Figure S1. Phylogenetic relationships among 50 selected taxa of angiosperms used as a constraint tree for analyses of rate variation of *clpP* gene. (a) Cladogram of Geraniaceae taken from '(b)' tree. *C.* = *California*, *F.* = *Francoa*, *G.* = *Geranium*, *H.* = *Hypseocharis*, *M.* = *Melianthus*, *P.* = *Pelargonium*, *V.* = *Viviania*. Branches of Geraniaceae are numbered for use in likelihood ratio tests. (b) Maximum likelihood phylogenetic tree inferred from five plastid genes used as a constraint tree for rate variation of *clpP* gene. Bootstrap support values >50% are shown on the branches. Scale bar represents the number of substitution per site.



Figure S2. Maximum likelihood phylogenetic tree inferred from five plastid genes for 44 taxa of angiosperms used as a constraint tree for analyses of rate variation of *accD* gene. Bootstrap support values >50% are shown on the branches. Scale bar represents the number of substitution per site.



Figure S3. Phylogenetic relationships among 17 species of *Geranium* and three related species of Geraniaceae. (a) Maximum likelihood phylogram inferred from 98 organelle genes from 17 *Geranium* and three outgroups used as a constraint tree for rate analyses. Bootstrap support >50% is species shown on the branches. Subgeneric classification follows Aedo *et al.* (1998). (b) The tree in (a) was modified manually by adding two branches for the evolutionary rate variation of duplicated genes and branch numbers are labeled. The numbers '1' and '2' after each species in the cladogram represent paralogs in species that have experienced gene duplications (see Fig. 3). This cladogram was used as the species tree in subsequent analyses. Each branch and species in *Geranium* is given a number that is used for likelihood ratio tests reported in table S9. Scale bar indicates the number of substitutions per site.



Figure S4. Nonsynonymous (d_N) and synonymous (d_S) substitutions in *Geranium* plastid (red) and mitochondrial (blue) individual genes or groups of genes. The box represents values between quartiles, solid lines extend to minimum and maximum values, outliers are shown as circles and horizontal lines in boxes show median values.



Figure S5. Nonsynonymous (*dN***) and synonymous (***dS***) sequence divergence among** *Geranium* **mitochondrial genes**. Colored bars indicate *dN* (red) and *dS* (blue) values. Absent bars in some species indicate that the values of dN or dS are zero.



Figure S5. (continued)



Figure S5. (continued)



Figure S6. Nonsynonymous sequence divergence in plastid-encoded genes and gene groups. Clades A2 and B showed accelerated rates of sequence evolution and are highlighted in shaded gray. Branch lengths in all trees are drawn to the same scale based on the number of nonsynonymous substitution per site. Major clades and subclades (see supplementary fig. S1A) are labeled on *clpP* tree. *C.* = *California*, *E.* = *Erodium*, *G.* = *Geranium*, *M.* = *Monsonia*.



Figure S7. Synonymous sequence divergence in mitochondrial-encoded genes and gene groups.

Clades A2 and B have accelerated rates of sequences and are highlighted in shaded gray. Branch lengths in all trees are drawn to the same scale based on the number of synonymous substitution per site. Major clades and subclades (see Fig. S3A) are labeled on ATPase tree. C. = California, E. = Erodium, G. = Geranium, M. = Monsonia.



0.01

Figure S8. The effective number of codons (EN_c) plotted against GC content at third synonymous site positions (GC_{3s}) for plastid (a) and mitochondrial (b) genes in *Geranium*. Line shows the expected ENc as a function of GC_{3S} given random codon usage (Wright 1990).



Figure S9. Correlation between synonymous substitution rates and (a) GC content in third synonymous site positions (GC_{3S}), and (b) the effective number of codons (EN_c). Linear regression analyses using gene or gene groups (*atp9*, blue; *atp* gene group, green; other genes, red). Significance of fit was evaluated by Pearson's correlation in the R package.





Figure S10. Nonsynonymous (d_N) and synonymous (d_S) values for four mitochondrial genes. Color bars indicate the values of d_N (red) and d_S (blue). Significance of fit was evaluated by Wilcoxon rank sum tests in the R package. Numbers at nodes on tree are bootstrap values. *G*. = *Geranium*. Scale bar indicates the number of substitutions per site.



