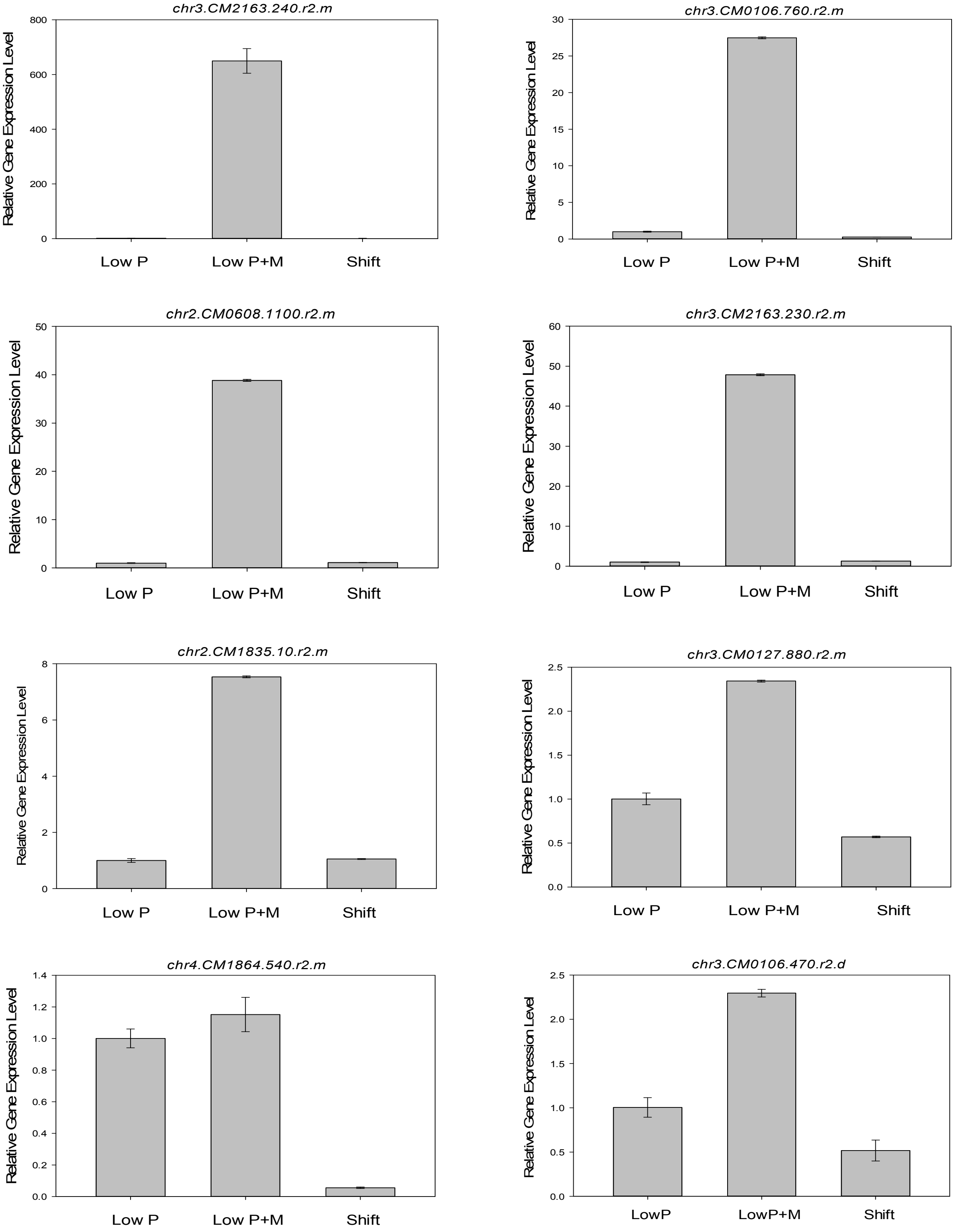
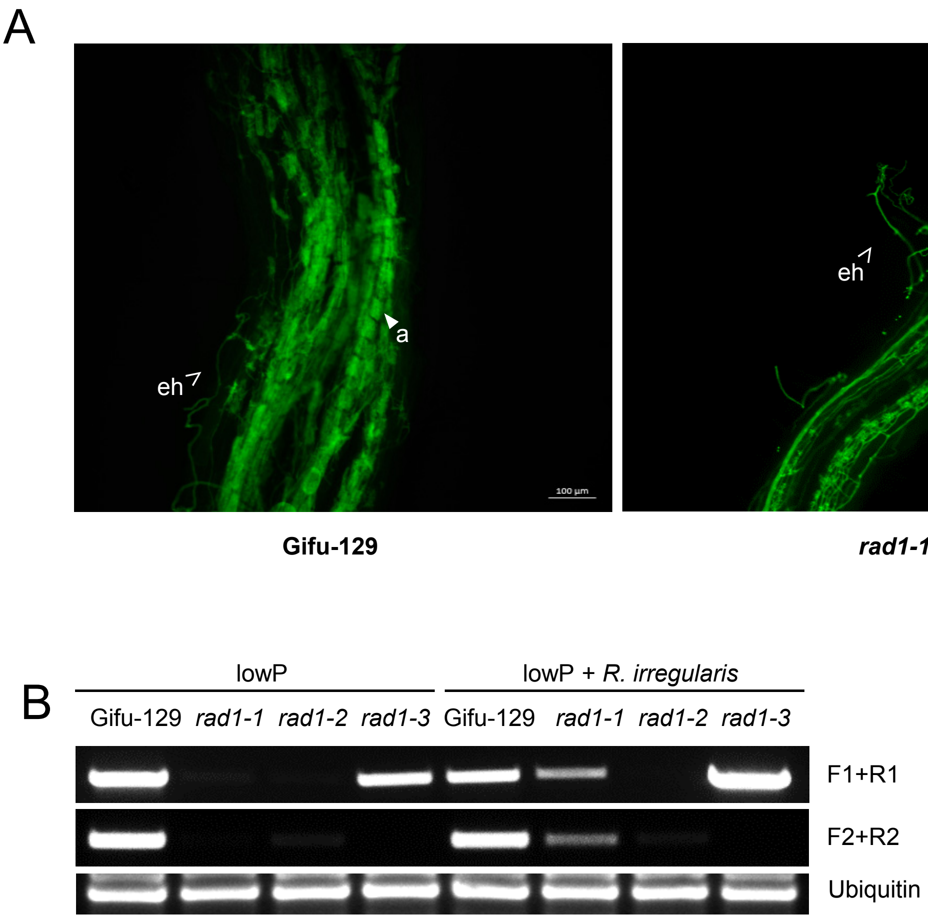
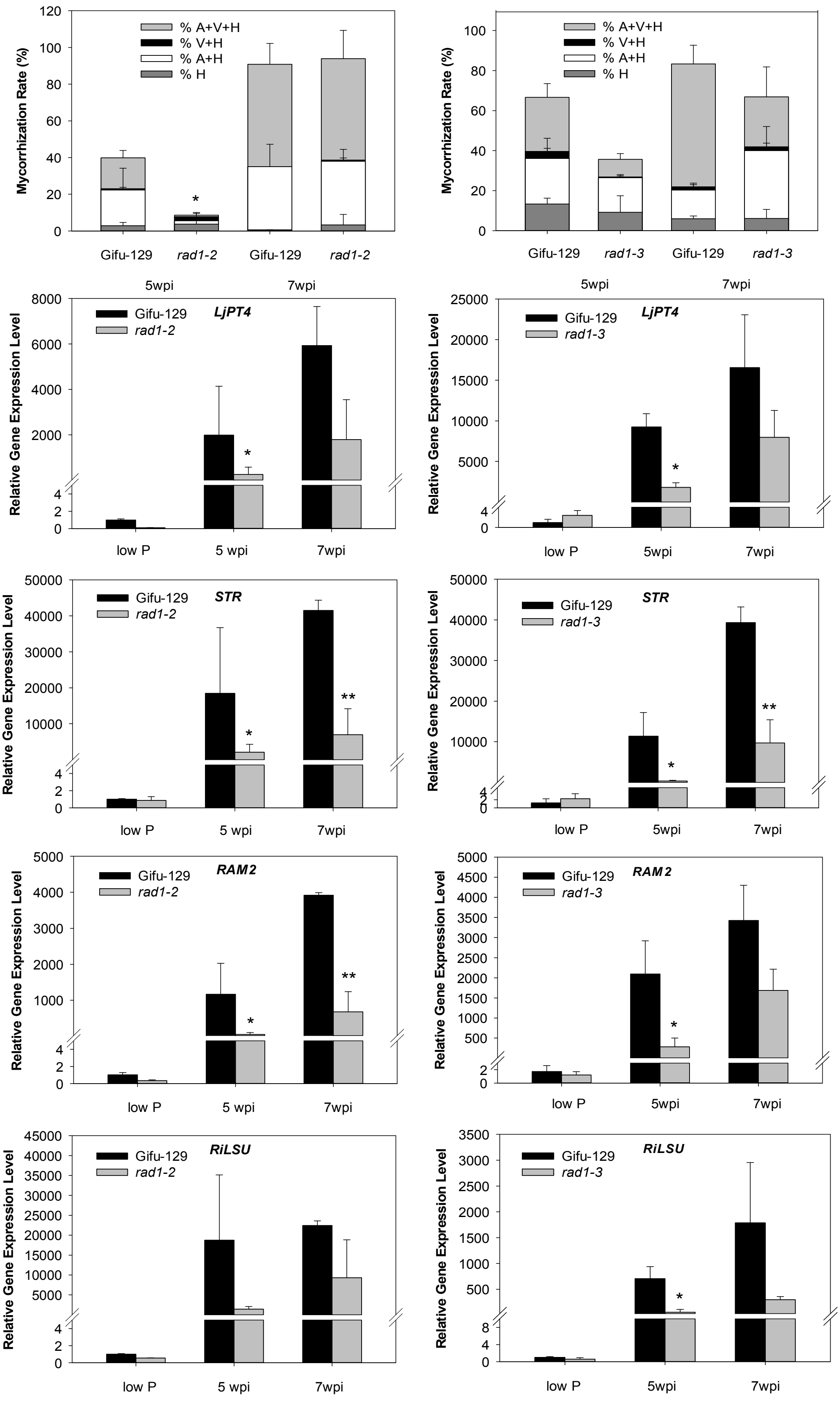


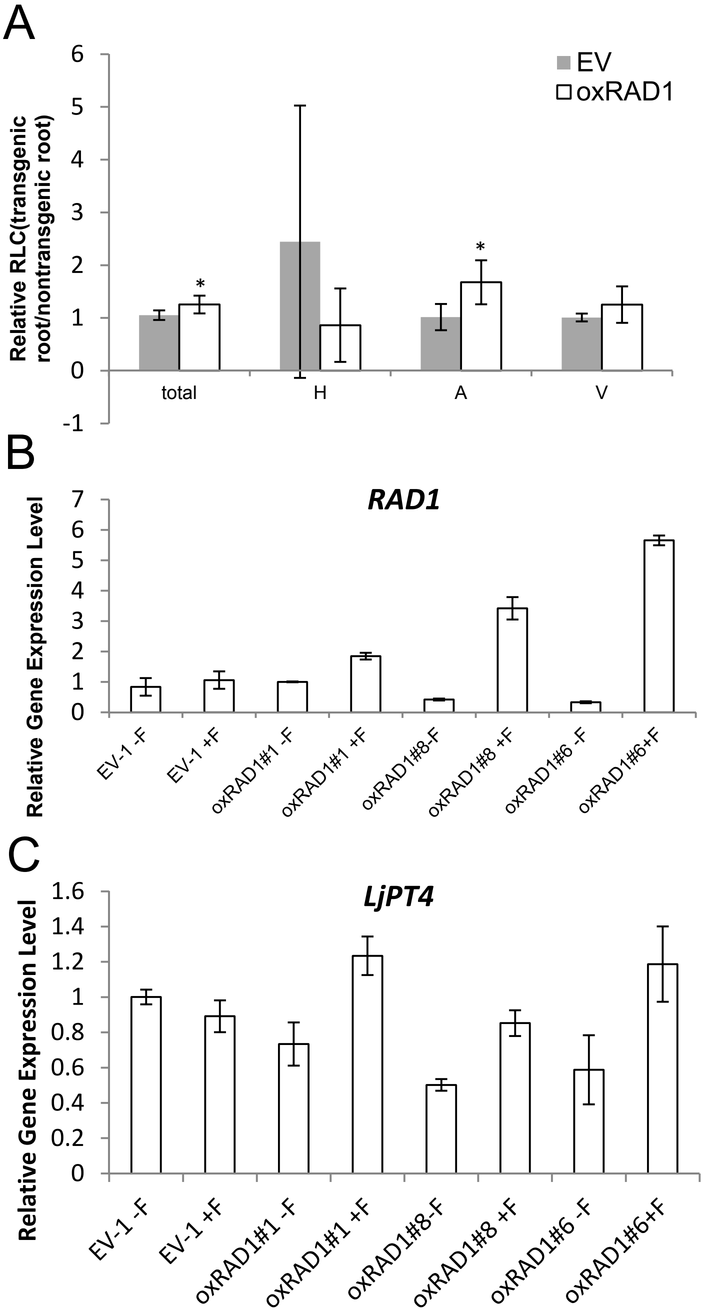
Supplemental Fig. S1. AM-inducible transcription factor genes repressed by high Pi condition. Gifu-129 plants grown at low Pi (5 µM) conditions were used as controls. After inoculation with *R. irregularis*, plants were grown at low Pi conditions for 8 weeks or 6 weeks at low Pi and re-supplied for two weeks with high Pi (7.5mM). Roots of four plants in one pot were harvested as one sample. Three technical replicates were used to calculate SD.



Supplemental Fig. S2. Mycorrhiza morphology in *rad1-1* and *RAD1* transcript analysis in *rad1 LORE1a* insertion lines.
(A) Arbuscules in Gifu-129 and *rad1-1* in the presence of *R. irregularis* at 8wpi. eh, external hypha. a, arbuscule. da, degenerated arbuscule. Scale bars are 100 µm.
(B) *RAD1* transcripts were detected in Gifu-129, *rad1-1, rad1-2* and *rad1-3* in the presence or absence of *R. irregularis* at 6wpi. 32 cycles were performed. Representative results from one out of three independent experiments are shown.



Supplemental Fig. S3. Mycorrhization rate and AM-inducible marker gene expression in *rad1-2* and *rad1-3* at 5 and 7 wpi .Three biological replicates were utilized. Error bars represent SD. Student's t-test was used to calculate confidence level. * indicates *P*<0.05. ** indicates *P*<0.01. Similar results was obtained from another independent experiment.

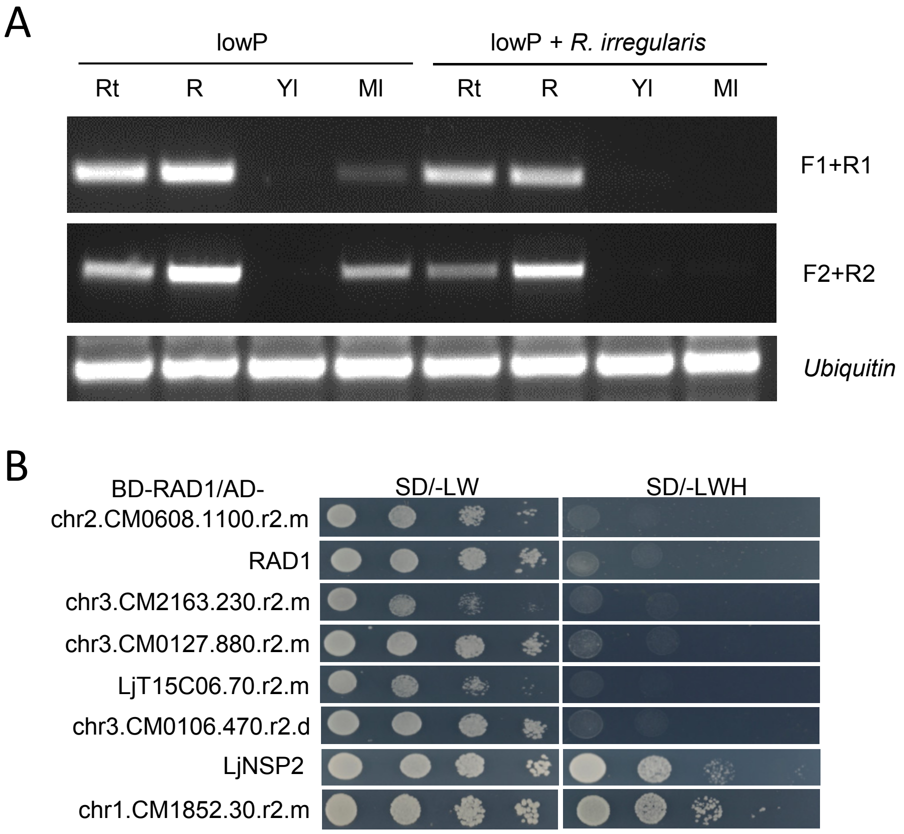


Supplemental Fig. S4. Mycorrhizal structures and *RAD1* gene expression in transgenic hairy roots overexpressing chimeric *35Spro:RAD1*.

(A) Relative root length colonized (RLC) in wild type hairy roots transformed with the empty vector (EV) or with a construct containing 35S*pro*:*RAD1* (oxRAD1). Percentage of RLC from transgenic hairy roots with EV (*n*=3) or oxRAD1 (*n*=8), respectively, was quantified. Relative RLC was calculated comparing the RLC from the transgenic red fluorescent hairy roots with the RLC from non-transgenic roots on the same composite plants. Significant differences of means between EV and oxRAD1 were determined by Student's t-test (* indicates $P<0.05$).

(B) *RAD1* gene expression is shown in transgenic red fluorescent hairy roots (+F) and non-transgenic non-fluorescent roots (-F) from three composite oxRAD1 plants and one control plant with EV in presence of *R. irregularis*. Error bars are from technical replicates.

