0.015312	0.005096	0.003062
						3	0.014174	0.006286	0.004581
Blue Ancestor + M1 + M3	F	Q	Y	V	A	1	0.004454	0.004093	0.001073
						2	0.005077	0.003843	0.001099
						3	0.005185	0.004062	0.000968
Blue Ancestor + M2 + M3	L	K	Y	V	A	1	0.014228	0.005472	0.005130
						2	0.017234	0.005847	0.006648
						3	0.015149	0.008228	0.004005
Blue Ancestor + M1 + M2 + M3 (= Red Ancestor)	F	K	Y	V	A	1	0.003317	0.001337	0.000995
						2	0.003290	0.001431	0.001099
						3	0.003290	0.001306	0.000811
Red Ancestor + M4	F	K	Y	I	A	1	0.004400	0.001243	0.000864
						2	0.004725	0.001400	0.000995
						3	0.004535	0.001369	0.000995
Red Ancestor + M5	F	K	Y	V	E	1	0.005727	0.002026	0.001021
						2	0.004833	0.001901	0.000759
						3	0.004779	0.001776	0.000916
Red Ancestor + M4 + M5	F	K	Y	I	E	1	0.003669	0.001118	0.000681

(= *I. gesnerioides*)						2	0.003154	0.000962	0.000602
						3	0.003371	0.001275	0.000654
I. baumii	L	Q	H	V	A	1	0.003588	0.004125	0.004816
(=Blue Ancestor + 105:N → Y)						2	0.002640	0.004250	0.004371
						3	0.003479	0.004062	0.004266
I. cyaneum	L	Q	H	V	A	1	0.002071	0.003498	0.004005
(=Blue Ancestor + 7 substitutions						2	0.002098	0.004031	0.004868
at sites other than M1 to M5)						3	0.002153	0.003843	0.004633

Table S3. Relative activity (RA) of DFR from present-day taxa, inferred ancestors, and possible intermediates. RA was calculated by dividing the activity for a single substrate by the sum of activities across all the substrates. Values reflect averages across three replicates (see table S2).

	Relative Activity on:		
Sequence	DHK	DHQ	DHM
Blue Ancestor	0.26	0.36	0.38
Blue Ancestor + M1	0.69	0.25	0.06
Blue Ancestor + M2	0.63	0.22	0.15
Blue Ancestor + M3	0.37	0.41	0.22
Blue Ancestor + M1 + M2	0.62	0.22	0.16
Blue Ancestor + M1 + M3	0.44	0.37	0.20
Blue Ancestor + M2 + M3	0.57	0.24	0.19
Blue Ancestor + M1 + M2 + M3 (= Red Ancestor)	0.59	0.24	0.17
Red Ancestor + M4	0.67	0.20	0.14
Red Ancestor + M5	0.65	0.24	0.11
Red Ancestor + M4 + M5 (= *I. gesnerioides*)	0.66	0.22	0.13
I. baumii (= Blue Ancestor + 105:N→Y)	0.27	0.35	0.38
I. cyaneum	0.20	0.36	0.43

Table S4. Comparisons of relative activities of DFR on different substrates for alleles that differ by one mutation. A. Possible substitutions involving M1 through M3 between Blue Ancestor (BA) and Red Ancestor (RA= BA+M1+M2+M3). B. Possible substitutions involving M4 and M5 between Red Ancestor and *I. gesnerioides*. Symbols F, P and B as in Table 3.

			Multivariate Analysis				Univariate Analysis									
							DHK			DHQ			DHM			
Comparison			Wilksλ	F	P	B	F	P	B	F	P	B	F	P	B	
A.																
BA	vs.	BA+M1	0.0222	205.35	<.0001	**	205.35	<.0001	**	29.59	<.0001	**	259.82	<.0001	**	
BA	vs.	BA+M2	0.0199	230.33	<.0001	**	230.33	<.0001	**	50.88	<.0001	**	104.61	<.0001	**	
BA	vs.	BA+M3	0.0657	66.32	<.0001	**	66.32	<.0001	**	4.63	0.047		45.38	<.0001	**	
BA+M1	vs.	BA+M1+M2	0.1222	33.52	<.0001	**	33.52	0.0008	*	2.14	0.1632		42.00	<.0001	**	
BA+M1	vs.	BA+M1+M3	0.1167	35.33	<.0001	**	35.33	<.0001	**	51.10	<.0001	**	10.98	0.0044	*	
BA+M2	vs.	BA+M2+M1	0.9521	0.23	0.8707		0.23	0.3996		0.05	0.8196		0.35	0.5639		
BA+M2	vs.	BA+M2+M3	0.5516	3.79	0.035		3.79	0.0038	*	1.00	0.3329		4.33	0.0492		
BA+M3	vs.	BA+M3+M1	0.2241	16.16	<.0001	**	16.16	<.0001	**	0.20	0.6636		36.84	<.0001	**	
BA+M3	vs.	BA+M3+M2	0.0910	46.6	<.0001	**	46.6	<.0001	**	68.66	<.0001	**	1.86	0.1918		
BA+M1+M2	vs.	BA+M1+M2+M3	0.8670	0.72	0.5588		0.72	0.1455		0.90	0.3569		0.31	0.5869		
BA+M1+M3	vs.	BA+M1+M2+M3	0.1160	35.56	<.0001	**	35.56	0.0001	**	58.69	<.0001	**	13.85	0.0019	*	
BA+M2+M3	vs.	BA+M1+M2+M3	0.8462	0.85	0.4902		0.85	0.3413		0.03	0.8579		0.97	0.3395		
B.																
RA	vs.	RA+M4	0.1732	16.71	0.0022	**	37.89	0.0003	**	20.67	0.0019	*	11.69	0.0091		
RA	vs.	RA+M5	0.1738	16.64	0.0022	**	21.14	0.0018	*	0.00	0.9551		36.97	0.0003	**	
RA+M4	vs.	RA+M4+M5	0.5906	2.43	0.1583		0.32	0.5885		4.28	0.0722		2.06	0.1893		
RA+M5	vs.	RA+M4+M5	0.5486	2.88	0.1223		0.99	0.3494		5.85	0.042		1.51	0.2545		