Figure S11. Phylograms of plastid-encoded polymerase (PEP) genes *rpoA*, *rpoB*, *rpoC1* and *rpoC2* showing nonsynonymous (d_N) and synonymous (d_S) substitution rates. Geranium reflexum and *G. phaeum* are highlighted in shaded gray. All trees are drawn to the same scale. *C. = California*, *E. = Erodium*, *G. = Geranium*, *M. = Monsonia*.copy Scale bar indicates the number of substitutions per site. For *rpoA*, copy 1 of this gene was used for *G. reflexum* and *G. phaeum*.



0.04

Figure S12. Rapid structural evolution of the plastid-encoded *clpP* gene. (a) Schematic diagram of the variable structure of *clpP* among selected angiosperms. Arrowheads indicate the positions of the first (black) and second (gray) intron. Pink boxes indicate the conserved domain of caseinolytic protease. *Geranium* and *Monsonia* highlighted in red have a variable length at N-terminus. (b) Phylograms showing nonsynonymous (d_N) and synonymous (d_S) substitution rates for the *clpP* genes among representative angiosperms, including Geraniaceae. Branch lengths are drawn to the same scale based on d_N (left) and d_S (right) substitutions per site. Branches with significantly higher d_N/d_S ratios detected by likelihood ratio test (LRT) are marked with asterisks (*, p < 0.05; **, p < 0.01; ***, p < 0.0001 after Bonferroni correction). Colored asterisks indicate different models (i.e. black, 'branch leading to' and red, 'branch within'). Clades A, B, A1, A2, and A3 corresponding to different lineages of *Geranium* (supplementary fig. S1A) are highlighted in gray. Closed black circles on branches indicate lineages that have divergent *clpP* genes. *C. = California, F. = Francoa, G. = Geranium*, *H. = Hypseocharis, M. = Melianthus, P. = Pelargonium, V. = Viviania.* Gray rectangles at tips of *ds* tree indicates presence or absence of introns 1 and 2 in *clpP* gene. Scale bar indicates the number of substitutions per site.



0.1

Figure S13. Amino acid alignment of plastid-encoded *clpP* gene for *Nicotiana tabacum*, *Geranium* (G.) and *Monsonia emarginata*. Red box indicates caseinolytic protease conserved domain.