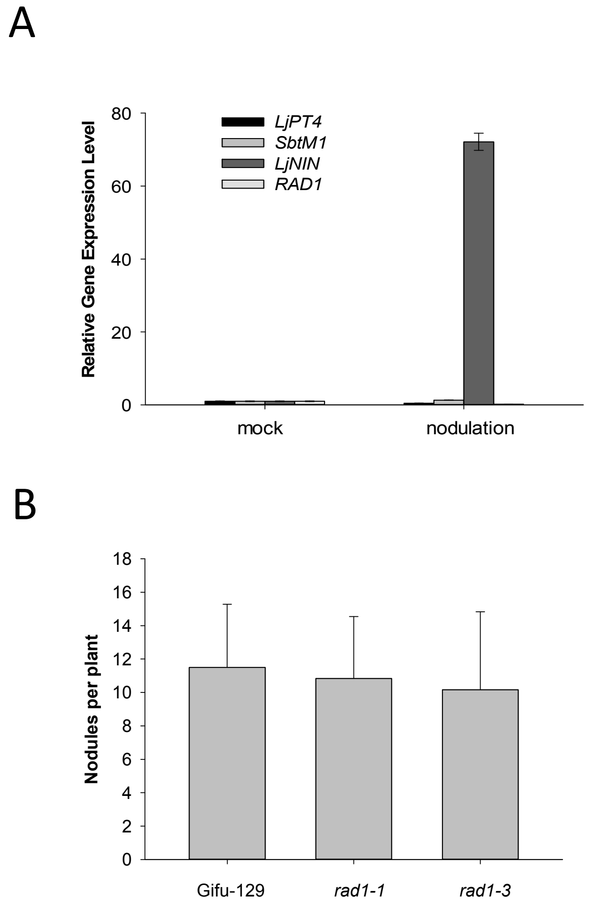
(C) *LjPT4* gene expression in transgenic fluorescent hairy roots (+F) and non-transgenic roots (-F) from three composite oxRAD1 plants and one control plant with EV in presence of *R. irregularis*. Error bars are from technical replicates.



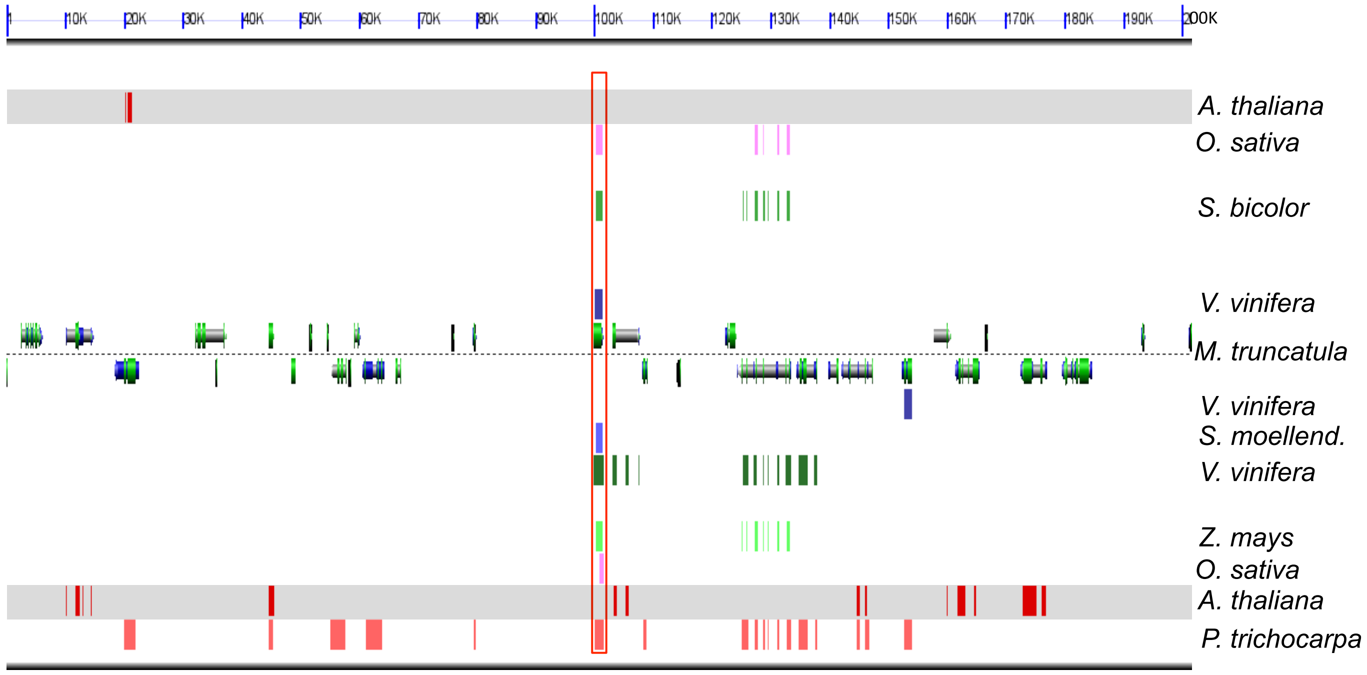
Supplemental Fig. S5. *RAD1* gene expression profile in mycorrhizal and non-mycorrhizal *Lotus* and GRAS proteins that interact with RAD1.

(A) RT-PCR analysis of *RAD1* transcript levels in different tissues of wild type in the presence or absence of *R. irregularis*. Rt, root tip; R, root ; Yl, young leaves; Ml, mature leaves. For *Ubiqutin* transcript amplification 28 PCR cycles, for *RAD1* 32 PCR cycles were used.

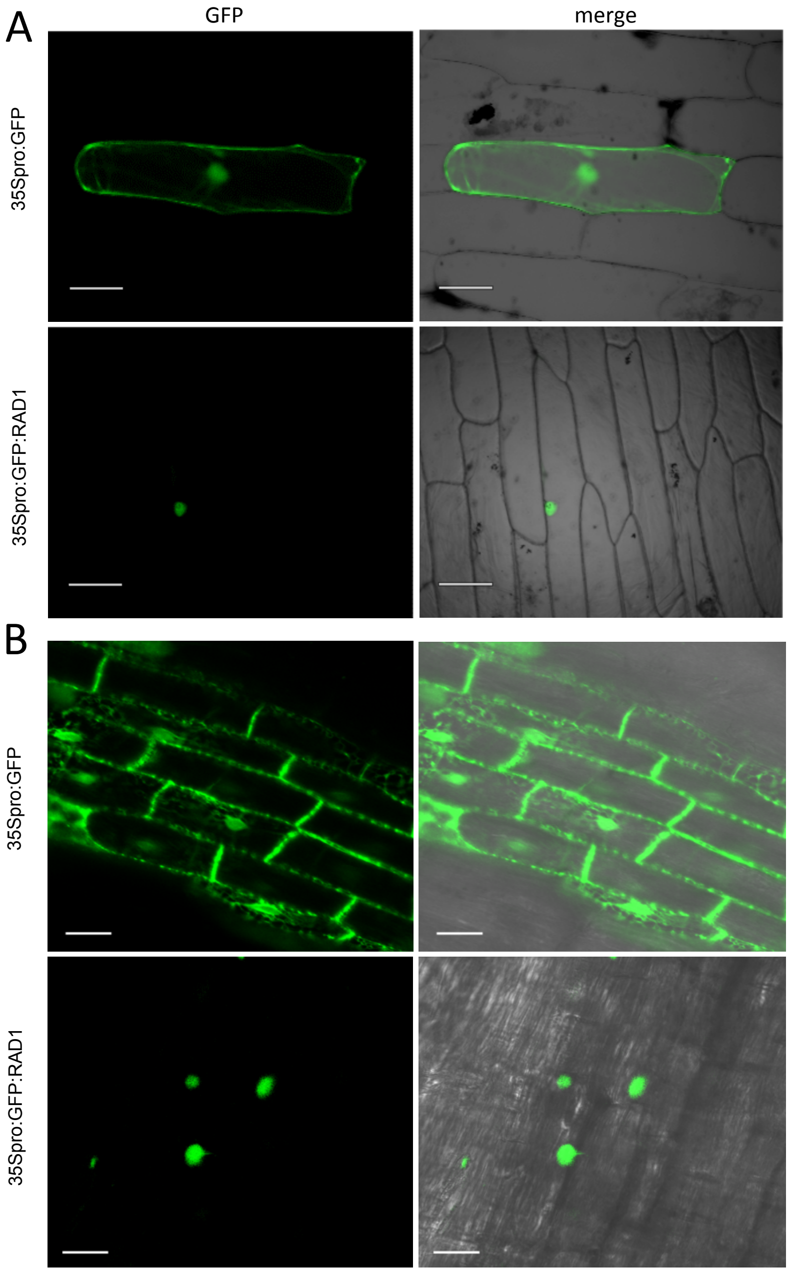
(B) Y2H analysis of interaction of RAD1 with mycorrhiza-regulated GRAS proteins. RAD1 was fused to the GAL4 DNA binding domain (BD). An AP2 transcription factor (chr2.CM0608.1100.r2.m) fused to the GAL4 activation domain (AD) was used as negative control. SD/-LW and SD/-LWH indicate a synthetic dropout medium lacking Leu and Trp and lacking Leu, Trp, and His, respectively.



Supplemental Fig. S6. *RAD1* gene expression in nodulated roots and nodulation in *rad1-1* and *rad1-3.*
(A) Relative expression levels of *LjPT4*, *SbtM1*, *LjNIN* and *RAD1* genes detected by qRT-PCR in nodulated plant roots.
(B) The number of nodules per plant in *rad1-1*, *rad1-3* and Gifu-129 is given (n=6). Two experiments were performed independently with similar results. Error bars represent SD.



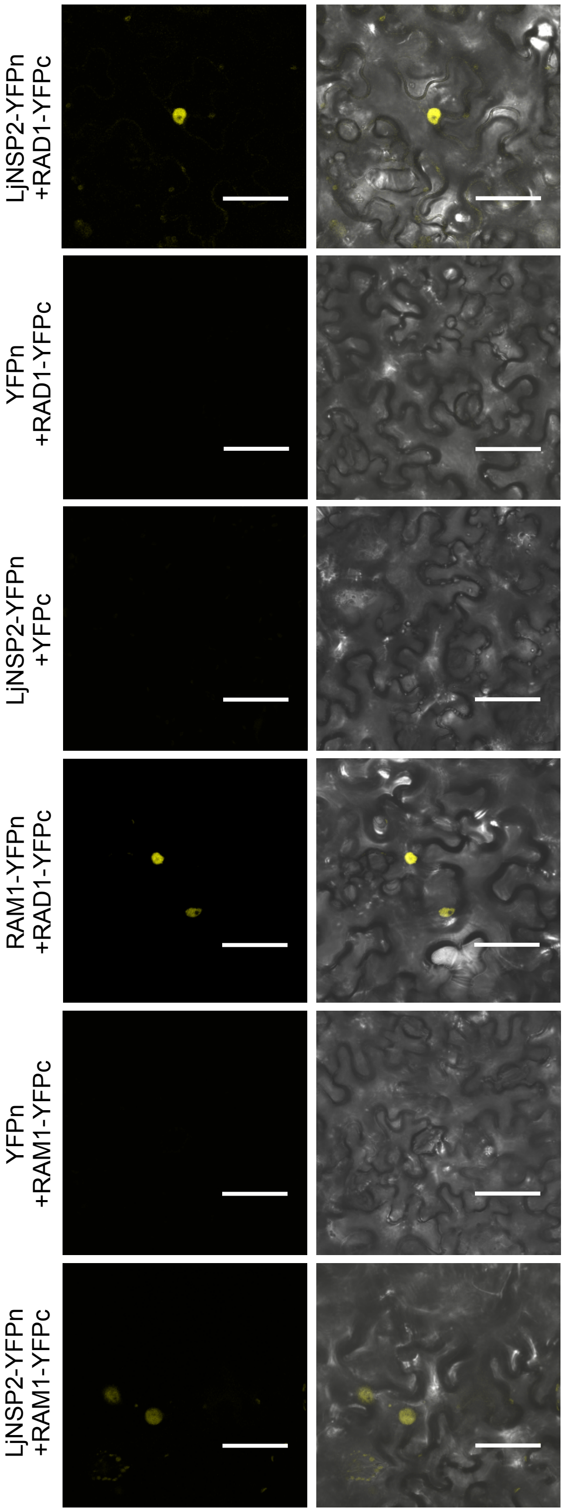
Supplemental Fig. S7. Synteny analysis of *RAD1* on AM host species and non-host *Arabidopsis thaliana*. Synteny analysis of a ~200 kb region encompassing the *RAD1* locus in *M. truncatula*. Green and blue boxes above and below the dashed lines represent *Medicago* genes. Orthologous genes in other species are indicated above and below: *Arabidopsis thaliana* (red), *Populus trichocarpa* (orange), *Vitis vinifera* (dark blue – RAD1-like; dark-green RAD1), *Sorghum bicolor* (green), *Oryza sativa* (pink), *Zea mays* (light green) and *Selaginella moellendorffii* (violet). The RAD1 locus is red-boxed.



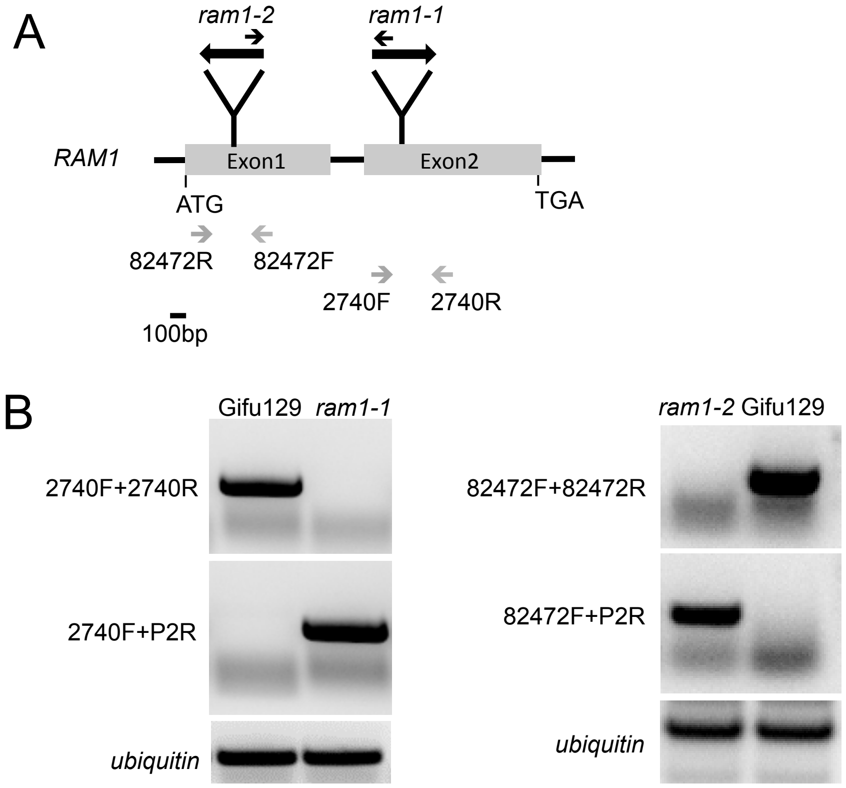
Supplemental Fig. S8. Nuclear localization of RAD1.

(A) Nuclear localization of RAD1 in onion epidermal cells after biolistic bombardment with *35Spro:GFP* or *35Spro:GFP:RAD1*, respectively. Images with green fluorescence (GFP) alone and merged with bright-field image (merge) are shown as indicated. Scale bar is 100 µm.

(B) Nuclear localization of RAD1 in *Lotus* hairy roots expressing *35Spro:GFP* or *35Spro:GFP:RAD1*, respectively. Scale bar is 20 µm.



Supplemental Fig. S9. Interaction between RAM1, RAD1 and LjNSP2 by BiFC assay in *N. benthamiana*. Leaves were infiltrated with mixtures of *A. tumefaciens* strain GV3101 to coexpress RAD1 and RAM1 or LjNSP2 fused to half of split YFP, respectively as indicated at left. Three days after infiltration, images were captured from five different leaves for each plasmid combination. Scale bars are 50 µm.



Supplemental Fig. S10. Gene structure of *RAM1* in wild type *Lotus* and *ram1* mutant.
(A) Gene structure of *RAM1* in *Lotus* and the positions of the *LORE1a* insertions in *ram1-1* and *ram1-2*. Gray arrows indicate the primers matching with the genome sequence used for genotyping. Black arrows indicate P2R primer on the left border of the *LORE1a* transposon.
(B) Identification of homozygous *ram1-1* and *ram1-2* through genomic DNA amplification. Primer pairs as described in (A) were used for PCR and are indicated at left. *Ubiquitin* gene was used as a control.

Supplemental Table S1: List of mycorrhiza up-regulated transcription factors generated by DESeq.

Seq. Name (*Lotus*CDSv2.5)	**Putative annotation of transcription factor**	**RNA-seq data**			
		fold change (+*R.i.* /-*R.i.*)	***P*adj**	**resVarA**	**resVarB**
chr6.CM0679.550.r2.d	MADS-box	M specific	7.0e-04	6.09	0.00
chr6.CM1613.300.r2.m	C2H2 Zinc finger	M specific	5.1e-03	0.24	0.00
chr1.LjB18K24.70.r2.a	MYB	M specific	1.42e-102	17.63	0.00
chr3.CM0111.130.r2.d	NAC	M specific	5.10e-17	0.82	0.00
chr3.CM2163.240.r2.m	GRAS	970.83	6.73e-89	8.54	5.1e-04
chr2.CM0608.1100.r2.m	AP2	586.44	1.18e-06	12.14	1.3e-03
chr3.CM2163.230.r2.m	GRAS	462.42	3.11e-111	12.07	8.84e-06
chr1.CM1852.30.r2.m	GRAS	438.56	8.73e-109	2.33	9.84e-06
chr5.CM0239.240.r2.m	GRAS	170.03	1.89e-40	0.64	0.01
chr2.CM0021.530.r2.m	C2H2 Zinc finger	165.30	5.66e-65	0.18	6.40e-05
chr4.CM0075.50.r2.m	ERF	164.31	3.43e-23	4.27e-04	0.01
chr2.CM0081.1990.r2.m	AP2	150.70	5.7e-69	10.70	3.88e-04
chr1.LjB18K24.100.r2.a	MYB	120.88	1.56e-80	15.0	2.46e-03
chr6.CM0041.30.r2.a	AP2	78.55	4.72e-62	0.08	1.72e-04
LjB17L21.50.r2.d	MADS-box	44.10	7.51e-07	2.13e-03	0.09
chr1.CM0029.1590.r2.a	ARF	43.89	5.02e-48	0.1	8.24e-04
chr1.CM1413.480.r2.d	WRKY	38.12	5.42e-26	0.49	0.08
LjSGA_032180.1	C2H2 Zinc finger	24.65	2.11e-30	0.76	0.02
chr5.CM1667.220.r2.a	GRAS	19.68	6.69e-15	0.63	0.28
LjSGA_122341.1	GRAS	13.04	9.07e-19	1.46	0.14
chr6.CM0139.1440.r2.d	GRAS	11.74	4.35e-35	0.74	4.56e-05
chr2.CM1835.10.r2.m	NAC	8.96	3.65e-09	1.05	1.5e-03
LjSGA_020219.1	GRAS	8.93	5.03e-03	0.60	0.10
LjSGA_012098.1	AP2	6.80	3.26e-09	0.04	0.25
chr4.CM0680.320.r2.m	GRAS	6.02	9.55e-17	11.47	0.10
LjSGA_029723.1	AP2	6.00	8.98e-14	1.73	0.10
chr1.CM0012.840.r2. m	NAC	6.04	2.71e-23	4.12	0.70
chr3.CM0127.880.r2.m	GRAS	5.77	6.12e-13	6.14	0.01

LjSGA_045015.1	GRAS	5.74	2.84e-22	4.13	3.60e-03
LjSGA_073109.1	GRAS	5.34	1.93e-22	5.29	0.15
LjSGA_026563.1	MYB	5.31	4.30e-03	0.95	0.01
chr6.CM0118.1050.r2.a.	ERF	4.98	6.92e-19	0.29	0.13
chr1.cm0122.2130.r2.d	GRAS	4.81	2.83e-13	1.06	0.22
chr3.CM0106.470.r2.d	GRAS	3.57	1.50e-13	2.51	0.12
chr1.CM0375.530.r2.a	C2H2 Zinc finger	3.54	9.86e-12	1.44e-03	1.05e-03
chr2.CM0018.140.r2.m	GRAS	3.20	8.07e-10	0.01	0.71
LjSGA_055804.0.1	C2H2 Zinc finger	3.17	9.21e-04	0.01	0.07
chr3.CM0106.760.r2.m	NIN-like	3.13	7.69e-08	1.24	1.93
chr4.CM1864.540.r2.m	GRAS	2.93	2.01e-10	40.77	1.24
LjT15C06.70.r2.m	GRAS	2.51	4.23e-04	11.13	1.06e-03
chr5.CM0200.2670.r2.d	LOB domain	2.46	6.97e-06	0.21	0.06
chr1.CM0284.80.r2.d	AP2-ERF	2.31	1.60e-03	1.21	2.98e-03
chr6.CM0314.840.r2.d	bZIP	2.30	3.35e-03	0.46	2.86e-03
chr3.CM0176.150.r2.m	NAC	2.27	3.43e-03	0.40	5.15e-05
chr3.CM0460.10.r2.d	GRAS	2.11	2.85e-04	1.50	0.19

"resVar" values indicate the degree of the variance among the biological samples. If resVarA or resVarB are <15, the fold changes are robust. “M specific” indicates mycorrhiza specific induction of gene expression.