Fig. S1. Maximum likelihood reconstruction of *Dfr* sequence evolution in the red-flowered clade. The species tree and the *Dfr* gene tree differ in the placement of *I. fuchsioides* allele A. However, both topologies place three mutations (M1, M2, M3) along the branch leading from the Blue Ancestor (BA) to the ancestor of the red-flowered species (Red Ancestor, RA). Two subsequent mutations (M4 and M5) occurred along the lineage leading to *I. gesnerioides* in both trees. Amino acids are denoted with standard single letter abbreviations and values in parentheses are the posterior probabilities of the state at the beginning and end of the branch, respectively.

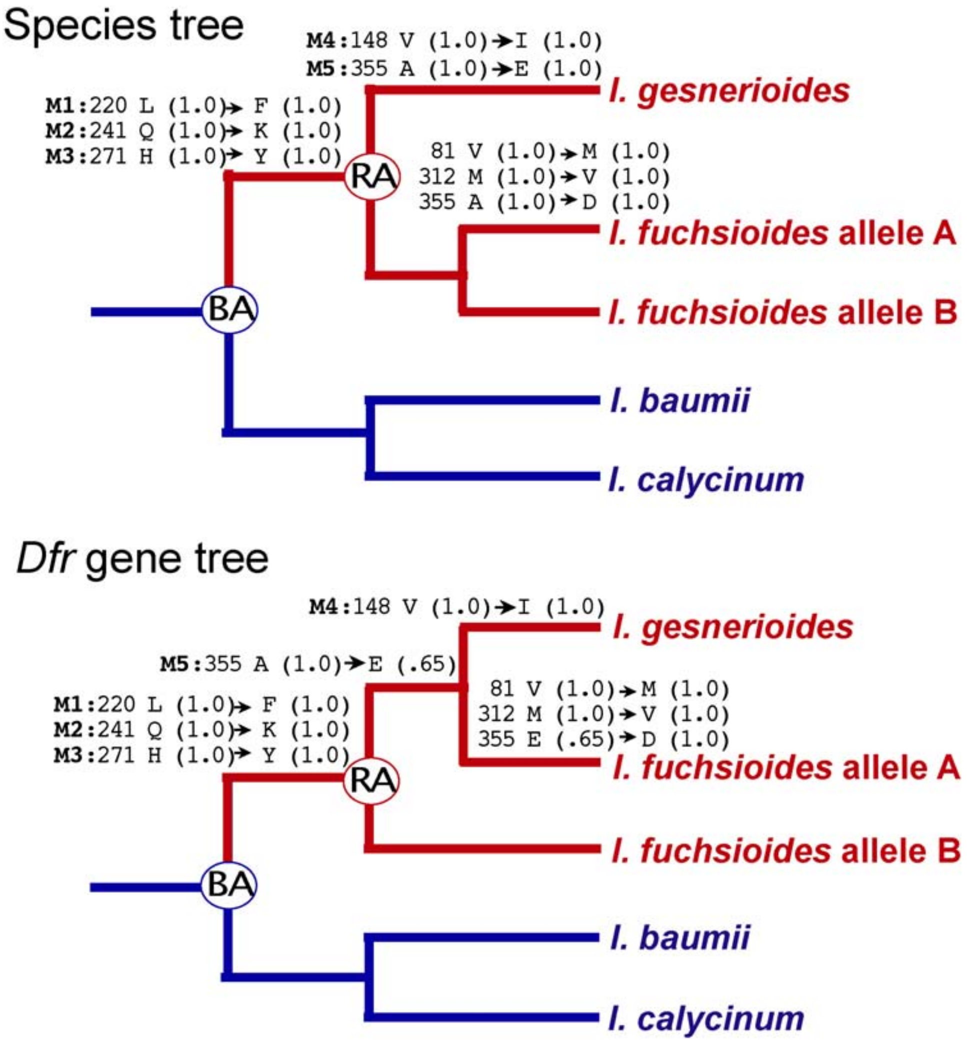

Fig.S2. Location of mutations associated with the transition to red flowers on the DFR enzyme. The crystal structure was created for *Vitis* (Petit et al 2007) with the NADPH co-factor (light gray) and DHQ (yellow) as the substrate. Amino acids involved in the color transition are shown in red. Numbers 1-3 correspond to M1, M2, M3, the amino acids that changed during the transition from blue to red flowers (fig. 2). Numbers 4 and 5 indicate M4 and M5, the substitutions subsequent the transition.