	1 10	20	30	40	50	60	70	80	90	100	110	120	130	140
Nicotiana_tabacum	MPIGVPKVPF	RSPG												
G incanum	MPVGVPRVHVVY	DSEGNI	PW KH - KK	KNFDEDY-	FGKKFDEEFGDI	D L G		EDS	ENEFVEGV			I	IN EGOE – EEFGF	KEFENEK
G sanguineum	MPVGVEKVHVTV	DSKGNI	PWRR-08	EGFDDDD-	FGDKFDDDFYD	S – – F G – – – – – – –		GGS	PT PG P				EGEEE-EGVGE	KGFENEK
G brycei	MPVGVPKVHVVV	DSEGNT	PWKR-KK	KNFDDDD-	FGKKFDRRFGDI	D T. D		EG S	ENELVEGV				IN REOR - EGEGE	KEFBNEK
G nodosum	MPVGVPKVYTTY	DSEGNI	PI KP-KK	ERFDDDD-	FGOKFDKDLGDI	DDVFGOKVDKEL	DDDDDFDDDFI	HGLGKG S	BKBFVBGV				GEEDE-KEGE	KOFONEN
C nlatvanthum	MPVGVPEVVVVV	DSKGDT	PW	KGEDDDD-	FGEKEVEGEGDI	DeeFGeeeeeD	DEGDDEGDDEG	DDFGKG == S	ENERVEG S	ENELNEEVEGS	ENELNEEVE	GSENEEVEGE	DEEOROEKEGE	RELENKK
C traverrii	NPVCWPFWHVVV	DSEGNT	DW	KGEDDDD-	FGFKFDFFFGDI	D F C			ENEL VECU				DEFER	FFFNFK
C_oraversii	M D U C U D V T U U U V	DEVENT		K D F D D D A	PCFYFAFFICAI	D FC	ADECANE	ACCDCD C	INFIT VACU	EDE		DED	DEDERECCUCI	CIENEY
G_praterise		DCRONI		KAPDDDA-	POERFALLOAD	D	ADFGANFC	JAG SD SD == S		EDE			DEDERLOGVOR	C D D D V D I
C_entressii		DDEGKT	PERFICUENCE AND A A A A A A A A A A A A A A A A A A	RREDDDL-	VNV PV PV P	3							DESER-RATOR	A DF BRED
G_renexum		DDEGKI	PERININKKK	RRV BALD -	KNKEKE KE				EVEN DENT				DATHT-ENPNB	ANTD SDT
G_pnaeum		DDEGKI	PERIHIHRAK	RRVEALD-	KNKE			K	O REALDENT				DATHT-ENPND	NTUSUT
G_subcaulescens	MPIGVERVPVPF	DSEGNI	PRIIRMANA	RRFDDNA-	FGHLFDEEVEDI	K L N		K K S	SERLDEDE	N L N		KAS:	LEEDDE-EEDDE	SDDNLNK
G_platypetalum	MPVGTPRVRVPY	DSEGNI	PHRVERRDANK	QREHEEG-	E NENEFFLEC	G		1	DEDFANSS				KGSG8	SGSDSDS
G_renardii	MPIRVER		KEANK	QRFHEED-	F N F N E F F L E C	G		I	DEEFAKSS				KGSGS	SGSDSDS
G_yeoi	MANAVERV PE		PMP EPDDE	VPFYDISA	X S S D X D P EQ Y Q I	D		P	DODT DOD B				<u> </u>	RELDDEV
G_robertianum	MANAVERV PE	YDPAYY	PMP EPDDE	VPFYDISA	Y S S D Y D P E Q Y Q I	D		P	DODLDODP			·	IQ D Q D Q D F	BELDDEV
G_maderense	MANAVRKV PE	Y D P A Y Y	P V P E P D D E	VPFYDISA	Y S S D Y D P E Q Y Q I	D		P	DQDLDPDI			I	1Q D Q	- ELDDEV
G_macrorrhizum	MAEAVPKVLF	Y D P A Y Y	PVP EPDHE	RPFYATSG	YSSDYAPEQYRI	D		P	DQDLDPDQ					- D L D P D Q
Monsonia_emarginata	MAFCIPKVRVRFLKKK	KKENKKKV	F D N F P D – E K	EELDENE-	-DLDKNEDLGAI	NEDLGENQDLDE	NEDLDENEDLO	GENQDLDE	NEDLDENA			· I	LGENEDLGEN	D L D ENA
	150	160	170	180	190	200	210	220	230	240	250	200	220	200
Nicotiana_tabacum								BED	A SWV DVYN	REVRERLEFEG	OEVDSEISN	QLIGEMVYI SI	EDE-TKDLYLF	FINSPGG
G incanum	KV GOASENEYENEN	YGNANYGN	DE		DFDFESOK	OVCFEDDDKDDN	ALEAA KREEK	K K B B B B D	NOWYDUFE	VIEREGVIEVG	REMINKFAN	TVVSLMTYDD	YNPNRTPOP	FINSPGG
G sanguineum	KV GOASDDESENGD	YF	FDOEFEFESOK	OVRFKADN	HFFDOEFESOK	VRFEDDDKDDN	AL EAA BKRKK		TOWYDEFE	VIEREGVIEVG	REMINKEAN	KVVSLMVYLD	DNPNRTPAVE	FINSPGG