Supplemental Table S2: Fourty-five mycorrhiza up-regulated transcription factors under different treatments.

Overall highly expressed genes					
Gene ID	Low P+M	High P+M	shift	Low P	High P
chr1.CM0375.530.r2.a	292.749139	65.0446598	61.1996559	82.7275152	63.6899635
chr3.CM0176.150.r2.m	91.522928	38.2383685	53.9730872	40.394509	26.0924937
chr1.CM0012.840.r2.m	349.15635	32.9624048	39.6685867	57.8409834	34.5023455
chr5.CM0200.2670.r2.d	231.955327	107.937444	170.9069145	94.3299251	179.063813
chr3.CM0460.10.r2.d	244.052245	42.2776687	73.5944351	115.4822037	79.2501851
chr2.CM0018.140.r2.m	255.58247	48.2073169	51.6239006	79.9909595	56.886244
chr3.CM0106.470.r2.d	491.400977	70.7606984	115.6729524	137.4609898	97.676021
chr1.CM0122.2130.r2.d	165.672537	30.088316	31.2386641	34.408076	25.3193636
chr6.CM0139.1440.r2.d	319.955846	31.5771559	33.8243808	27.2527556	26.925496
chr6.CM0314.840.r2.d	83.713264	51.1785688	48.5987186	36.375022	52.2947531
chr3.CM0106.760.r2.m	183.9800069	65.5690423	74.88890491	58.78141048	62.03393774

Genes which were only responsive to AM under low P conditions

Gene ID	Low P+M	High P+M	shift	Low P	High P
chr2.CM0021.530.r2.m	301.577336	0.3172006	3.7890705	1.8244533	1.2694607
LjSGA_032180.1	144.082462	6.8876778	9.0619715	5.8439403	7.6067857
chr1.CM1413.480.r2.d	99.969242	6.2532766	4.9788498	2.6226361	9.6793125
chr2.CM1835.10.r2.m	44.693256	8.3700899	15.0244424	4.9887356	5.4743866
chr3.CM0111.130.r2.d	43.419953	0.5308103	0	0	0
LjSGA_026563.1	16.807706	6.9976965	2.8933424	3.1642823	7.1603485
chr1.LjB18K24.100.r2.a	485.867106	3.812835	5.8745779	4.019487	2.9554226
chr1.LjB18K24.70.r2.a	601.412944	0	2.5113982	0	0.4264798
LjB17L21.50.r2.d	18.859199	0	0.4258908	0.4276024	0.4264798
chr6.CM0679.550.r2.d	8.276469	0.3172006	0.8517816	0	0
chr3.CM0127.880.r2.m	117.540714	12.0600506	19.4895084	20.3825447	12.1882979
chr5.CM1667.220.r2.a	56.110791	1.4824121	1.2337258	2.8507238	1.2694607
chr5.CM0239.240.r2.m	145.41238	2.226832	0.8517816	0.8552047	0
chr1.CM1852.30.r2.m	800.135821	1.7996127	2.5113982	1.8244533	1.2694607
chr3.CM2163.230.r2.m	843.667924	1.1652115	0.8517816	1.8244533	0.4165011
chr3.CM2163.240.r2.m	470.488665	0.3172006	3.3631797	0.4846243	1.259482
chr1.CM0029.1590.r2.a	240.24622	5.5152845	5.4047406	5.4733599	4.6513631
LjSGA_029723.1	123.780115	11.8592965	14.5985516	20.6106325	10.0958136
LjSGA_012098.1	56.634212	4.7644368	5.0227963	8.3240837	6.7737834
chr6.CM0041.30.r2.a	320.224498	2.7512145	3.3192332	4.0765089	6.3173675
chr2.CM0081.1990.r2.m	339.365693	3.5992253	4.6408521	2.2520557	2.9554226
chr2.CM0608.1100.r2.m	284.202937	0	0	0.4846243	0
chr4.CM0075.50.r2.m	70.258661	0	0	0.4276024	0.4264798

Low Pi induced genes regardless of AM status

Gene ID	Low P+M	High P+M	shift	Low P	High P
LjT15C06.70.r2.m	103.902922	2.9712519	5.1106894	41.3637576	6.3173675
chr4.CM1864.540.r2.m	1502.546805	18.216164	41.9738268	513.4417307	32.7066181
LjSGA_073109.1	475.682844	23.3235125	25.1443536	88.3434921	16.8496397
LjSGA_045015.1	380.155738	14.3125937	10.0016461	66.2504137	14.2807822
chr4.CM0680.320.r2.m	176.496151	5.295247	0.807835	29.3052966	3.3719237
LjSGA_020219.1	376.56175	5.6124476	13.3208793	42.1905135	8.4198305
LjSGA_122341.1	96.630563	0	0	7.411857	0.4264798
chr1.CM0284.80.r2.d	103.803575	7.1012875	6.9764641	45.0126642	4.2248833
chr6.CM0118.1050.r2.a	343.723987	19.4006588	24.5123047	69.0726886	23.563551

Intermediates

Gene ID	Low P+M	High P+M	shift	Low P	High P
chr6.CM1613.300.r2.m	6.53636	1.6960217	0.4258908	0	0
LjSGA_055804.0.1	44.863102	16.0993508	16.0823821	14.168024	15.1736565

Normalized counts per transcript from different treatments are shown above.

Supplemental Table S3: Sequences of oligonucleotide primer pairs used.

Primers for RT-PCR:

Gene ID	**primer name and primer sequence**
Lj*NIN*	qLjNIN-F: TGGATCAGCTAGCATGGAAT
	qLjNIN-R: TCTGCTTCTGCTGTTGTCAC
SbtM1	qSbtM1-F: TGTATGCTGCTGCTGAAAAAAACAACT
	qSbtM1-R: CTTCTTGACCTTTTGCAATAAATGGGATTC
STR	qSTR-F: CTGGACAAGATCACCGTCCT
	qSTR-R: GTGGCCATCAAGCTGGTATT
RAD1	q-F1: CCGAGGCTCATGCCTAGGTCCACT
	q-R1: CCCCAATGGGTTTCCATGCCTATCC
	q-F2: ATGGTGGAGCAGGATTCAAG
	q-R2: TTAACATTTCCAGCAAGAAG
RAM1	qRAM1-F: GGAGGTTTCTTGAGGCACTG
	qRAM1-R: CCTTCCATGATCTTCCTCCA
RAM2	qRAM2-F: GGGATGGACCCGTTTTACTT
	qRAM2-R: GACTTTGTTTGCGGCCTTAG
Lj*PT4*	qLjPT4-F: TCCAAGCGGAGCAAGACAAG
	qLjPT4-R: TTCTGTGTGAGGTTCTGGCTGTAG
Ubiquitin	qUbiquitin-F: TTCACCTTGTGCTCCGTCTTC
	qUbiquitin-R: AACAACAGCACACACAGACAATCC
R.irregularis large subunit rRNA (RiLSU)	LR1 : GCATATCAATAAGCGGAGGA
	8.22: AACTCCTCACGCTCCACAGA
chr3.CM0106.760.r2.m	q30-F: AGGATGCTGCAAAGAGCATT
	q30-R: CTCCCTGGACAGAGTCAAGC
chr3.CM0127.880.r2.m	q21-F: GAAGCCATGGGACAAGAAAA
	q21-R: TCCAAAGTGCTCACAACAGC

chr2.CM0608.1100.r2.m	q3-F: AGAGGAGTAGCAAGGCACCA
	q3-R: TGGGGTTTGCTCTTGAGAAG
chr3.CM2163.230.r2.m	q-8-F: TGCCTGTGAAGACAGGTGAA
	q-8-R: GCCACTCAGCAAAAGTCCTC
chr3.CM2163.240.r2.m	q-5-F: TCTTCACCCAAGTGGAAGG
	q-5-R: CCAGCATGATTCTCTCCACA
chr2.CM1835.10.r2.m	q17F: GCAACTGATTCTTGGGCAAT
	q17R: GAATGGCTGATGGTGTTGTG
chr3.CM0106.470.r2.d	q26-F: TCCTCACTTGAAACTCACTG
	q26-R: CATTCGTTGACCAACCTCCT

Primers for genotyping:

LORE1a mutant	**primer name and primer sequence**
rad1-1	1039F: GGAATGCTATTCACACTCACACTC
	1039R: GAACGAGGAACTCAACACATC
rad1-2	576F: CCGAGGCTCATGCCTAGGTCCACT
	576R: CCCCAATGGGTTTCCATGCCTATCC
rad1-3	2260F: AGTCGTGGCGCTTTGAATTCGGTG
	2260R: TGAGTGGTTTCGAATCACCAGAGCG
ram1-1	2740F: GCTTGTGCTGAAGCAGTGGCCAAA
	2740R: TGCGGAGGGAGTGTGCTAATTCGG
ram1-2	82472F: TCAATTGCTGGCAGGGCAAGGTTC
	82472R: TGCAGTCCAAGAAAAATGCCACTCA
Ubiquitin	GS-UB-F:CGTGAAGGCTAAGATCCAGGATAAG
	GS-UB-R:CGATACTACTTGTTCAAGAGGGGC
	P2R: CCATGGCGGTTCCGTGAATCTTAGG

Note: Endogenous fragment: F+R, LORE1a insertion fragment: F+P2R

Primers for Y2H:

plasmids	**primer name and primer sequence**
pAS-RAD1/ pACT-RAD1	RAD1-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGTCCCCTCCTCTTTATAGTG
	RAD1-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAACATTTCCAGCAAGAAGCTG
pACT-RAM1	RAM1-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGATCAATTCAATGTGTGG
	RAM1-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGCATCTCCATGCAGAGG
pACT-AP2	AP2-attB1:

	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGAGTTTGCTTCTGTAAAATC
	AP2-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATTCCATAGGGGGAAACATG
pACT-Chr3.CM2163.230.r2.m	Chr3.CM2163.230.r2.m-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGAGGGAGCTGAGATATGAC
	Chr3.CM2163.230.r2.m-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATTTCCTGAAACTCCATGCTG
pACT-Chr3.CM0127.880.r2	Chr3.CM0127.880.r2.m-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGAAGACAAGGGCTTAAAAC
	Chr3.CM0127.880.r2.m-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAAACTTCCAGGCTGATATAG
pACT-cGRAS	Chr3.CM0106.470.r2.d-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGACATTACTCCTTATAC
	Chr3.CM0106.470.r2.d-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGTAGCTAGCTGGGGCT
pACT-LjT15C06.70.r2.m	LjT15C06.70.r2.m-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGACATGGAGATTGACATTG
	LjT15C06.70.r2.m-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCCTACAAATTAGAGTCATCTG
pACT-LjNSP2	LjNSP2-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGAAATGGATATAGATTG
	LjNSP2-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATGCACAATCTGATTCTG
pAS- RAD1ΔN	RAD1-121-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTatgGTTGGAGCTGAAGAAGATG
	RAD1-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAACATTTCCAGCAAGAAGCTG
pAS- RAD1ΔPS	RAD1-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGTCCCCTCCTCTTTATAGTG

	RAD1-340-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAAGAGCTTCATCATCATTC
pAS- RAD1 ΔS	RAD1-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGTCCCCTCCTCTTTATAGTG
	RAD1-420-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACATCTTTGCCCTCTTAGTG
pAS- RAD1 ΔNL	RAD1-200-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTatgTCCATGATGAACATCATG
	RAD1-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAACATTTCCAGCAAGAAGCTG

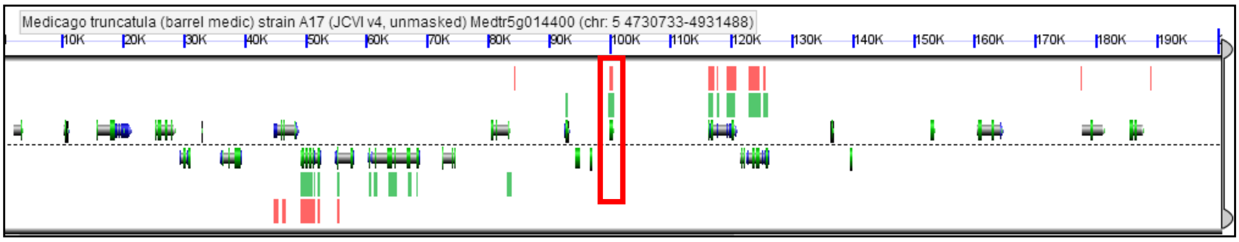
Primers for BiFC:

pBatTL-RAD1-YFPn/ pBatTL-RAD1-YFPc	RAD1-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGTCCCCTCCTCTTTATAGTG
	RAD1withoutTGA-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCacatttccagcaagaagctg
pBatTL-RAM1-YFPn/ pBatTL-RAM1-YFPc	RAM1-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGATCAATTCAATGTGTGG
	RAM1withoutTGA-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCGCATCTCCATGCAGAGG
pBatTL-LjNSP2-YFPn	LjNSP2-attB1: GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGAAATGGATATAGATTG
	LjNSP2withoutTGA-attB2: GGGGACCACTTTGTACAAGAAAGCTGGGTCTGCACAATCTGATTCTG

Supplemental Text S1. Synteny analysis of potential symbiotic transcription factors. A ~200kb sized region in the *Medicago truncatula* genome containing transcription factors identified by RNAseq was compared to the syntenic region in *Arabidopsis thaliana* (col-0, pink) and *Populus trichocarpa* (green) using CoGe:GEvo (https://genomevolution.org/CoGe/GEvo.pl) as described in Delaux et al. 2014. *Medicago* was used as a proxy for *Lotus* given that this tool is not available for *Lotus*. The target gene in *Medicago* is red-boxed together with orthologs from *Populus* and *Arabidopsis* when present. The vertical orange stripes are regions that contain unknown nucleotides that were inserted during the scaffolding of assembled contigs. However, the position and relative size of the gap (orange stripe) is known due to the size of the genomic library and/or a genomic map.

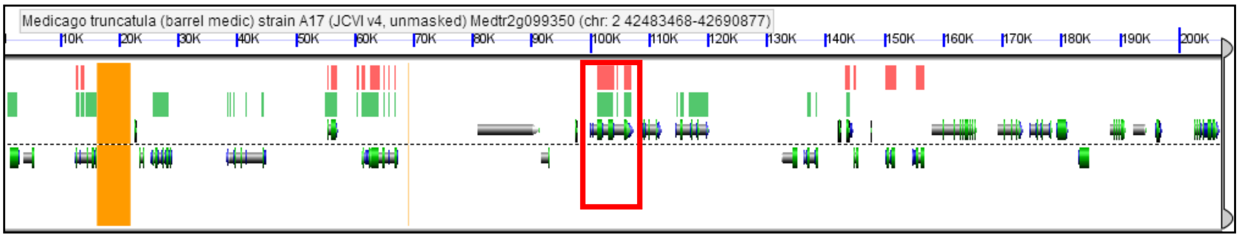
Examples of genes conserved in host and non-host species

chr2.CM0021.530.r2.m



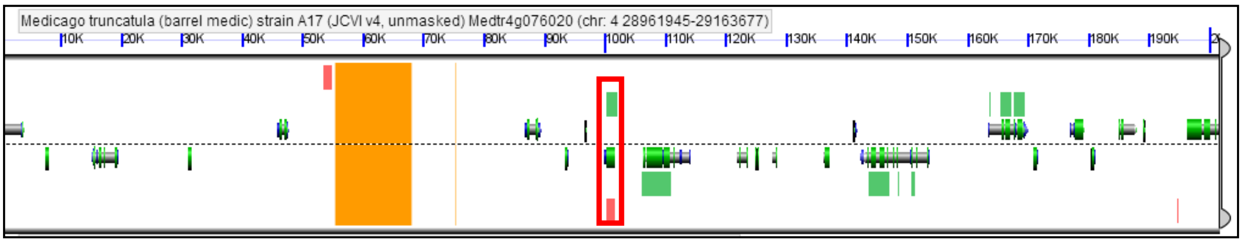
Strong syntenic support for both *Arabidopsis* and *Populus*.

chr3.CM0106.760.r2.m



Strong syntenic support for both *Arabidopsis* and *Populus*.

chr3.CM0460.10.r2.d

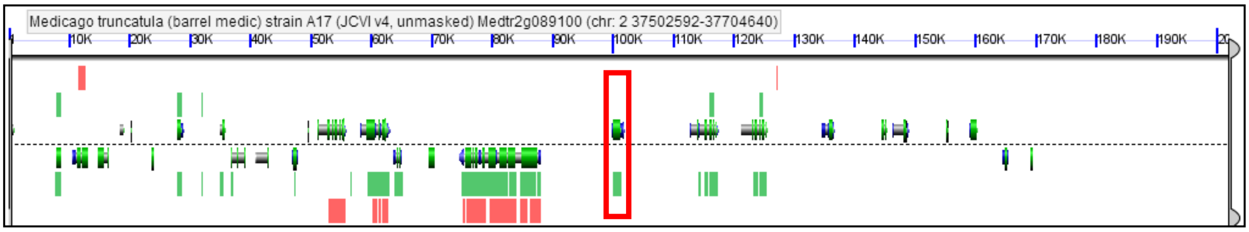


Strong syntenic support for *Populus* but low support for *Arabidopsis*.