G brycei	KV GOASEYEYDDDD	Y			EDLESOK	VRFEDDEKEDN	ALEAARKRKK	EBBBBO	TRWYDEF	VIHREGVIEVG	REMTMERAN	TVVSLMVYDD	DNPNRTPRVF	FINSPGG
G nodosum	KI KFGKOL ENDRNNGP	Ŷ			FDGELESOK	VRFEDDEKDEN	ALEAAEKRKK	- EKDEEDOE	HNWHDTFF	VTHRECVTEVC	REMEMBEAN	GVVSLMVYDD	DNPDBT PDVT	TNSPGG
G platvanthum	TV GOASODDYHNYF	¥			NEDNNDFESOK	VRFEDDEKDEN	ALVAAEKRKK	- REREER	TONDDEFE	VTHREGITEVC	REMANKEAN	SVVSLMVYDD	DKPNBTPMV	TNSPGG
G traversii	KV GOASEDEGENDD	v		FS	VESDEDEESOK	VEFEDDEKDDN	AT. FAADSSCR		TONYDER	VTHE BOLLTEVC	REMARKEDN	TWVSTMVYTD	OR PNRT PONT	TNSPCC
Coratence	KU COASELOVEDWE	G FP	PDNKUT PNFKK	VCOASEDE	VNENNEEDCAR	VCFFDDDKDDN	AT FAALS		X X MD DT F	TTPPFCVVFVC	DTMMMFT	CWV ST WVVT D	TOND TOTDN	TNCDCC
G_praterise G_andrassii	GOASEDESNEET	D	F K K K V D D K D K K	VGQADUDU	DDDFOK	VEFEDDKDDN	ALAAAEKPKK		NEWHDIER	VTHPECTVEVC	DEMENSEDN	SVVSTMVVTD		FTNSPCC
C reflexive		v			PREPORT	A D U D D U D D D D	DYFAAOYPYY		TO NEW YORK T	WINCH TRIC	O T T D Z Z T A N	SWUCT WURT DO	DOUDBUB	TNGDCC
G_reliexum	DV NYNDOODDVND	v			br SKQ br 1		P I DAROXXXXX			VIIC EGHTFIG		G V V G L M V H L D V	EDOUDDUDE	TNCDCC
G_phaeum	VC CREIDEDED	1	DEDID	EDED	LF 3 3 2 LF 1	RAR V DOV DP DD S	PILAAKAAAAA				DRUDNERAN			TNSPGG
G_subcaulescens	KSSEEDEDEDDDF	6		5080	LLW I SMN		PIERAKKKKK			VLAR BOVLELG			GARDATPIN	TNOPCO
G_platypetaium	GIDSODDEDFLN	G			EEFKPUSQKME	EVREDUDIKUUS	PIEARKARKA		VNWYDDDD	VVEREGELFFG	EENTTKFAN	SVI SLMIYLE.	ENKPNET PERM	<u>INSPGG</u>
G_renardii	GIDSQDDEDFLN	G			LEF EPN SQL	EV REDEDERED S	PILAAKARAA	ALLERAK	VNWIDLL	VVHREGELEEG	O LA TITKE AK	SVISLMIILE	IN KPNRIPHU	INSPGG
G_yeoi	PFYDIFA	¥			SSDYDPMQ		E E		PLWTDLLH	WLYDGCYLELG	RELKABLAN	TLISLILLD	PKSPSIQUAR	TNSPGG
G_robertianum	PFYD1FA	¥			SSDYDPMQ		BEKQEEI		PLWTDLLH	VINDGETIELG	RELKAPLAN	HLISLIIHLDI	APRSPSIQUES	FINSPGG
G_maderense	PFYDIFA	¥			SSDYDPMQ	YKDPDLDDE	BEKQEEI	X QEEKPE	PLWTDLLH	VLYDGCFLELG	RELKAELAN	TLISLIIHLDI	PKSPSIQUE	FINSPGG
G_macrorrhizum	DPDQDPDQDIHQYL	¥			QML1	RLAFDKDEEKK	EEKKEE	а — — — К ррк Р G	PAWTDELC	VLYTGCYLFLG	GELKGKLVN	NELSEVIELDI	PKAPSTOWN	FINSPGG
Monsonia_emarginata	DL GENADLGENEDL	G		ENQ	DLDENEDLGEN	EDSDEÐSÐEÐLD	EYENEDGKKK	K – – K K K K K S V	FLWVDLFH	H IFRKRMIFCC	Q D L N V D I A N	NIAGAMIFINI	JEDSSLK-QEVE	FINSRGG
	290 30)	310	320	330	340 35	0 360	37	0	380	390	400	410 43	20 426
Nicotiana_tabacum	WVIPGVAIYDMMQFVR	P D V H T H	CMGLAASMGSE	ILVGGEIT	KRLAFPHARVM	IHOPASSFYEAO	EGEFVLEMEE	LKLRETLER	VYVQRTGE	PLWVVSEDMER	DVFMSATEA	QAYGIVDLVAV	/ E	
G_incanum	FVFAGLAIYDFMDFLR	YTNKVQTE	VIGHAASMASV	VLIGGGM-	- RVAFTHARVM	IHQPRMKAFQDE	SSQIALEARV	LDLRHMITE	IYBRKTTI	RYSVIIVDLER	DKFMTAIOA	R D Y G I V D G V A H	? R K D D V L J	DETF
G_sanguineum	FWFAGLAIYDSMNILR	YGNKVQTI	VIGIAASMASV	VLIGGGY-	- RVAFTNARVM	IHQPRMKAFEDE	SSQIALEARV	LDLRHSITE	IYERKTTE	PWLINVDLER	DKFLSAEEA	R D Y G I V D G V A F	RKTD DDLFP	DETE
G_brycei	FVFAGLAIYDTMYLLR	YANKVOTH	VIGIAASMASV	VLIGGGM-	-RVAFTHARVM	IHQPRMKAFEDE	SSEIALEARV	LDLRHITTE	IYERKTTR	PYSVITVDLER	DKFMTAIOA	RKYGIVDGVAH	RKTDEDVF	DETE
G_nodosum	FVFAGLAIYDTMDILR	YSNKVQTV	VIGIAASMASV	VLIGGGL-	- RVAFTNAKVM	IHQPRMKAFEDE	SSQIALEARV	LDLRHSITE	IYERKTTR	PWVIIVDLES	DKFMSAIEA	R D Y G I V D G V A H	RKTDDDVL	DETF
G_platyanthum	FVFAGLAIYDFMDLLR	YGNKVQTI	VIGHAASMASV	VLIGGGL -	- RVAFTNAKVM	IHQPRMKAFEDE	SSEIALEARV	LDLRHSITE	IYBRKTTR	PYSTITVDLEY	DKFMSATEA	R D Y G I V D G V A F	RKRDDDVF	DETF
G_traversii	FVFAGLAIYDEMDELR	YGNKVOTH	VIGHAASMASV	VLIGGGL-	-RVAFTNAKVM	IHQPRMKAFEDE	SSEIALEARV	LDLRHSITE	IYBRKTTR	PWWAITVDLEC	DKFMSATOA	OKYGIVDGVAH	RKRDDDVLF	DETF
G pratense	EWFAGLALYDSLNLIR	YGNKVOWF	VIGIAASMASV	VLIGEGI-	-RWAFINAKVM	IHOPRMKPFEDE	ASIGUGINGRVI	LDURHITIE	TYERKTTR	PWWINNRDOFC	DKEMTADDA	OAYCIVDVLAE	RKRDDDVFF	DIF
G endressii	FVFAGLALYDTTELLN	YGTKVOTV	VIGHAASMASV	VIIIGCCL-	-RVAFTNAKVM	THOPRMKAFEDE	SSEDADEARV	LDDRHSTAE	IYERKTTR	SMSDITVDLEC	DKEMSATOA	KLYCIVDGVAE	RKRDDDVL	DETE
G reflexum	FALGGLAIYDTLDEVR	- GKEVOTH	AIGHAASIASV	VLIGCCL-	- RVAMHNAKVM	TROPRMEMERDE	ASDUVIDSRI	LDDEL STRE	TYVRKTGE	THMSTTRDDEV	EKEMTAKDA	ENVELIDGVA	A KONT D DTMEP	IL INS
G phaeum	EALGGIATYDETDIVR	GKAVOTH	ATGUAASTASV	VI. T GGGI -	- RMAMHNATVM	ROPRMEVERDE	A SIDILY TRISPET	TDERNSTRE	TYVRKTCK	THMCTTRDTTV	EKEMTAKDA	ENVETTORY	NET D DTMEP	ALL TRAS
G subcaulescens	EVILSGISTYDMDIVR	-VKTVOT	ALGUAASMASV	VI. TGGGT -	- BMALAHAKVM	THOPRMKAFEDE	STIDIA TORAR VI	LDDENSTOR	TYVRKTGE	THIS TRADMEN	DTELTADAA	RAVETVOGT	BESDD SDDGL7	ABLWSGT
C platypetalum	EWEAGT STYDEMSTMP	DGNKUONI	VICIAASMAST.	VITTCECT.	- RVN SPENRVM	THOPPMEPEDE	SISISII A TRASPR V	TDIETNIGE	TVAORTCE	SUSESSE	DTEMPAKOA	L SVR TVDGV TI	EHP	ATTCSL.
C renardii	EVER SCHOLDENS LTR	DGNKMONI	VVGTAASMAST	VIELCEGT -	BUDSPENRVM	THOPPMKPEDD	STRATATENEV	LDDEVNDOR	TYVORDGH	SEISKINSEDMER	ATCMONKON	L SVR TVDAV TT	DER POTY	ST. T
G veni	EWI, SGMATEDAMI, VIN	GNDD	CVGTAASMATT	MARACKO -	BINDRESCOM	THOPSMKREPEDE	TO U DAVI V TO ELLI TO	ADDERD	TVVBVBBB	PWWVTMSDTER	DTEMSARES	Y LVCDVDCW St	7	
C robertianum	EWI, SOMATED AMI, VIN		CUCHAASMATV	WANT A CKO	PINT PHOLIT	THOPSMERED	UV DAVUT DEH T	ADTOHDING	TVVDVDHE	DWWVTMSDTET	DTEMSSHER	VI.VET VDEV SI	7	
C maderense	EWIT SCHATED AMT. VIN		CVCDAASMAT	VALUE A CKO-	PILRESCT	THOPSMERED	AV SAVUT SHUM		TVVDVDUD	DWWVTMSDT	DTEMSNUM	VI.VET VDEW OT	7	
C macroschizum	VUL SOLATER AMURIAN		GVCDAACMATV		- PITAT PERAT						D C Z M C N U O N	VEVETWORVET	7	
G_macrormizum			OVCIAACMATV		MERTODANTA						E D B B B B B B B B B B B B B B B B B B		125	
monsonia_emarginata	THEFT AVIDUATALS	ABVMDD	C V G BAASAASL	UNDER D RD	AMPICPAAKIN.	THORNER DE BAK	HELENADET HE	HI THERE AV UR	LIAKETFE	DEWVMARDER	CIT NOT PARTY		E 11	