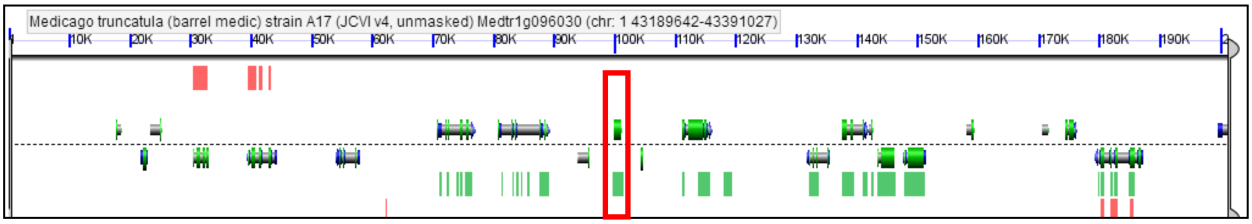
Genes missing in in non-host species

Present in Lupinus (potential common symbioses transcription factors)

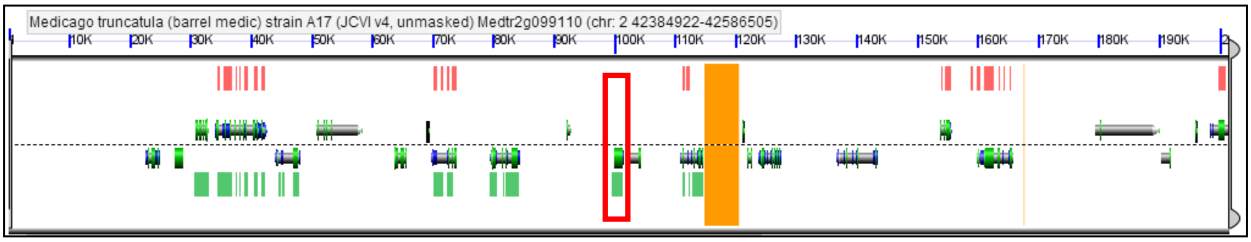
LjSGA_045015.1



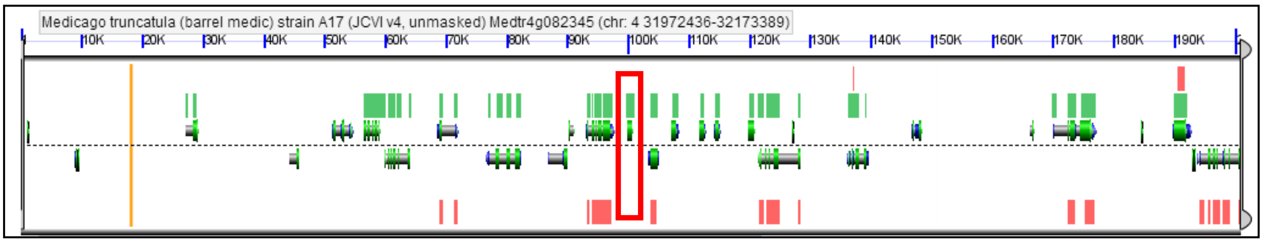
chr5.CM1667.220.r2.a



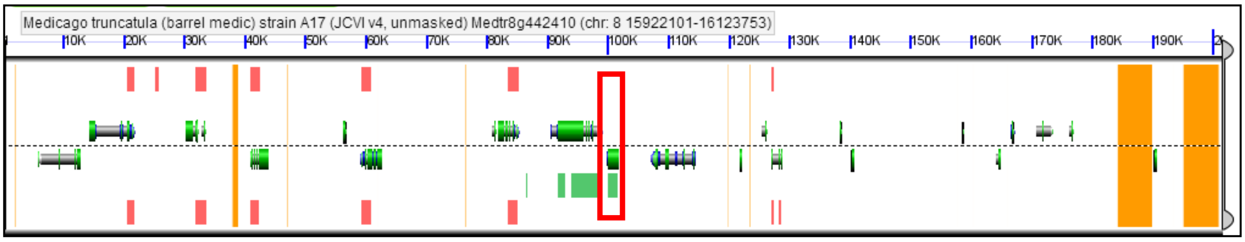
chr3.CM0106.470.r2.d



chr4.CM0075.50.r2.m

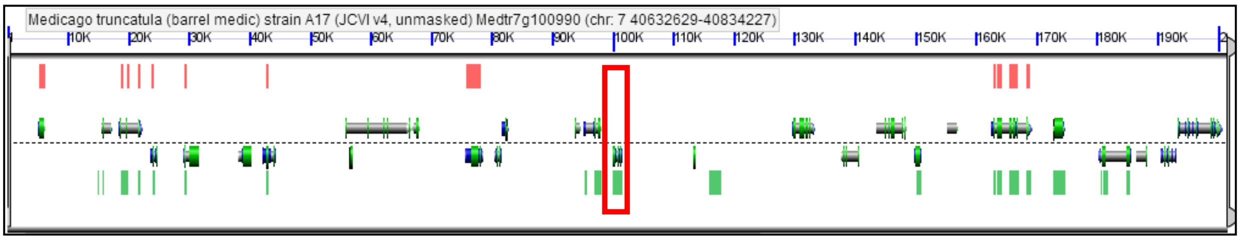


LjSGA_122341.1

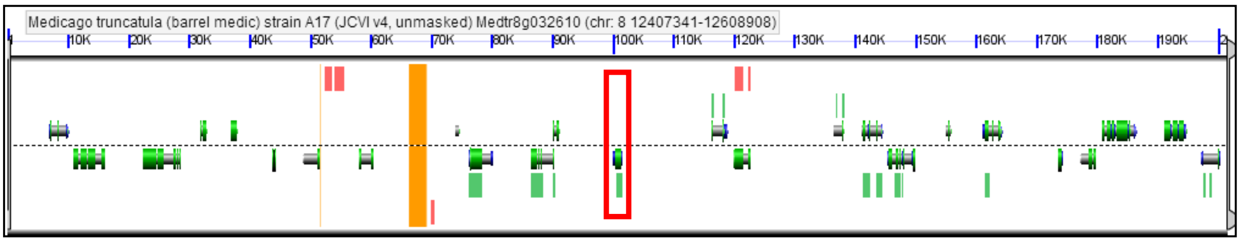


__Absent in Lupinus (potential AMS specific transcription factors)__

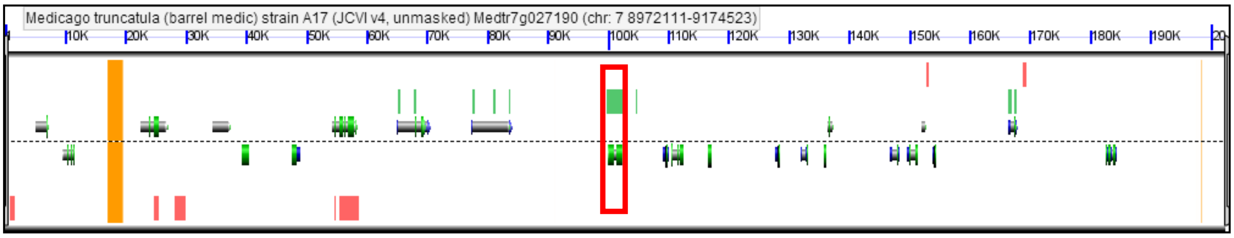
chr1.CM0012.840.r2.m



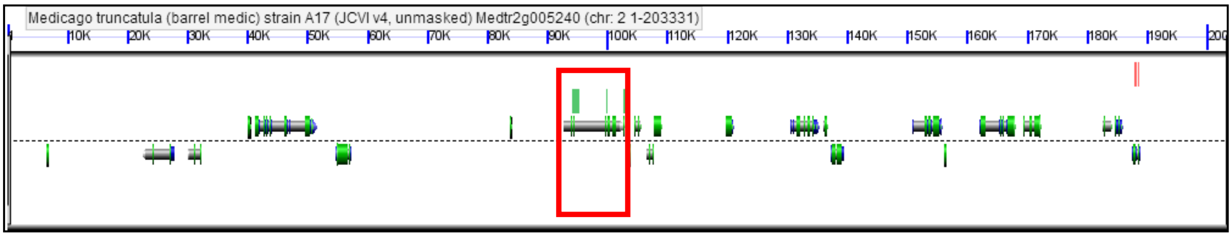
chr1.CM0284.80.r2.d



chr1.CM1852.30.r2.m

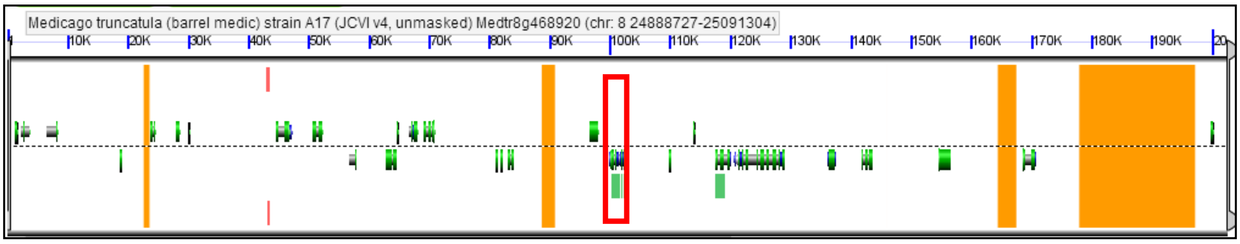


chr1.CM0029.1590.r2.a

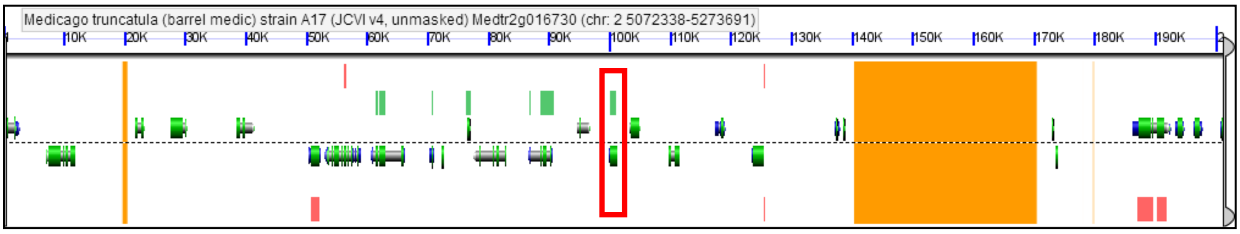


Low syntenic support even for *Populus.*

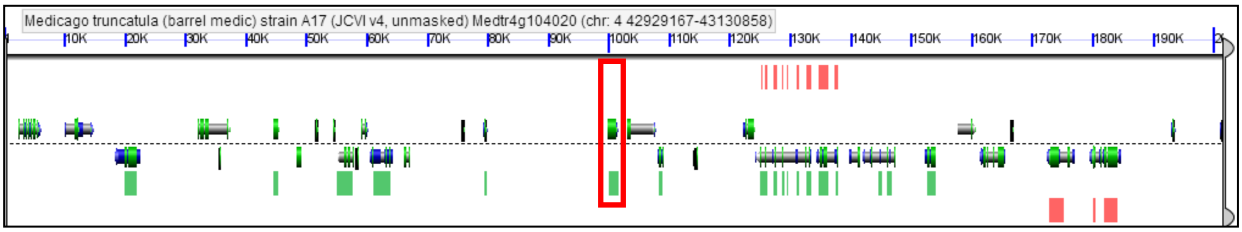
chr2.CM0608.1100.r2.m



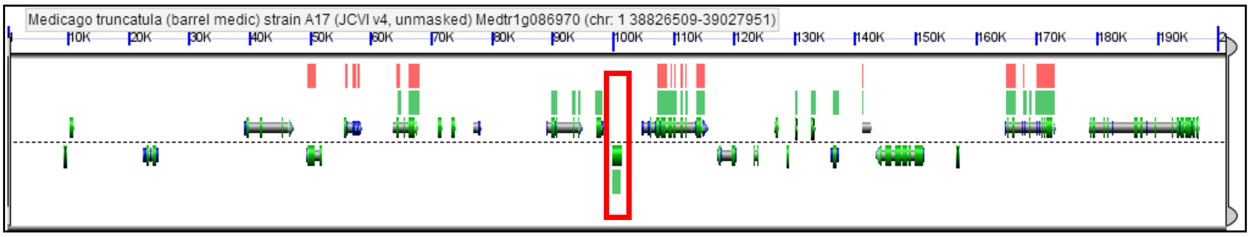
chr6.CM0118.1050.r2.a



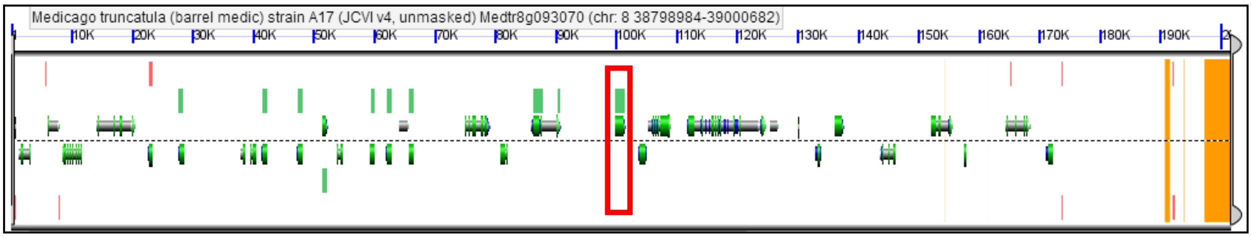
chr4.CM1864.540.r2.m



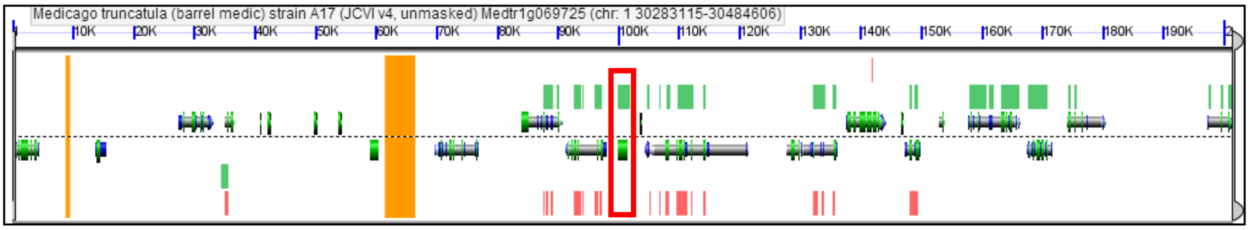
chr5.CM0239.240.r2.m



chr4.CM0680.320.r2.m



LjSGA_020219.1



Supplemental Text S2: Sequences used for phylogenetic tree.

Aquilegia coerulea

>Aco Aquca_123_00013
MISDTFHDEKIEGVRGLDLSLSAMAFYPFPYLSTLENSVSTWFLPSTDDTRNH
KRLKPTLSSDELIGSNNSPYSGGSNNSLFSGGSCITRANSTNSLNSLPKLQFRD
HIWTYTQRFLAAEAVEEAAAAMIGEDGERVDEGSGDGMRLVQLLVSCAEAV
ACRDRSHASTLLSELRANALVFGTSFQRVASCFVQGLADRLALVQPLGAVGV
VAPTMNFMACSSEKKEEALRLVYDVCPHIQFGHFVANTSILEAFEGESLVHV
VDLGMTMGLPHGHQWRRLIHSLANRAGQPPRRLRITGVGNSGDQLQTIGDEL
EAYAQSLGINFDILQLHCVVKESRGALNSVLQIIHELSPKVLVLVEQDSSHNGP
FFLGRFMEALHYYSAIFDSLDAMLPKYDTKRAKIEQFYFAEEIKNIVSCEGPA

RVERHERVDQWRRRMSRAGFQPVPIKMLAQAKQWLGKVKVCDGYTIVEEK
GCLVLGWKSKPLVAASCWKC
Arabidopsis thaliana
>AtSCL1
MVEQTVVREHIKARVMSLVRSAEPSSYRNPKLYTLNENGNNNGVSSAQIFDP
DRSKNPCLTDDSYPSQSYEKYFLDSPTDEFVQHPIGSGASVSSFGSLDSFPYQS
RPVLGCSMEFQLPLDSTSTSSTRLLGDYQAVSYSPSMDVVEEFDDEQMRSKIQ
ELERALLGDEDDKMVGIDNLMEIDSEWSYQNESEQHQDSPKESSSADSNSHV
SSKEVVSQATPKQILISCARALSEGKLEEALSMVNELRQIVSIQGDPSQRIAAY
MVEGLAARMAASGKFIYRALKCKEPPSDERLAAMQVLFEVCPCFKFGFLAA
NGAILEAIKGEEEVHIIDFDINQGNQYMTLIRSIAELPGKRPRLRLTGIDDPESV
QRSIGGLRIIGLRLEQLAEDNGVSFKFKAMPSKTSIVSPSTLGCKPGETLIVNFA
FQLHHMPDESVTTVNQRDELLHMVKSLNPKLVTVVEQDVNTNTSPFFPRFIE
AYEYYSAVFESLDMTLPRESQERMNVERQCLARDIVNIVACEGEERIERYEA
AGKWRARMMMAGFNPKPMSAKVTNNIQNLIKQQYCNKYKLKEEMGELHFC
WEEKSLIVASAWR
>AtGAI
MKRDHHHHHHQDKKTMMMNEEDDGNGMDELLAVLGYKVRSSEMADVAQ
KLEQLEVMMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSSNAE
YDLKAIPGDAILNQFAIDSASSSNQGGGGDTYTTNKRLKCSNGVVETTTATAE
STRHVVLVDSQENGVRLVHALLACAEAVQKENLTVAEALVKQIGFLAVSQIG
AMRKVATYFAEALARRIYRLSPSQSPIDHSLSDTLQMHFYETCPYLKFAHFTA
NQAILEAFQGKKRVHVIDFSMSQGLQWPALMQALALRPGGPPVFRLTGIGPP
APDNFDYLHEVGCKLAHLAEAIHVEFEYRGFVANTLADLDASMLELRPSEIES
VAVNSVFELHKLLGRPGAIDKVLGVVNQIKPEIFTVVEQESNHNSPIFLDRFTE
SLHYYSTLFDSLEGVPSGQDKVMSEVYLGKQICNVVACDGPDRVERHETLSQ
WRNRFGSAGFAAAHIGSNAFKQASMLLALFNGGEGYRVEESDGCLMLGWH
TRPLIATSAWKLSTN
>AtSCL3
MVAMFQEDNGTSSVASSPLQVFSTMSLNRPTLLASSSPFHCLKDLKPEERGLY
LIHLLLTCANHVASGSLQNANAALEQLSHLASPDGDTMQRIAAYFTEALANRI
LKSWPGLYKALNATQTRTNNVSEEIHVRRLFFEMFPILKVSYLLTNRAILEAM
EGEKMVHVIDLDASEPAQWLALLQAFNSRPEGPPHLRITGVHHQKEVLEQM
AHRLIEEAEKLDIPFQFNPVVSRLDCLNVEQLRVKTGEALAVSSVLQLHTFLA
SDDDLMRKNCALRFQNNPSGVDLQRVLMMSHGSAAEARENDMSNNNGYSP
SGDSASSLPLPSSGRTDSFLNAIWGLSPKVMVVTEQDSDHNGSTLMERLLESL
YTYAALFDCLETKVPRTSQDRIKVEKMLFGEEIKNIISCEGFERRERHEKLEK
WSQRIDLAGFGNVPLSYYAMLQARRLLQGCGFDGYRIKEESGCAVICWQDR
PLYSVSAWRCRK
>AtSCL4
MAYMCTDSGNLMAIAQQVIKQKQQQEQQQQHHQDHQIFGINPLSLNPWPN
TSLGFGLSGSAFPDPFQVTGGGDSNDPGFPFPNLDHHHATTTGGGFRLSDFGG
GTGGGEFESDEWMETLISGGDSVADGPDCDTWHDNPDYVIYGPDPFDTYPSR
LSVQPSDLNRVIDTSSPLPPPTLWPPSSPLSIPPLTHESPTKEDPETNDSEDDDFD