caseinolytic protease

Figure S14. Nucleotide alignment of the conserved domain sequences of the plastid-encoded *clpP* gene. Exons are annotated on *Arabidopsis thaliana*. The C to U editing site at alignment position is highlighted in pink.



Figure S15. Correlation between nonsynonymous (d_N) and synonymous (d_S) substitution rates of *clpP* (a) and *accD* (b) genes. Boxplot distribution of the values of d_N and d_S for *clpP* and *accD* genes between divergent groups (D, red) and other angiosperms (O, blue) that contain canonical *clpP* and *accD* genes. The box represents values between quartiles, solid lines extend to minimum and maximum values, outliers are shown as circles and horizontal lines in boxes show median values. Significance of fit was evaluated by Pearson's correlation and pairwise Wilcoxon rank sum tests in the R package. (a) Linear regression analyses included divergent groups (red circles) and other angiosperms (blue triangles). In particular, Geraniaceae data points are highlighted in closed red circles. (b) Linear regression analyses included four branches leading to divergent lineages (see Fig. S15, closed red circles) and other angiosperms (blue triangle). *Geranium, California*, and *Pelargonium* terminal branch values and the value of the internal branch were marked with red open circles.







Figure S16. Phylograms of plastid *a ccD* gene showing nonsynonymous (d_N) and synonymous (d_B substitution rates. Rectangles indicate the conserved domain (acetyl-CoA carboxylase beta subunit D; pink) in the predicted ORFs. Arrowheads indicate the lineages with a disrupted *accD* gene. *C. = California*, *G. = Geranium*, *F. = Francoa*, *M. = Melianthus*, *P. = Pelargonium*, *V. = Viviania*. Scale bar indicates the number of substitutions per site.