LEPPLLKAIYDCARISDSDPNEASKTLLQIRESVSELGDPTERVAFYFTEALSNR
LSPNSPATSSSSSSTEDLILSYKTLNDACPYSKFAHLTANQAILEATEKSNKIHI
VDFGIVQGIQWPALLQALATRTSGKPTQIRVSGIPAPSLGESPEPSLIATGNRLR
DFAKVLDLNFDFIPILTPIHLLNGSSFRVDPDEVLAVNFMLQLYKLLDETPTIV
DTALRLAKSLNPRVVTLGEYEVSLNRVGFANRVKNALQFYSAVFESLEPNLG
RDSEERVRVERELFGRRISGLIGPEKTGIHRERMEEKEQWRVLMENAGFESVK
LSNYAVSQAKILLWNYNYSNLYSIVESKPGFISLAWNDLPLLTLSSWR
>AtSCL5
MRLSVFIIPLVESRQASGIINKQSTSLLIRFSLYLEASISTKSFFSKSQRISQTQSPI
CLSANYYQPDNLDMEATQKHMIQEGSSMFYHQPSSVKQMDLSVQTFDSYCT
LESSSGTKSHPCLNNKNNSSSTTSFSSNESPISQANNNNLSRFNNHSPEENNNS
PLSGSSATNTNETELSLMLKDLETAMMEPDVDNSYNNQGGFGQQHGVVSSA
MYRSMEMISRGDLKGVLYECAKAVENYDLEMTDWLISQLQQMVSVSGEPV
QRLGAYMLEGLVARLASSGSSIYKALRCKDPTGPELLTYMHILYEACPYFKF
GYESANGAIAEAVKNESFVHIIDFQISQGGQWVSLIRALGARPGGPPNVRITGI
DDPRSSFARQGGLELVGQRLGKLAEMCGVPFEFHGAALCCTEVEIEKLGVRN
GEALAVNFPLVLHHMPDESVTVENHRDRLLRLVKHLSPNVVTLVEQEANTN
TAPFLPRFVETMNHYLAVFESIDVKLARDHKERINVEQHCLAREVVNLIACEG
VEREERHEPLGKWRSRFHMAGFKPYPLSSYVNATIKGLLESYSEKYTLEERD
GALYLGWKNQPLITSCAWR
>AtSCL6
MPLPFEEFQGKGISCFSSFSSSFPQPPSSPLLSHRKARGGEEEEEEVPAAEPTSV
LDSLISPTSSSTVSSSHGGNSAVGGGGDATTDEQCGAIGLGDWEEQVPHDHE
QSILGLIMGDSTDPSLELNSILQTSPTFHDSDYSSPGFGVVDTGFGLDHHSVPPS
HVSGLLINQSQTHYTQNPAAIFYGHHHHTPPPAKRLNPGPVGITEQLVKAAEV
IESDTCLAQGILARLNQQLSSPVGKPLERAAFYFKEALNNLLHNVSQTLNPYS
LIFKIAAYKSFSEISPVLQFANFTSNQALLESFHGFHRLHIIDFDIGYGGQWASL
MQELVLRDNAAPLSLKITVFASPANHDQLELGFTQDNLKHFASEINISLDIQVL
SLDLLGSISWPNSSEKEAVAVNISAASFSHLPLVLRFVKHLSPTIIVCSDRGCER
TDLPFSQQLAHSLHSHTALFESLDAVNANLDAMQKIERFLIQPEIEKLVLDRSR
PIERPMMTWQAMFLQMGFSPVTHSNFTESQAECLVQRTPVRGFHVEKKHNS
LLLCWQRTELVGVSAWRCRSS
>AtSCL7
MAYMCTDSGNLMAIAQQLIKQKQQQSQHQQQEEQEQEPNPWPNPSFGFTL
PGSGFSDPFQVTNDPGFHFPHLEHHQNAAVASEEFDSDEWMESLINGGDASQ
TNPDFPIYGHDPFVSFPSRLSAPSYLNRVNKDDSASQQLPPPPASTAIWSPSPPS
PQHPPPPPPQPDFDLNQPIFKAIHDYARKPETKPDTLIRIKESVSESGDPIQRVG
YYFAEALSHKETESPSSSSSSSLEDFILSYKTLNDACPYSKFAHLTANQAILEAT
NQSNNIHIVDFGIFQGIQWSALLQALATRSSGKPTRIRISGIPAPSLGDSPGPSLI
ATGNRLRDFAAILDLNFEFYPVLTPIQLLNGSSFRVDPDEVLVVNFMLELYKL
LDETATTVGTALRLARSLNPRIVTLGEYEVSLNRVEFANRVKNSLRFYSAVFE
SLEPNLDRDSKERLRVERVLFGRRIMDLVRSDDDNNKPGTRFGLMEEKEQW
RVLMEKAGFEPVKPSNYAVSQAKLLLWNYNYSTLYSLVESEPGFISLAWNNV
PLLTVSSWR

```
>AtSCL8
MESGFSGGGGGSDFYGGGGGRSIPGGPGTVINVGNNNPQTTYRNQIPGIFFDQ
IGNRVAGGNGFSGKRTLADFQAAQQHQQQQQQPFYSQAALNAFLSRSVKP
RNYQNFQSPSPMIDLTSVNDMSLFGGSGSSQRYGLPVPRSQTQQQQSDYGLF
GGIRMGIGSGINNYPTLTGVPCIEPVQNRVHESENMLNSLRELEKQLLDDDDE
SGGDDDVSVITNSNSDWIQNLVTPNPNPNPVLSFSPSSSSSSSSPSTASTTTSVC
SRQTVMEIATAIAEGKTEIATEILARVSQTPNLERNSEEKLVDFMVAALRSRIA
SPVTELYGKEHLISTQLLYELSPCFKLGFEAANLAILDAADNNDGGMMIPHVI
DFDIGEGGQYVNLLRTLSTRRNGKSQSQNSPVVKITAVANNVYGCLVDDGGE
ERLKAVGDLLSQLGDRLGISVSFNVVTSLRLGDLNRESLGCDPDETLAVNLAF
KLYRVPDESVCTENPRDELLRRVKGLKPRVVTLVEQEMNSNTAPFLGRVSES
CACYGALLESVESTVPSTNSDRAKVEEGIGRKLVNAVACEGIDRIERCEVFGK
WRMRMSMAGFELMPLSEKIAESMKSRGNRVHPGFTVKEDNGGVCFGWMGR
ALTVASAWR
>AtSHR
MDTLFRLVSLQQQQQSDSIITNQSSLSRTSTTTTGSPQTAYHYNFPQNDVVEE
CFNFFMDEEDLSSSSSHHNHHNHNNPNTYYSPFTTPTQYHPATSSTPSSTAAA
AALASPYSSSGHHNDPSAFSIPQTPPSFDFSANAKWADSVLLEAARAFSDKDT
ARAQQILWTLNELSSPYGDTEQKLASYFLQALFNRMTGSGERCYRTMVTAA
ATEKTCSFESTRKTVLKFQEVSPWATFGHVAANGAILEAVDGEAKIHIVDISS
TFCTQWPTLLEALATRSDDTPHLRLTTVVVANKFVNDQTASHRMMKEIGNR
MEKFARLMGVPFKFNIIHHVGDLSEFDLNELDVKPDEVLAINCVGAMHGIAS
RGSPRDAVISSFRRLRPRIVTVVEEEADLVGEEEGGFDDEFLRGFGECLRWFR
VCFESWEESFPRTSNERLMLERAAGRAIVDLVACEPSDSTERRETARKWSRR
MRNSGFGAVGYSDEVADDVRALLRRYKEGVWSMVQCPDAAGIFLCWRDQP
VVWASAWRPT
>AtRGA
MKRDHHQFQGRLSNHGTSSSSSSISKDKMMMVKKEEDGGGNMDDELLAVL
GYKVRSSEMAEVALKLEQLETMMSNVQEDGLSHLATDTVHYNPSELYSWLD
NMLSELNPPPLPASSNGLDPVLPSPEICGFPASDYDLKVIPGNAIYQFPAIDSSS
SSNNQNKRLKSCSSPDSMVTSTSTGTQIGGVIGTTVTTTTTTTTAAGESTRSVI
LVDSQENGVRLVHALMACAEAIQQNNLTLAEALVKQIGCLAVSQAGAMRK
VATYFAEALARRIYRLSPPQNQIDHCLSDTLQMHFYETCPYLKFAHFTANQAI
LEAFEGKKRVHVIDFSMNQGLQWPALMQALALREGGPPTFRLTGIGPPAPDN
SDHLHEVGCKLAQLAEAIHVEFEYRGFVANSLADLDASMLELRPSDTEAVAV
NSVFELHKLLGRPGGIEKVLGVVKQIKPVIFTVVEQESNHNGPVFLDRFTESL
HYYSTLFDSLEGVPNSQDKVMSEVYLGKQICNLVACEGPDRVERHETLSQW
GNRFGSSGLAPAHLGSNAFKQASMLLSVFNSGQGYRVEESNGCLMLGWHTR
PLITTSAWKLSTAAY
>AtSCL11
MDALLQVSVDGFRFENGSGSCCKPRNNLESGNNLFPDFHESQNQSSPNDSPPT
VCLDNSPVLKYINDMLMDEEDFVGISRDDLALQAAERSFYEIIQQQSPESDQN
TSSSSDQNSGDQDFCFPSTTTDSSALVSSGESQRKYRHRNDEEDDLENNRRNK
QPAIFVSEMEELAVKLEHVLLVCKTNQEEEEERTVITKQSTPNRAGRAKGSSN
```

KSKTHKTNTVDLRSLLTQCAQAVASFDQRRATDKLKEIRAHSSSNGDGTQRL
AFYFAEALEARITGNISPPVSNPFPSSTTSMVDILKAYKLFVHTCPIYVTDYFA
ANKSIYELAMKATKLHIVDFGVLYGFQWPCLLRALSKRPGGPPMLRVTGIEL
PQAGFRPSDRVEETGRRLKRFCDQFNVPFEFNFIAKKWETITLDELMINPGETT
VVNCIHRLQYTPDETVSLDSPRDTVLKLFRDINPDLFVFAEINGMYNSPFFMT
RFREALFHYSSLFDMFDTTIHAEDEYKNRSLLERELLVRDAMSVISCEGAERF
ARPETYKQWRVRILRAGFKPATISKQIMKEAKEIVRKRYHRDFVIDSDNNWM
LQGWKGRVIYAFSCWKPAEKFTNNNLNI
>AtSCL13
MQTSQKHHSAAGLHMLYPQVYCSPQFQAKDNKGFSDIPSKENFFTLESSTAS
GSLPSYDSPSVSITSGRSPFSPQGSQSCISDLHHSPDNVYGSPLSGVSSLAYDEA
GVKSKIRELEVSLLSGDTKVEEFSGFSPAAGKSWNWDELLALTPQLDLKEVL
VEAARAVADGDFATAYGFLDVLEQMVSVSGSPIQRLGTYMAEGLRARLEGS
GSNIYKSLKCNEPTGRELMSYMSVLYEICPYWKFAYTTANVEILEAIAGETRV
HIIDFQIAQGSQYMFLIQELAKRPGGPPLLRVTGVDDSQSTYARGGGLSLVGE
RLATLAQSCGVPFEFHDAIMSGCKVQREHLGLEPGFAVVVNFPYVLHHMPDE
SVSVENHRDRLLHLIKSLSPKLVTLVEQESNTNTSPFLSRFVETLDYYTAMFES
IDAARPRDDKQRISAEQHCVARDIVNMIACEESERVERHEVLGIWRVRMMM
AGFTGWPVSTSAAFAASEMLKAYDKNYKLGGHEGALYLFWKRRPMATCSV
WKPNPN
>AtSCL14
MGSYPDGFPGSMDELDFNKDFDLPPSSNQTLGLANGFYLDDLDFSSLDPPEA
YPSQNNNNNNINNKAVAGDLLSSSSDDADFSDSVLKYISQVLMEEDMEEKPC
MFHDALALQAAEKSLYEALGEKYPSSSSASSVDHPERLASDSPDGSCSGGAFS
DYASTTTTTSSDSHWSVDGLENRPSWLHTPMPSNFVFQSTSRSNSVTGGGGG
GNSAVYGSGFGDDLVSNMFKDDELAMQFKKGVEEASKFLPKSSQLFIDVDSY
IPMNSGSKENGSEVFVKTEKKDETEHHHHHSYAPPPNRLTGKKSHWRDEDED
FVEERSNKQSAVYVEESELSEMFDKILVCGPGKPVCILNQNFPTESAKVVTAQ
SNGAKIRGKKSTSTSHSNDSKKETADLRTLLVLCAQAVSVDDRRTANEMLRQ
IREHSSPLGNGSERLAHYFANSLEARLAGTGTQIYTALSSKKTSAADMLKAYQ
TYMSVCPFKKAAIIFANHSMMRFTANANTIHIIDFGISYGFQWPALIHRLSLSR
PGGSPKLRITGIELPQRGFRPAEGVQETGHRLARYCQRHNVPFEYNAIAQKWE
TIQVEDLKLRQGEYVVVNSLFRFRNLLDETVLVNSPRDAVLKLIRKINPNVFIP
AILSGNYNAPFFVTRFREALFHYSAVFDMCDSKLAREDEMRLMYEKEFYGRE
IVNVVACEGTERVERPETYKQWQARLIRAGFRQLPLEKELMQNLKLKIENGY
DKNFDVDQNGNWLLQGWKGRIVYASSLWVPSSS
>AtSCL15
MKIPASSPQDTTNNNNNTNSTDSNHLSMDEHVMRSMDWDSIMKELELDDDS
APNSLKTGFTTTTTDSTILPLYAVDSNLPGFPDQIQPSDFESSSDVYPGQNQTT
GYGFNSLDSVDNGGFDFIEDLIRVVDCVESDELQLAQVVLSRLNQRLRSPAGR
PLQRAAFYFKEALGSFLTGSNRNPIRLSSWSEIVQRIRAIKEYSGISPIPLFSHFT
ANQAILDSLSSQSSSPFVHVVDFEIGFGGQYASLMREITEKSVSGGFLRVTAVV
AEECAVETRLVKENLTQFAAEMKIRFQIEFVLMKTFEMLSFKAIRFVEGERTV
VLISPAIFRRLSGITDFVNNLRRVSPKVVVFVDSEGWTEIAGSGSFRREFVSAL

EFYTMVLESLDAAAPPGDLVKKIVEAFVLRPKISAAVETAADRRHTGEMTWR
EAFCAAGMRPIQLSQFADFQAECLLEKAQVRGFHVAKRQGELVLCWHGRAL
VATSAWRF
>AtSCL16
MQIPTLIDSMANKLHKKPPPLLKLTVIASDAEFHPPPLLGISYEELGSKLVNFA
TTRNVAMEFRIISSSYSDGLSSLIEQLRIDPFVFNEALVVNCHMMLHYIPDEILT
SNLRSVFLKELRDLNPTIVTLIDEDSDFTSTNFISRLRSLYNYMWIPYDTAEMF
LTRGSEQRQWYEADISWKIDNVVAKEGAERVERLEPKSR
>AtRGL1
MKREHNHRESSAGEGGSSSMTTVIKEEAAGVDELLVVLGYKVRSSDMADVA
HKLEQLEMVLGDGISNLSDETVHYNPSDLSGWVESMLSDLDPTRIQEKPDSE
YDLRAIPGSAVYPRDEHVTRRSKRTRIESELSSTRSVVVLDSQETGVRLVHAL
LACAEAVQQNNLKLADALVKHVGLLASSQAGAMRKVATYFAEGLARRIYRI
YPRDDVALSSFSDTLQIHFYESCPYLKFAHFTANQAILEVFATAEKVHVIDLGL
NHGLQWPALIQALALRPNGPPDFRLTGIGYSLTDIQEVGWKLGQLASTIGVNF
EFKSIALNNLSDLKPEMLDIRPGLESVAVNSVFELHRLLAHPGSIDKFLSTIKSI
RPDIMTVVEQEANHNGTVFLDRFTESLHYYSSLFDSLEGPPSQDRVMSELFLG
RQILNLVACEGEDRVERHETLNQWRNRFGLGGFKPVSIGSNAYKQASMLLAL
YAGADGYNVEENEGCLLLGWQTRPLIATSAWRINRVE
>AtSCL18
MLTSFKSSSSSSEDATATTTENPPPLCIASSSAATSASHHLRRLLFTAANFVSQS
NFTAAQNLLSILSLNSSPHGDSTERLVHLFTKALSVRINRQQQDQTAETVATW
TTNEMTMSNSTVFTSSVCKEQFLFRTKNNNSDFESCYYLWLNQLTPFIRFGHL
TANQAILDATETNDNGALHILDLDISQGLQWPPLMQALAERSSNPSSPPPSLRI
TGCGRDVTGLNRTGDRLTRFADSLGLQFQFHTLVIVEEDLAGLLLQIRLLALS
AVQGETIAVNCVHFLHKIFNDDGDMIGHFLSAIKSLNSRIVTMAEREANHGD
HSFLNRFSEAVDHYMAIFDSLEATLPPNSRERLTLEQRWFGKEILDVVAAEET
ERKQRHRRFEIWEEMMKRFGFVNVPIGSFALSQAKLLLRLHYPSEGYNLQFL
NNSLFLGWQNRPLFSVSSWK
>AtRGL2
MKRGYGETWDPPPKPLPASRSGEGPSMADKKKADDDNNNSNMDDELLAVL
GYKVRSSEMAEVAQKLEQLEMVLSNDDVGSTVLNDSVHYNPSDLSNWVES
MLSELNNPASSDLDTTRSCVDRSEYDLRAIPGLSAFPKEEEVFDEEASSKRIRL
GSWCESSDESTRSVVLVDSQETGVRLVHALVACAEAIHQENLNLADALVKR
VGTLAGSQAGAMGKVATYFAQALARRIYRDYTAETDVCAAVNPSFEEVLEM
HFYESCPYLKFAHFTANQAILEAVTTARRVHVIDLGLNQGMQWPALMQALA
LRPGGPPSFRLTGIGPPQTENSDSLQQLGWKLAQFAQNMGVEFEFKGLAAESL
SDLEPEMFETRPESETLVVNSVFELHRLLARSGSIEKLLNTVKAIKPSIVTVVE
QEANHNGIVFLDRFNEALHYYSSLFDSLEDSYSLPSQDRVMSEVYLGRQILNV
VAAEGSDRVERHETAAQWRIRMKSAGFDPIHLGSSAFKQASMLLSLYATGD
GYRVEENDGCLMIGWQTRPLITTSAWKLA
>AtSCR
MAESGDFNGGQPPPHSPLRTTSSGSSSSNNRGPPPPPPPPLVMVRKRLASEMSS
NPDYNNSSRPPRRVSHLLDSNYNTVTPQQPPSLTAAATVSSQPNPPLSVCGFS