Figure S17. Characterization of the nuclear-encoded acetyl-CoA carboxylase beta subunit (accD) genes of Hypseocharis and Monsonia. (a) Architecture of the nuclear-encoded accD gene from Hypseocharis bilobata and Monsonia emarginata. Pink boxes indicate the conserved domain of the acetyl-CoA carboxylase subunit β (Lee et al. 2004), green boxes in N-terminus indicate a transit peptide.
(b) Amino acid sequence alignment of the nuclear-encoded accD regions from *H. bilobata* and *M. emarginata*. Colors indicate subunits of ACC corresponding to the schematic diagram in (a).



B Target peptide

Hypseocharis Monsonia	NATVLRTPPFRALSP NAVSQKLTLALFGGAC	PVATVPATATR	-KPSCISETQ IRSQQISES-	40 54 RNAKTSCISSLW - TTOGSPIFSGP	GORKPF RFVR L KSSVAKV OFLK A	79 VPF AS QEENGT E ADQ SS NRSRYL K	[®] ELEAAQTPE TAISSFAAQHE	9 SKAQA QQ PQP BI L FLGAV QK ANI BI S	100 110 VEBGEEEAE NKLNVIEKTISEG	120 NGEWSTEBETSGKAL NG-ALDOVT	130 KGEEEEEE SDKALEG KKMDKE	150 EEFYD QVREFAESMK EEDDFFVYGRPFKEHLA
Hypseocharis Monsonia	RENABELKEEEELKR -DNTEYQQYGYGPEE	I SP I S ELQGEEEEE I S	GEEDEHYEDR DEPYODR	IESYKKETGLTD LVRYKKDTGLLE	AVOTGIGKINGI AIOTGKCKINGV	PVAFGVMEFOFV PVAFGSMEFGFM	GGSMGNVVGEK GGSMGSVVGEK	TRLIEYATDQEL TRLIECATNEFL	PLILVCSSGGARM PLILVCSSGGARM	QEGSYSLMQMAKISS QEGSESLMQMAKISG	NIHNYROQSKKNH Alyyfroksvorsnh	KLFYVTILASPTTGGV (OKLFYAAILASPTTGGV
Hypseocharis Monsonia	IASFGMLGDIIIAEPN LASFGMLGDVIISEPK	ATIAFAGARVIE TFIAFAGPRVIE	OTLKVPVPED SILKEVVPEG	AQE <mark>AEPLFEKGI</mark> SHAAEPLLEKGI	FDLLVPRNLLKG IDEIVDRYDLKS	VLTHLELHGFF VLAMLLDFHYYS	400 211 NONETI 211 PYVF SONPVI	410 NSVDLN I SRQ NLAGAT I QY B SG	GDEVEDGPSSAAS	putatuve cat	alytic site	

acetyl-CoA carboxylase beta subunit D

Figure S18. Nuclear acetyl-CoA carboxylase (ACC) divergence among selected angiosperms.

(a) Maximum likelihood tree showing nonsynonymous (d_N) and synonymous (d_S) substitution rates for the nuclear ACC homologs. All trees are drawn to the same scale. *G*. = *Geranium*. The numbers '1' and '2' after each species in the phylograms represent paralogs for ACC. (b) Boxplot distribution of the values of d_N and d_S for *Geranium* ACC orthologs (ACC1 and ACC2). The box represents values between quartiles, solid lines extend to minimum and maximum values, outliers are shown as circles and horizontal lines in boxes show median values. Significance of fit was evaluated by Pearson's correlation and pairwise Wilcoxon rank sum tests in the R package. Scale bar indicates the number of substitutions per site.



Figure S19. Chronogram of Geraniaceae divergence times. Times shown are the median age estimates from the BEAST analysis. Blue bars indicated 95% highest posterior density intervals in age estimates. G. = Geranium. The numbers '1' and '2' after each species in the phylograms represent paralogs for ACC. Ma = million years ago.