GLPVFPSDRGGRNVMMSVQPMDQDSSSSSASPTVWVDAIIRDLIHSSTSVSIPQ
LIQNVRDIIFPCNPNLGALLEYRLRSLMLLDPSSSSDPSPQTFEPLYQISNNPSPP
QQQQQHQQQQQQHKPPPPPIQQQERENSSTDAPPQPETVTATVPAVQTNTAE
ALRERKEEIKRQKQDEEGLHLLTLLLQCAEAVSADNLEEANKLLLEISQLSTP
YGTSAQRVAAYFSEAMSARLLNSCLGIYAALPSRWMPQTHSLKMVSAFQVF
NGISPLVKFSHFTANQAIQEAFEKEDSVHIIDLDIMQGLQWPGLFHILASRPGG
PPHVRLTGLGTSMEALQATGKRLSDFADKLGLPFEFCPLAEKVGNLDTERLN
VRKREAVAVHWLQHSLYDVTGSDAHTLWLLQRLAPKVVTVVEQDLSHAGS
FLGRFVEAIHYYSALFDSLGASYGEESEERHVVEQQLLSKEIRNVLAVGGPSR
SGEVKFESWREKMQQCGFKGISLAGNAATQATLLLGMFPSDGYTLVDDNGT
LKLGWKDLSLLTASAWTPRS
>AtSCL9
MITEPSLTGISGMVNRNRLSGLPDQPSSHSFTPVTLYDGFNYNLSSDHINTVVA
APENSVFIREEEEEEDPADDFDFSDAVLGYISQMLNEEDMDDKVCMLQESLD
LEAAERSLYEAIGKKYPPSPERNLAFAERNSENLDRVVPGNYTGGDCIGFGNG
GIKPLSSGFTLDFRNPQSCSSILSVPQSNGLITIYGDGIDESSKNNRENHQSVWL
FRREIEEANRFNPEENELIVNFREENCVSKARKNSSRDEICVEEERSSKLPAVF
GEDILRSDVVDKILVHVPGGESMKEFNALRDVLKKGVEKKKASDAQGGKRR
ARGRGRGRGRGGGGGQNGKKEVVDLRSLLIHCAQAVAADDRRCAGQLLKQ
IRLHSTPFGDGNQRLAHCFANGLEARLAGTGSQIYKGIVSKPRSAAAVLKAH
QLFLACCPFRKLSYFITNKTIRDLVGNSQRVHVIDFGILYGFQWPTLIHRFSMY
GSPKVRITGIEFPQPGFRPAQRVEETGQRLAAYAKLFGVPFEYKAIAKKWDAI
QLEDLDIDRDEITVVNCLYRAENLHDESVKVESCRDTVLNLIGKINPDLFVFGI
VNGAYNAPFFVTRFREALFHFSSIFDMLETIVPREDEERMFLEMEVFGREALN
VIACEGWERVERPETYKQWHVRAMRSGLVQVPFDPSIMKTSLHKVHTFYHK
DFVIDQDNRWLLQGWKGRTVMALSVWKPESKA
>AtSCL21
MDNVRGSIMLQPLPEIAESIDDAICHELSMWPDDAKDLLLIVEAISRGDLKLV
LVACAKAVSENNLLMARWCMGELRGMVSISGEPIQRLGAYMLEGLVARLA
ASGSSIYKSLQSREPESYEFLSYVYVLHEVCPYFKFGYMSANGAIAEAMKDEE
RIHIIDFQIGQGSQWIALIQAFAARPGGAPNIRITGVGDGSVLVTVKKRLEKLA
KKFDVPFRFNAVSRPSCEVEVENLDVRDGEALGVNFAYMLHHLPDESVSME
NHRDRLLRMVKSLSPKVVTLVEQECNTNTSPFLPRFLETLSYYTAMFESIDVM
LPRNHKERINIEQHCMARDVVNIIACEGAERIERHELLGKWKSRFSMAGFEPY
PLSSIISATIRALLRDYSNGYAIEERDGALYLGWMDRILVSSCAWK
>AtSCL22
MPLPFEQFQGKGVLGFLDSSSSPGYKIWANPEKLHGRVEEDLCFVVNNGGFS
EPTSVLDSVRSPSPFVSSSTTTLSSSHGGPSGGGAAAATFSGADGKCDQMGFE
DLDGVLSGGSPGQEQSIFRLIMAGDVVDPGSEFVGFDIGSGSDPVIDNPNPLFG
YGFPFQNAPEEEKFQISINPNPGFFSDPPSSPPAKRLNSGQPGSQHLQWVFPFSD
PGHESHDPFLTPPKIAGEDQNDQDQSAVIIDQLFSAAAELTTNGGDNNPVLAQ
GILARLNHNLNNNNDDTNNNPKPPFHRAASYITEALHSLLQDSSLSPPSLSPPQ
NLIFRIAAYRAFSETSPFLQFVNFTANQTILESFEGFDRIHIVDFDIGYGGQWAS
LIQELAGKRNRSSSAPSLKITAFASPSTVSDEFELRFTEENLRSFAGETGVSFEIE

LLNMEILLNPTYWPLSLFRSSEKEAIAVNLPISSMVSGYLPLILRFLKQISPNVV
VCSDRSCDRNNDAPFPNGVINALQYYTSLLESLDSGNLNNAEAATSIERFCVQ
PSIQKLLTNRYRWMERSPPWRSLFGQCGFTPVTLSQTAETQAEYLLQRNPMR
GFHLEKRQSSSPSLVLCWQRKELVTVSAWKC
>AtSCL23
MTTKRIDRDLPSSDDPSSAKRRIEFPEETLENDGAAAIKLLSLLLQCAEYVATD
HLREASTLLSEISEICSPFGSSPERVVAYFAQALQTRVISSYLSGACSPLSEKPLT
VVQSQKIFSALQTYNSVSPLIKFSHFTANQAIFQALDGEDSVHIIDLDVMQGLQ
WPALFHILASRPRKLRSIRITGFGSSSDLLASTGRRLADFASSLNLPFEFHPIEGII
GNLIDPSQLATRQGEAVVVHWMQHRLYDVTGNNLETLEILRRLKPNLITVVE
QELSYDDGGSFLGRFVEALHYYSALFDALGDGLGEESGERFTVEQIVLGTEIR
NIVAHGGGRRKRMKWKEELSRVGFRPVSLRGNPATQAGLLLGMLPWNGYT
LVEENGTLRLGWKDLSLLTASAWKSQPFD
>AtPAT1
MYKQPRQELEAYYFEPNSVEKLRYLPVNNSRKRFCTLEPFPDSPPYNALSTAT
YDDTCGSCVTDELNDFKHKIREIETVMMGPDSLDLLVDCTDSFDSTASQEING
WRSTLEAISRRDLRADLVSCAKAMSENDLMMAHSMMEKLRQMVSVSGEPIQ
RLGAYLLEGLVAQLASSGSSIYKALNRCPEPASTELLSYMHILYEVCPYFKFG
YMSANGAIAEAMKEENRVHIIDFQIGQGSQWVTLIQAFAARPGGPPRIRITGID
DMTSAYARGGGLSIVGNRLAKLAKQFNVPFEFNSVSVSVSEVKPKNLGVRPG
EALAVNFAFVLHHMPDESVSTENHRDRLLRMVKSLSPKVVTLVEQESNTNTA
AFFPRFMETMNYYAAMFESIDVTLPRDHKQRINVEQHCLARDVVNIIACEGA
DRVERHELLGKWRSRFGMAGFTPYPLSPLVNSTIKSLLRNYSDKYRLEERDG
ALYLGWMHRDLVASCAWK
>AtRGL3
MKRSHQETSVEEEAPSMVEKLENGCGGGGDDNMDEFLAVLGYKVRSSDMA
DVAQKLEQLEMVLSNDIASSSNAFNDTVHYNPSDLSGWAQSMLSDLNYYPD
LDPNRICDLRPITDDDECCSSNSNSNKRIRLGPWCDSVTSESTRSVVLIEETGV
RLVQALVACAEAVQLENLSLADALVKRVGLLAASQAGAMGKVATYFAEAL
ARRIYRIHPSAAAIDPSFEEILQMNFYDSCPYLKFAHFTANQAILEAVTTSRVV
HVIDLGLNQGMQWPALMQALALRPGGPPSFRLTGVGNPSNREGIQELGWKL
AQLAQAIGVEFKFNGLTTERLSDLEPDMFETRTESETLVVNSVFELHPVLSQP
GSIEKLLATVKAVKPGLVTVVEQEANHNGDVFLDRFNEALHYYSSLFDSLED
GVVIPSQDRVMSEVYLGRQILNLVATEGSDRIERHETLAQWRKRMGSAGFDP
VNLGSDAFKQASLLLALSGGGDGYRVEENDGSLMLAWQTKPLIAASAWKLA
AELRR
>AtSCL26
MNYPYEDFLDLFFSTHTDPLATAASTSSNGYSLNDLDIDWDCDFRDVIESIMG
DEGAMMEPESEAVPMLHDQEGLCNSASTGLSVADGVSFGEPKTDESKGLRL
VHLLVAAADASTGANKSRELTRVILARLKDLVSPGDRTNMERLAAHFTNGLS
KLLERDSVLCPQQHRDDVYDQADVISAFELLQNMSPYVNFGYLTATQAILEA
VKYERRIHIVDYDINEGVQWASLMQALVSRNTGPSAQHLRITALSRATNGKK
SVAAVQETGRRLTAFADSIGQPFSYQHCKLDTNAFSTSSLKLVRGEAVVINC
MLHLPRFSHQTPSSVISFLSEAKTLNPKLVTLVHEEVGLMGNQGFLYRFMDLL

```
HQFSAIFDSLEAGLSIANPARGFVERVFIGPWVANWLTRITANDAEVESFASW
PQWLETNGFKPLEVSFTNRCQAKLLLSLFNDGFRVEELGQNGLVLGWKSRRL
VSASFWASCQTNQ
>AtSCL27
MPLSFERFQGEGVFGLSSSSFYSDSQKIWSNQDKTEAKQEDLGYVVGGFLPEP
TSVLDALRSPSPLASYSSTTTTLSSSHGGGGTTVTNTTVTAGDDNNNNKCSQ
MGLDDLDGVLSASSPGQEQSILRLIMDPGSAFGVFDPGFGFGSGSGPVSAPVS
DNSNLLCNFPFQEITNPAEALINPSNHCLFYNPPLSPPAKRFNSGSLHQPVFPLS
DPDPGHDPVRRQHQFQFPFYHNNQQQQFPSSSSSTAVAMVPVPSPGMAGDD
QSVIIEQLFNAAELIGTTGNNNGDHTVLAQGILARLNHHLNTSSNHKSPFQRA
ASHIAEALLSLIHNESSPPLITPENLILRIAAYRSFSETSPFLQFVNFTANQSILES
CNESGFDRIHIIDFDVGYGGQWSSLMQELASGVGGRRRNRASSLKLTVFAPPP
STVSDEFELRFTEENLKTFAGEVKIPFEIELLSVELLLNPAYWPLSLRSSEKEAI
AVNLPVNSVASGYLPLILRFLKQLSPNIVVCSDRGCDRNDAPFPNAVIHSLQY
HTSLLESLDANQNQDDSSIERFWVQPSIEKLLMKRHRWIERSPPWRILFTQCG
FSPASLSQMAEAQAECLLQRNPVRGFHVEKRQSSLVMCWQRKELVTVSAWK
C
>AtSCL28
MLAGCSSSSLLSPTRRLRSEAVAATSATVSAHFPMNTQRLDLPCSSSFSRKETP
SSRPLGRSISLDNSNNNNNKPIERKTKTSGCSLKQNIKLPPLATTRGNGEGFSW
NNDNNNRGKSLKRLAEEDESCLSRAKRTKCENEGGFWFEHFTGQDSSSPALP
FSLTCSGDDEEKVCFVPSEVISQPLPNWVDSVITELAGIGDKDVESSLPAAVKE
ASGGSSTSASSESRSLSHRVPEPTNGSRNPYSHRGATEERTTGNINNNNNRND
LQRDFELVNLLTGCLDAIRSRNIAAINHFIARTGDLASPRGRTPMTRLIAYYIE
ALALRVARMWPHIFHIAPPREFDRTVEDESGNALRFLNQVTPIPKFIHFTANE
MLLRAFEGKERVHIIDFDIKQGLQWPSFFQSLASRINPPHHVRITGIGESKLELN
ETGDRLHGFAEAMNLQFEFHPVVDRLEDVRLWMLHVKEGESVAVNCVMQ
MHKTLYDGTGAAIRDFLGLIRSTNPIALVLAEQEAEHNSEQLETRVCNSLKYY
SAMFDAIHTNLATDSLMRVKVEEMLFGREIRNIVACEGSHRQERHVGFRHWR
RMLEQLGFRSLGVSEREVLQSKMLLRMYGSDNEGFFNVERSDEDNGGEGGR
GGGVTLRWSEQPLYTISAWTTGGN
>AtSCL29
MLLEETEPPNQTLDHVLSWLEDSVSLSPLPGFDDSYLLHEFDGSQTWEWDQT
QDPEHGFIQSYSQDLSAAYVGCEATNLEVVTEAPSIDLDLPPEIQQPNDQSRK
RSHDGFLEAQQVKKSARSKRKAIKSSEKSSKDGNKEGRWAEKLLNPCALAIT
ASNSSRVQHYLCVLSELASSSGDANRRLAAFGLRALQHHLSSSSVSSSFWPVF
TFASAEVKMFQKTLLKFYEVSPWFALPNNMANSAILQILAQDPKDKKDLHIID
IGVSHGMQWPTLLEALSCRLEGPPPRVRITVISDLTADIPFSVGPPGYNYGSQL
LGFARSLKINLQISVLDKLQLIDTSPHENLIVCAQFRLHHLKHSINDERGETLK
AVRSLRPKGVVLCENNGECSSSADFAAGFSKKLEYVWKFLDSTSSGFKEENS
EERKLMEGEATKVLMNAGDMNEGKEKWYERMREAGFFVEAFEEDAVDGA
KSLLRKYDNNWEIRMEDGDTFAGLMWKGEAVSFCSLWK
>AtSCL30
```

```
MDAILPVPVDGFRFDTGSGSCCKPRNNLESGTTNRFTCFNESESQSNPSPTESK
VCSDYLPVFKYINDMLMEEDLEGQSCMLEDSLALQAAERSFFEVLQDQTPIS
GDLEDGSLGNFSSITSLHQPEVSEESTRRYRHRDDDEDDDLESGRKSKLPAIST
VDELAEKFEEVLLVCQKNDQGEATEKKTRHVKGSSNRYKQQKSDQPVDMR
NLLMQCAQAVASFDQRRAFEKLKEIREHSSRHGDATQRLGYHFAEALEARIT
GTMTTPISATSSRTSMVDILKAYKGFVQACPTLIMCYFTANRTINELASKATTL
HIIDFGILYGFQWPCLIQALSKRDIGPPLLRVTGIELPQSGFRPSERVEETGRRL
KRFCDKFNVPFEYSFIAKNWENITLDDLVINSGETTVVNCILRLQYTPDETVSL
NSPRDTALKLFRDINPDLFVFAEINGTYNSPFFLTRFREALFHCSSLFDMYETT
LSEDDNCRTLVERELIIRDAMSVIACEGSERFARPETYKQWQVRILRAGFRPA
KLSKQIVKDGKEIVKERYHKDFVIDNDNHWMFQGWKGRVLYAVSCWKPAK
K
>AtSCL31
MESNYSGVVNGYDVSFLPTSIPDLGFGVPSSSDFDLRMDQYYHQPSIWVPDQ
DHHFSPPADEIDSENTLLKYVNQLLMEESLAEKQSIFYDSLALRQTEEMLQQV
ISDSQTQSSIPNNSITTSSSSNSGDYSNSSNSSVRIENEVLFDNKHLGDSGVVSFP
GSNMLRGGEQFGQPANEILVRSMFSDAESVLQFKRGLEEASKFLPNTDQWIF
NLEPEMERVVPVKVEEGWSAISKTRKNHHEREEEEDDLEEARRRSKQFAVNE
EDGKLTEMFDKVLLLDGECDPQIIEDGENGSSKALVKKGRAKKKSRAVDFRT
LLTLCAQSVSAGDKITADDLLRQIRKQCSPVGDASQRLAHFFANALEARLEGS
TGTMIQSYYDSISSKKRTAAQILKSYSVFLSASPFMTLIYFFSNKMILDAAKDA
SVLHIVDFGILYGFQWPMFIQHLSKSNPGLRKLRITGIEIPQHGLRPTERIQDTG
RRLTEYCKRFGVPFEYNAIASKNWETIKMEEFKIRPNEVLAVNAVLRFKNLR
DVIPGEEDCPRDGFLKLIRDMNPNVFLSSTVNGSFNAPFFTTRFKEALFHYSAL
FDLFGATLSKENPERIHFEGEFYGREVMNVIACEGVDRVERPETYKQWQVRM
IRAGFKQKPVEAELVQLFREKMKKWGYHKDFVLDEDSNWFLQGWKGRILFS
SSCWVPS
>AtSCL32
MTKTRILNPTRFPSPKPLRGCGDANFMEQLLLHCATAIDSNDAALTHQILWVL
NNIAPPDGDSTQRLTSAFLRALLSRAVSKTPTLSSTISFLPQADELHRFSVVELA
AFVDLTPWHRFGFIAANAAILTAVEGYSTVHIVDLSLTHCMQIPTLIDAMASR
LNKPPPLLKLTVVSSSDHFPPFINISYEELGSKLVNFATTRNITMEFTIVPSTYSD
GFSSLLQQLRIYPSSFNEALVVNCHMMLRYIPEEPLTSSSSSLRTVFLKQLRSL
NPRIVTLIEEDVDLTSENLVNRLKSAFNYFWIPFDTTDTFMSEQRRWYEAEIS
WKIENVVAKEGAERVERTETKRRWIERMREAEFGGVRVKEDAVADVKAML
EEHAVGWGMKKEDDDESLVLTWKGHSVVFATVWVPI
>AtSCL33
MGSYSAGFPGSLDWFDFPGLGNGSYLNDQPLLDIGSVPPPLDPYPQQNLASA
DADFSDSVLKYISQVLMEEDMEDKPCMFHDALSLQAAEKSLYEALGEKYPV
DDSDQPLTTTTSLAQLVSSPGGSSYASSTTTTSSDSQWSFDCLENNRPSSWLQ
TPIPSNFIFQSTSTRASSGNAVFGSSFSGDLVSNMFNDTDLALQFKKGMEEASK
FLPKSSQLVIDNSVPNRLTGKKSHWREEEHLTEERSKKQSAIYVDETDELTDM
FDNILIFGEAKEQPVCILNESFPKEPAKASTFSKSPKGEKPEASGNSYTKETPDL
RTMLVSCAQAVSINDRRTADELLSRIRQHSSSYGDGTERLAHYFANSLEARLA
```

GIGTQVYTALSSKKTSTSDMLKAYQTYISVCPFKKIAIIFANHSIMRLASSANA
KTIHIIDFGISDGFQWPSLIHRLAWRRGSSCKLRITGIELPQRGFRPAEGVIETGR
RLAKYCQKFNIPFEYNAIAQKWESIKLEDLKLKEGEFVAVNSLFRFRNLLDET
VAVHSPRDTVLKLIRKIKPDVFIPGILSGSYNAPFFVTRFREVLFHYSSLFDMC
DTNLTREDPMRVMFEKEFYGREIMNVVACEGTERVERPESYKQWQARAMR
AGFRQIPLEKELVQKLKLMVESGYKPKEFDVDQDCHWLLQGWKGRIVYGSSI
WVPFFFYVGRATRVLIMDPNFSESLNGFEYFDGNPNLLTDPMEDQYPPPSDTL
LKYVSEILMEESNGDYKQSMFYDSLALRKTEEMLQQVITDSQNQSFSPADSLI
TNSWDASGSIDESAYSADPQPVNEIMVKSMFSDAESALQFKKGVEEASKFLP
NSDQWVINLDIERSERRDSVKEEMGLDQLRVKKNHERDFEEVRSSKQFASNV
EDSKVTDMFDKVLLLDGECDPQTLLDSEIQAIRSSKNIGEKGKKKKKKKSQV
VDFRTLLTHCAQAISTGDKTTALEFLLQIRQQSSPLGDAGQRLAHCFANALEA
RLQGSTGPMIQTYYNALTSSLKDTAADTIRAYRVYLSSSPFVTLMYFFSIWMI
LDVAKDAPVLHIVDFGILYGFQWPMFIQSISDRKDVPRKLRITGIELPQCGFRP
AERIEETGRRLAEYCKRFNVPFEYKAIASQNWETIRIEDLDIRPNEVLAVNAGL
RLKNLQDETGSEENCPRDAVLKLIRNMNPDVFIHAIVNGSFNAPFFISRFKEAV
YHYSALFDMFDSTLPRDNKERIRFEREFYGREAMNVIACEEADRVERPETYR
QWQVRMVRAGFKQTIKPELVELFRGKLKKWRYHKDFVVDENSKWLLQG
WKGRTLYASSCWVPA
>Al 493792
MTTKRIDRDFPSSDDPSAAAKRRVNDIEFPEETLENDGAAAAIKLLSLLLQCA
EYVATNHLREASTLLSEISEICSPFGSSPERVVAYFAQALQTRVISSYLSGACTP
LSEKPLTVVQSQRLFSALQTFNSVSPLIKFSHFTANQAIFQALDGEDSVHIIDLD
VMQGLQWPALFHILASRPRKLRSIRITGFGSSSDLLASTGRRLADFASSLNLPF
EFHPIEGKIGNLIDPSQLGTRQGEAVVVHWMQHRLYDVTGNDLETLEILRRLK
PNLITVVEQELSYDDGGSFLGGFVEALHYYSALFDALGDGLGEESGERFTVEQ
IVLATEIRNIVAHGGRRRRRMKWKEELNRVGFRPVSLRGNPAMQAGLLLGM
LPWNGYTLVEENGTLRLGWKDLSLLTASAWKSQPFD*

Brachypodium dystachion

> Bd Bradi2g57940
MGMSPSPFTTSYCLTEQDASICTETQELDYHHSYYYGIEDVSLDEVELELAAS
SGRAHKATKVDYHSSPYHISWPPPPQQASADVESSRVRKKQFRDVLESCKQK
VEAMEAMEQHSPPLSGGGVGFQDQQGDSGHGVAVGGEGSSNSGGGTDGMR
LVQLLVACAEAVACRDRAQAASLLRELQAGAPVHGTAFQRVASCFVQGLAD
RLALAHPPALGPASMAFCIPQSSSSASCAGRGEALAVAYEVCPYLRFAHFVA
NASILEAFEGESKVHVVDLGMTLGLDRAHQWRALLDGLAARGVARPARVR
VTGVGARVDAMRAVGLELEAYAEELGMCVEFRAIDRTLESLHVDDLGVEAD
EAVAINSVLELHCVVKESRGALNSVLQTIRKLAPKAFVLVEQDAGHNGPFFL
GRFMEALHYYAALFDALDAALPRYDARRARVEQFHFGAEIRNVVGCEGAAR
VERHERADQWRRRMSRAGFQSMPIKMAAKAREWLEENAGGTGYTVAEEKG
CLVLGWKGKPVIAASCWKC*

Carica papaya

>Cp evm.TU.supercontig_34.30

MNSWDSFSKHETLNHELAIRRFCPARLEQEQGVSDQWSIDTVNDSVFSLPHD
QYHHDHNYIHQNVTTHGLPASDVNEFVDSFINMDHHDGQYDDHDHQNDDD
GNKSLQVKTQYLLDHHHDDHLHQQNWSGESDYNRYFPMMMPGGDHEVDL
YGSSVSHGVDQGLYLVHLLLACAEAVGCRDTKLADSILAQIWSSATLMGDSL
QRVSYCFAMGLKSRLSLLQTVTANGTFTNGGNHVCWVTMEEKMEAFHLLY
QTTPFIAFGFMAANEAICQATQGKYSLHIIDLGMEHTLQWPSLIRNLASRPEGP
PKKLKITAIINDQYRIELEGSTKLLVEEANSLGIPLEIYMISEPATPTLMTREKLG
LEEGEALLVNSIMHLHKYVKESRGSLKAILQAIKKLSPTLLTVVEQDANHNGP
FFLGRFLESLHYYSAIFDSLEASLLRNSPQRMKIEKYHFAEEIRNIVAYEGLDR
VERHERADQWRRQLGRAGYQVMGLKCLSQARMMLSVYGCDGYTLSCEKG
CLLLGWKGRPIMLASAWQANNSSS*
Citrus cinensis
>Cs orange1.1g040125m.g
MVYDFLFNTHKLYNDQRSNQAIGLQLGLSIVDCYPYLPMMSDNSAASSMML
QPRDQKRLKRTISVADSIAGDGSPTSTLSRSSSSSSLSNLPRLQFRDHIWTYTQ
RYLAAEAVEEAAAAMTKAVDGCGGDQQDGTADGMRLVQLLIACAEAVAC
RDKSHASALLSELRANALVFGSSFQRVASCFVQGLADRLASVQPLGAVGSFA
PSMNIMDIAGSREKEEAFRLVYEICPHIQFGHFVANSSILEAFEGESFVHVVDL
GMTLGLPRGQQWRRLIESLANRAGQPPRRLRITAVGLCVEKFQSIGDELKDY
AKTYGINLEFSVVESNLENLQTKDIKVLENEVLVVNSILQLHCVVKESRGALN
SVLQIIHELSPKVLVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPKY
DTKRAKIEQFYFAEEIKNIVSCEGPARVERHERVDQWRRRMSRAGFQAAPMK
MINQAQKWLKNNKVCEGYTVVEEKGCLVLGWKSKPIIATSCWKC*
Citrus clementina
>Ccl Ciclev10031305m.g
MEWGMYWRHGLCLDDKNNEVMGLDLNYSATACYRCFNLSTLANHSCTWS
TSDETRNHKRIKKTSMIDESSGSSSRLCNSLNSLPRLQFRDHVWVYTQRYLAI
EAMEEAALAMMAGKDTEVKEEGNGDGMKLVQQLIACAEAVACRDKAHAS
ALLLELRVNALVFGTSFQRVASCFVQGLSDRLALVQPLGAVGVVGPAAKSM
AITSQRDESLSLFYEICPQIQFGHFVANASILEAFEGESLVHVVDLGMTLGLPH
GRQWHSLMQSLVNRSGKVPKRLKITGVGNCSERLGEIGDELKRYADGLKLSF
EFLAVEKSLETLQAKDINVEDGEVLVMNSILELHCVVKESRGALNSVLQRLH
QLSPKVVMLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTKRAK
IEQFYFAEEIKNIVSCEGPARVERHERVDQWRRRMSRAGFQSVPIKMLMQAK
QWLRKVQTCEGYTIIEEKGCLVLGWKSKPIIAASCWKCS*
>Ccl Ciclev10019777m.g
MVYDFLFNTHKLYNDQRSNQAIGLQLGLSIVDCYPYLPMMSDNSAASSMILQ
PRDQKRLKRTISVGDGSPTSTLSRSSSTSSLSNLPRLQFRDHIWTYTRRYLAAE
AVEEAAAAMIKSEDGCDGDQQDGTADGMRLVQLLIACAEAVACRDKSHAS
ALLSELRANALVFGSSFQRVASCFVQGLADRLASVQPLGAVGSFAPSMNIMDI
AGSREKEEAFRLVYEICPHIQFGHFVANSSILEAFEGESFVHVVDLGMTLGLPR
GQQWRRLIESLANRAGQPPRRLRITAVGLCVEKFQSIGDELKDYAKTYGINLE
FSVVESNLENLQTKDIKVLENEVLVVNSILQLHCVVKESRGALNSVLQIIHELS
PKVLVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTKRAKIEQ

FYFAEEIKNIVSCEGPARVERHERVDQWRRRMSRAGFQAAPMKMINQAQKW
LKNNKVCEGYTVVEEKGCLVLGWKSKPIIATSCWKC*
>20816852_peptide|Cclementina|Ciclev10018153m.g|Ciclev10018153m
MINSVCGTVGTLKSENSSSSIKIHPTSPNESSISETKRTPQSSDLEQPSSLTPPSL
NFPPLKFEIDGDVEVQSPDCSIWETFFSDHLDGDFMISSPVRNLPSPRTSAYNN
NYNYNYAQSMQGQSLSGCSPPRFSAQVGAFSSSLRAKGQSPLHKVFNSPGNQ
YMQPENLALTTAIEDFLEDYQKDGYGMFSPMKICGGGGGGGSSPQLIDMPTT
VPATLECLPMTNNNPSRFSGPVSESSSTAGASQHDNDIYQMGSIAAAPLSQQL
QQEQLQERQQTQQQQLQAQQQQQQNLNHSLMVPLPISSEQEQDSGLQLVHL
LLACAEAVAKEDFMLARRYLHHLNRVVSPLGDSMQRVASCFTEALSARLAA
TLTTKPSTSTPTPFSPFPPNSLEVLKIYQIVYQACPYVKFAHFTANQAIFEAFEA
EERVHVIDLDILQGYQWPAFMQALAARPGGAPFLRITGVGATIESAKETGRCL
TELAHSLHVPFEFHPVGEQLEDLKPHMFNRRVGEALAVNAVNRLHRVPSNCL
GNLLAMIRDQAPNIVTIVEQEASHNGPYFLGRFLEALHYYSAIFDSLDATFPPD
SAQRAKVEQYIFAPEIRNIVACEGGERTARHERLEKWRKIMEGKGFRGVPLSA
NAVTQSKILLGLYSCDGYRLTEDNGCLLLGWQDRALLAASAWRC*

Citrus sinensis

>Cs orange1.1g037028m.g
MAPDFQWRHGLCLDDKDNEAMGLDLNYSATACYRCFNLSTLANHSCTWST
SDETRNHKRIKKTSMIDESSGSSSRLCNSLNSLPRLQFRDHVWVYTQRYLAIE
AMEEAALAMMIDKDTEVKEEGNGDGMKLVQQLIACAEAVACRDKAHASAL
LSELRVNALVFGTSFQRVASCFVQGLSDRLALVQPLGAVGVVGSAAKSMAIT
SERDESLSLVYEICPQIQFGHFVANASILEAFEGESLVHVVDLGMTLGLPHGR
QWHSLMQSLVNRSGKVPKRLKITGVGNCSERLGEIGDELKRYADGLKLNFEF
LAVEKSLETLQAKDINVEDGEVLVMNSILELHCVVKESRGALNSVLQRLHQL
SPKVVMLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTKRAKIE
QFYFAEEIKNIVSCEGPARVERHERVDQWRRRMSRAGFQSVPIKMLMQAKQ
WLRKVQTCEGYTIIEEKGCLVLGWKSKPIIAASCWKCS*

Cucumis sativus

>Csa Cucsa.196250
MVLDGDGGSFFSTDFTSVTKEDEDTMGDSGAAHWLSLLDDTTASSRWVISFS
DEFRHKRLKIETESTPTEDGSGNSSNNSSSLSRSNSCDSLSTGFRAHIWTYNQR
YLAAEAVEEAAAAIINAEESAAEEDASADGMRLLHLLVACAEAVACRDRSH
ASILLSELRANALVFGSSFQRVASCFVQGLADRLALVQPLGYVGFGLPIMSRV
DHSSDRKKKDEALNLAYEIYPHIQFGHFVANSSILEVFEGENSVHVLDLGMAF
GLPYGHQWHSLIERLAESSNRRLLRVTGIGLSVNRYRVMGEKLKAHAEGVG
VQVEVLAVEGNLENLRPQDIKLHDGEALVITSIFQMHCVVKESRGALTSVLR
MIYDLSPKALVLVEQDSNHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTR
RAKIEQFYFAEEIKNIVSCEGMARVERHERVDQWRRRMSRAGFQASPIKVMA
QAKQWIGKFKANEGYTIVEEKGCLVLGWKSKPIVAASCWKC*

Eucalyptus grandii

>Eg Eucgr.H04688
MSLGYLQPELYSEDQYDESHVLDLNLSTLECYSVSSTSSLQSIHRCFLDETRN
CKRLKVEPSIGESIESKEIVHRGCSGMRTNSSNSLRKLQFRDHIWAYTQRYLA

VEAMEEAAAAIMQDAMKLVQRLIACAEAVACRDKTHASALLSELSSKALVF
GTSFQRVASCFVQGLVDRLALVQPSGAVGIVGQSAREMVMTNEKEEALRLV
YEICPYIQFSHFIANESILEAFEGESSIHVVDLGMTQGLPHVHQWRNLISSLAKR
PGSSSRCLKISGVGNCEEGLQTIGDKLKHYAKGLGVNFEFLMVQSNLENLQA
EDFSILEGEVLVINSILQLHCVVKESRGALNSVLQILHKLSPKLLVLVEQDSSH
NGPFFLGRFMEALHYYSAIFDALDAMLPKYDTRRAKIEQFFFAEEIKNIVSCE
GPARVERHERIDQWRRRMSRAGFQSAPIKMLTKAKQWLAKFKVCEGYTIVE
EKGCLVLGWNSKPIIAASCWKCS*
>Eg Eucgr.K02752.1
MLVYCDPEEEEEARGKKRPKRTYSFAESIGSNSPRAYTGGFGSGGISRSSSMSS
LDTFPRLHFRDHIWTYSQQYVAVEAVEEAAAAVSVGGGEEEDGTADGMRLL
QLLVACAEAVACRDKLQASSLLSELRSRALVFGTSFQRVASCFVQGLADRLS
LLQPLGTLGFISPAVVAAASGAAAPDEKEEALRLAYKICPYIQFGHFVANTSIL
EAFEGENFVHVVDLGMSLGLPGGHQWRRLIQSLAGRHGQPPRRLRITAIGLS
AEKFCSIGDELSAYALNFGINLDILQLHCVVKESRGALNSVLQMIHNLSPKVL
VLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDVMLPKYDTRRAKVEQFYFA
EEIKNIVSCEGPARVERHERVDQWRRRMTRAGFQLAPIKMMAQAKQWLGK
NKVCEGYTVVEEKGCLVLGWNSKPIVAASCWKC*
Fragaria vesca
>Fve gene22702-v1.0-hybrid
MAPGHGRRFMTEHECTDHEIRSNRMSDLELHGLNHSSMACYSSYPYMASLD
QEGSANDSSSLINPYSVHESRDRKRLKRTISIDDSLSSITSINGGMMSRSSSTNS
LTTLPRLHFRDHIWTYTQRYLAAEAVEEAAAAMINAEENGAEQEDGTADGM
RLVQLLIACAEAVACRDKSHASALLSELRANALVFGSSFQRVASCFVQGLAS
RLALVQPLGAVGFVPSKLNAKDYALMDKKEEALRLVYEICPHIQFGHFVANS
SILEAFEGESFVHVVDLGMTLGLPHGDQWRGLIHSLTTRAGQPPVSRLRITGV
GLCADRLQIIGEELEAYADSLGLSVEFTCVESNLENLKPEDIKLYEGEVLVVNS
ILQLHCVVKESRGALNSVLQMIHELSPKVLVLVEQDSSHNGPFFLGRFMEALH
YYSAIFDSLDAMLPKYDTRRAKMEQFYFAEEIKNIISCEGPARVERHERVDQ
WRRRMSRAGFQGAPIKMMAQAKQWLGKNKVCEGYTVVEEKGCLVLGWKS
RPIVAASCWKC*
Glycine max
>Gm Glyma17g13680
MAPLYSAFNGEIDGENLPLAFDIWATNLWAYGYYPHQPAISENSSSKLVDFPF
CDGTIIRDNKRVKRTVCFPIYNSVSCHSFFNTNSSSRNSIPKLHFRDHIRTYTQR
YLAAEPVEEASEDTNSSESSGGEEDGCADGVRLVQLLIACAEAVACRDKSHA
SILLSELKANALVFGSSFQRVASCFVQGLTERLNLIQPIGSAGPMMAPAMNIM
DAASDEMEEAYRLVYELCPHIQFGHYLANSTVLEAFEGESFVHVVDLGMSLG
LRHGHQWRALIQSLANRASGERVRRLRITGVGLCVRLQTIGEELSVYANNLGI
NLEFSVVNKNLENLKPEDIEVREEEVLVVNSILQLHCVVKESRGALNSVLQMI
HGLGPKVLVMVEQDSSHNGPFFLGRFMESLHYYSSIFDSLDVMLPKYDTKRA
KMEQFYFAEEIKNIVSCEGPLRMERHERVDQWRRRMSRAGFQAAPIKMVAQ
SKQWLLKNKVCEGYTVVEEKGCLVFGWKSRPIVAVSCWKC*
>Gm Glyma05g03020

MAPPPYSAYNGEIDGENLPLAFDIWENLWAYGYYPHQPISEINSTSTLVDFPFC
DGTIVRDNKRVKRTVCFPIYNSINCHSFFNTNNSSSRNSIPKLHFRDHIRTYTQ
RYLAAEPVEDTNSSESSGGEEDGCADGVRLVQLLIACAEAVACRDKSHASILL
SELKANALVFGSSFQRVASCFVQGLIERLNLIQPIGPAGPMMPSMMNIMDVAS
DEMEEAFRLVYELCPHIQFGHYLANSTILEAFEGESFVHVVDLGMSLGLRHG
HQWRGLIQNLAGRVGGERVRRLRITGVGLCERLQTIGEELSVYANNLGVNLE
FSVVEKNLENLKPEDIKVREEEVLVVNSILQLHCVVKESRGALNSVLQMIHGL
GPKVLVMVEQDSSHNGPFFLGRFMESLHYYSSIFDSLDVMLPKYDTKRAKME
QFYFAEEIKNIVSCEGPLRMERHERVDQWRRRMSRAGFQAAPIKMVAQAKQ
WLLKNKVCEGYTVVEEKGCLVLGWKSRPIVAVSCWKC*
Gossipium raimondii
>Gr Gorai.006G270800
MSHGFLPNEFQKSDKIDEVIGLDLTLSAMGMTFCSHRFMPLIVDNDDNNWSR
RFFEEETRDNKRLKRTINGGNSDGGMSRNGSNGSLSSLSSLHFRDHILTYSKR
YLAAEAVEEAAAAAAAMAAGYEENGIEEDETAYGMRLVQLLIACAEAVACR
DKSHASALLSELRANALVFGSSFQRVASCFVQGLADRLASVQPPGTAVGRVQ
LPPVMNIMDIMSDSDKKQEAFRLVYEICPHIKFGHFVANLAILEAFEGESLVH
VVDLGMTLGLPQGHQWRHLIQSLAAIRGGAGKPHCTRLRITAVGLCVDRFD
VIGKDLKTYAKSMGINLEFNIVESNLENLKPENIKVSDDEVLVVNSILQLHCV
VKESRGALNSVLQMIHELSPKLLVLVEQDSSHNGPFFLGRFMEALHYYSAIFD
SLDAMLPKYDTRRAKMEQFYFAEEIKNIVSCEGADRVERHERVDQWRRRMS
RAGFQATPLRLMSQAKQWLGNNKVCEGYTVVEDKGCLVLGWNAKPIVAVS
CWKC*
Lotus japonicus
>LjRAD1(chr4.CM1864.540.r2.m)
MSPPLYSVLLEDENSVFLLDLDLSSPMGFHAYPHLPILDSSIANWSLPFSISDET
FRESKKLKRTMIPISSADFSISSSSSLSVSVNSIPRLNFRDHIRTYKRYLAAEELP
EDTNSSESVVGAEEDGCADGMRLVQLLIACAEAVACRDKAHASMLLSELKS
NALVFGSSFQRVASCFVQGLAERLTLIQPIGSGAGVSQSMMNIMDAASEEME
EAYRLVYETCPHIQFGHFVANSTILEAFEGESFVHVVDLGMSLGLPHGHQWR
GLIHSLANRASGHGRVRRLRITAIGLCIARLQAIGDELSDYANNLGINLEFSVV
QKNLENLQPEDIKVNDDEALVVNSILQLHCVVKESRGALNSVLQMIHGLSPK
VLVMVEQDSSHNGPFFLGRFMESLHYYSAIFDSLDAMLPKYDTKRAKMEQF
YFAEEIKNIVSCEGPLRMERHERVDQWRRRMSRAGFQAAPIKMVAQAKQWL
LKNKICDGYTVVEEKGCLVLGWKSKPIVAASCWKC
>LjRAM1(chr1.CM1852.30.r2.m)
MINSMCGSSVSLKSENSRNKPQPTSPNESVLQSKKNATQSSADLEQTSLNLTP
PSLNLPALKFDLDGDVEVQSPDSSMWESFFSDHLLDGDFMISSPVRNNVPSPQ
ASTFNSNYNYAHQGIQSQSLSGCSPPRFSSPLGAFNSNKGKGLSPLHRVFNSPN
NQYMQHVENLALPAIEEFLEEYQGDHGLGGGGGYSNSSNKVSSDIGSSSECF
DMQNHIPSMLDSLTMQNSSRYCGSVSEDSSVHGGSSQLSQDSDFYHQMGSM
ASASLSQALQQERYQEKQQKHQTQQQQPQQQQNLTVPIPIGMDQEQDS
GLQLVHLLLACAEAVAKEEYMLARRYLHHLNRVVTPLGDSMQRVAACFTES
LSARLAATLTTKPQSISNGTSMPRSSSSSCLSPFPSNSIEVLKIYQIVYQACPYV

```
KFAHFTANQAIFEAFEAEERVHVIDLDILQGYQWPTFMQALAARPGGAPFLRI
TGVGPCIDSVRETGRCLTELAHSLRIPFEFHPVGEQLEDLKPHMFNRRVGEAL
AVNTVNRLHRVPGSHLGNLLSMIRDQAPNIVTLVEQEASHNGPYFLGRFLEA
LHYYSAIFDSLDATFPPESAQRAKVEQYIFAPEIRNIVACEGAERIERHERLEK
WRKIMEGKGFRGVALSPNAVTQSRILLGLYSCDGYRLTEDKGCLLLGWQDR
AIIAASAWRC
> chr3.CM2163.240.r2.m
MIEEEQLSFITLFSSVHSVSFSVSLMKLSFGLGSPYAWLKDMNEEDRGKYLAR
IINACAMFIESGSIEYADNALEYLAHIASPYGDAMQRVATYVSEGLAFQVLKN
LKGVPKALSLPKTLSTSEELLVKNLFFQFYPFLKIAYVISSHAIAEAMKGEEVI
HVIDLSASEAAQWIYLMQRLKEQQKKPLLLKITAIHEKKEVLEQMALHLGVE
AQKFNFHLQFNAIVSSLENLDLESLPVEKGEPLAISSVLQLHSLLATDDSMSSC
RNPPSESMNQSTFKEVLGKKKINQSSGSSLLSFPICASSPKMEGFLNGLRKLQP
RLVVITEQESNVNGSSLTERVDKALDFYGALFKCLESTFLGTAVERIMLERTL
LGDEIKNIVACEGVERKERHEKLETWIPRLELAGFGRETIRYDWMMQAKKEL
QGYGNGYKLLQENKCLFVCWNDKPLFSVSAWKVY
> chr3.CM0127.880.r2.m
MEDKGLKLVHLLKDTAVFTESCNFIDADIGLYYISQFATPEGDSMQRVATYFS
EALACCQVSKNLRGVPKVLSLSTKLSTPEEQLVRNFFFELYPFLKIAYKTTNQ
AIIEAMGQEKFINILDLSACDATQWIYLMKSLKEHLPDPPDVKIKVTCIHEKYE
VLEQMGLHLRLEAERLNFDFKFNAVVSTLENLDLEKLPLEKGEPLAISCVLQL
HSLLATSDEMVRTMNYAPAEASMNQYAQMLGKQNKNKNKINSSPDSALSPL
SLSPPKMECFLNGLWKLQPKVMVITEQEANVNGSTLTDRMENALQFYGALF
DCLEATFPRTLVDRTLLEKMLLGKQIKNIIACEGVERKERYEVVRTWIPRLQL
AGFGMVSISPNGMIQAKTLLQNYVGGYHTVQDKNCLFMCWEGRPLFSISAW
KF
> chr3.CM2163.230.r2.m
MRELRYDSQGLKNPISLLIECAKCVASGSIKNADIGLELISQISSPDGDAVQRM
VTYFSEALGYRIVKHLPGVYKALNSSKTSSSSEDFLVQKYFYELCPFLKFSYLI
TNQAIVEAMESENWVHIIDLHCSEPAQWVNLLLTLKNRHGGPPHLKITGIHEK
KEVLDQMSLHLTAEAGNLDFPLQFNPIVSKLEDVDFENLPVKTGEALAITSVL
QLHSLLATDDDMSGKFSPAGAASMNLQRAVHMGQRTFAEWLERDMINAYIL
SPDSALSPLSFGASPKMGVFLNAMRKLQPKLMVITEQESNLNGSNLMERIDRS
LYFYSALFDCLDSTVMRTSVERQKLESMLLGEQIKNIIACEGVDRKERHEKLE
KWIRRLELAGFVKVPLSYNGRIEAKSLLQRYGHKYKFREENDGLLICWSDRP
LFSVSAWSFRK
>LjNSP1( chr3.CM0416.1260.r2.d)
MTMEPNSTTSDHILDWLEGSVSFFPSFLDEPCNNSGYIQEYHLWDQYQTDAN
TGSSPNATNSTTATIVATSATSTTSLEPCGSNNNQPLSDLPKKRNATDESSLKP
PQNKNKRIKTRPMNEPENGDAVRKNKKGGAKANGSNCNSGNSKEGRWAEQ
LLNPCAAAIAGGNVNRVQHLLYVLHELASPTGDPNHRLAAHGLRALTHHLSS
SSSSPTSSGTITFASTEPRFFQKSLLKFYEVSPWFSFPNNIANASILQVLAEEANI
TSRTLHILDIGVSHGVQWPTLLDALSRRSGGPPSVVRLTVVTAENDQNMETPF
SKAPPGYNYYPRLLGYAQSIKINLQINRIENHSLQTLNAQSISASPDEILIVCAQ
```

```
FRLHHLNHNSPDERSEFLKVLGNMEPRGVILSENNTECCCSGCGNFAAGFTRR
VEYLWRFLDSTSSAFKGRESDERRVMEGEAAKALTNQREMNEEKEKWCGR
MKEAGFAGEVFGEDAVDGGRALLRKYDSNWEMKVEEKNTSVGLWWKGQP
VSFCSLWKLDGNDQGGTS
>LjNSP2( chr1.CM1976.90.r2.m)
MEMDIDCIHHLDFSGHSTLTNTPSSDNDNYGCSWNHWSPVVNWDAFTGNQD
DFHHLIDSMIDDNNTGPAFSDHTASTTSEEEEEEEATTTTMTTTTTTTTTPEA
ADDDFKGLRLVHLLMAGAEALTGANKNRELARVILVRLKELVSHTDGTNME
RLAAYFTEALQGLLEGAGGAYNSSSKHHVIGGPHHEPQNDALAAFQLLQDM
SPYVKFGHFTANQAIVEAVAHERRVHIVDYDIMEGVQWASLMQALASNPNG
PHLRITALSRSGVGRRSMATVQETGRRLTAFATSLGQPFSFHHSRLESDETFRP
AGLKLVRGEALVFNCMLNLPHLTYRSPNSVASFLTAAKALRPRLVTVVEEEV
GSALGGFVERFMDSLHHFSAVFDSLEAGFPMQGRARALVERVFLGPRIVGSL
ARIYRTGGGGEERGSWREWLRAAGFSGVAVSSANHCQSNLLLGLFNDGYRV
EELGSNKLVLHWKTRRLLSASLWTCSSESDCA
> LjT15C06.70.r2.m
MDMEIDIDDLLDFYGHSSTTPSSDDGGNWVPCSPLVDWDSFTAAADHDDFH
HIIDSFMDYTLPAGDLTPEEESIGERSATTTEEEDDEATAADHSGKGLRLVHLL
MAAAEALTGANKSHDLARAILIRLKELVSHTANTNMERLAAYFTDALQGLL
DGAAGAHNLNKNSVVAGPHREDPQTDMLAAFQLLQDMSPYIKFAHFTANQ
AILEAVAHERRVHIIDFDVSEGAQWASLIQALSSRKDGPSGPHLRITALSRSCG
RGGGRRSIATVQETGRRLTAFAASVGQPFSFHQCRLDPDETFRTSSLKLVRGE
ALVFNCVLHLPHLNYRAPDSIASFLSGARELSPKLVTLAEEEVGPVGDAGFVG
LFMDSLHRYSAMCDSLEAGFPMQRWARGLVERVFLGPRITSSVARLYRTGEE
DEEGSGSWGEWLVASGFRGVPISFTNHCQAKLLLGLFNDGYRVEELSNNKLV
LSWKSRRLLSASVWTSSDDSNL
> chr5.CM0239.240.r2.m
MQLLGGLTNTSNPPPTTEHEKEDHFSKEQNQQQQQHDHFSFDNLMFQDFYF
EPALEPKAAFQEPKEQKNRGVESSQFFGFNKHLVGSNKQVPPSSLASLDLLKN
YGSRFKRPRGQNILSKTKTCASPLKLSTEDILRVAGARFIQHSSSTQWQDNLC
TPMHPYANGFSGFSEEEIKDIELAQLLLAAAERVSCQQFERASFLLQNFQGNS
PHHGSAVQRVIFHFAQALQERINKENGTTLKGSDERVKERHLIELLEVDTNITI
LCHQKMPFYQVMQFAGVQAILENVASGTKIHLININIGCGVMCTGLMQALSE
RQENPVEILKITAVGVANKTKIEETGKRLMSFAESLNLPFSYNMIFVTDMMEI
KVEQFDVEDDEAVAVYSPYFLRSLVSDSDSLENLMKVLRKIKPNIMINLGVE
ANHNSPSFVNRFVEALFYYSAYFDCLDTCMKEEHECRMRIEEMLSEKIRNIVA
MEDRERTVRSVKIDVWRRFFARYRMVETRLSESAVYQANLVAKKFACGNSC
TIDKNGKCLLIGWKGTPIHSISAWKFP
> LjSGA.122341.1 partial
QGLDLVHMLLACAEAVGCRDTQQAEFLLSRIWALASPSGDSLQRVSYCFATG
LKCRLLLLPQSAIANGTFTRSSMNVPLITRENKLEAFHLLYQTTPYIAFGFMAA
NEAIFQASKGKSSIHIVDLGMEHTLQWPSLIRALVSRPEGPPKLRITGLIRIEDN
NPKLQTRINALLEEASSLGMP
> LjSGA.020219.1
```

```
MASRVMNSMLGVCSPLIDHRTIHSALQVFNNMSPFIKFAHFTSNQAILEAVNR
CGSIHIIDLDIMQGLQWPPFFHILATRIEGPPEVRMTGMGSSMELLVETGKNLT
NFARRLGISLKFNPVVTKFGEVDVSILKARPGETLAVHWLQHSLYDATGPDW
KTLRLLEELEPRIITLVEQDVNHGGAFLDRFVGSLHYYSTLFDSLGACLHSDD
DRRHSVEHGLLSREINNILAIGGPARSGEDKFRQWRSELARNCCFVQVPMSV
NSMAQAQLILNMFSPALGYSLAQVDGTLRLGWKDTSLYTASAWTCGGGAS
> LjSGA.045015.1 partial
GDTEQLSNESNTFHFESLLSLLNPSVLNLREDEALVINCQNWLRYLSDDDRNS
NSLRDAFLSLIKGLSPQIVLLVDEDCDLSASSLASRITACFNHLWIPFDALDTFL
PKDSSQRTEFESDIGQKLINIIGFEGHQRIERLESGVMMSQRMKNAGYFSVPFC
DETVKEVKGLLDEHAGGWGMKRDEGMLVLTWKGNNCVFVTAWVPSGMR
DHVGMDATL
> LjSGA.073109.1
MKAELKPPTSISFQNPTTTPLFNTPHTSTPLSGALIGCLGSLDGACIEKLLLHCA
SALESNNITLAQQVMWVLNNVASPQGDTNQRLTAWFLRALISRASRICPSVM
SFKGSNNIPRRSMTVTELTGYVDLIPWHRFGFCASNNEILKAVTGFQRVHVLD
FSITPCMQWPTFIDALAKRPEGPPSLRITVPSFRPQVPPLVNISIHEVGLRLGNF
AKFRDVPFEFHVIGDENLLPSEQLSNESTRFQFESLLSLLNPSVLNIREDEALVI
NCQNWLRYLSDDNRNSNSLRDVF*V*LKS
> chr3.CM0106.470.r2.d
MDITPYTSKKTHFNHHHHEAPIDLHRINHPQHSNPTSTSRSSESSETCDDGKW
ASKLLRECATAISERNSTKTHHLLWMLNELSSPYGDPDQKLASYFLQALFCK
ATDSGFKCYKTLSSVAEKNHSFDSARKLILKFQEVSPWTTFGHVASNGALLE
ALEGETKIHIVDISNTLCTQWPTLLEALATRNDETPHLKLTVVLLSNKIVGSV
MKEVGQRMEKFARLMGVPFEFNVVSGLRSLTKEGLGVQENEAIAVNCVGAL
RRVEVKERESVIGMFKSLSPKVVTFVEEEGDFVTSRDDFVKNFEECLKFYSLY
FEMLEESFPPTSNERLMLERECSRSIVRVLACNDGDEDGGGGGDQGGDCCCE
TRERGTQWSERLGSAFSPVGFSDDVVDDVKALLKRYQAGWSLVVPQGNDH
NLSGIYLTWKEEPVVWASAWKPQLAT
> chr2.CM0018.140.r2.m
MESMLQEEGSSSVTSSPLQFFSMMSLSPSIGSPYPWLRELKSEERGLYLIHLLL
TCANHVAAGSLENANITLEQISQLASPDGDTMQRIAAYFTEALADRILKTWPG
LHRALNSTRIVMVSEEILVQKLFFELFPFLKVAYILTNQAIIEAMEGEKMVHIID
LNAAEPAQWIALLQVLSARPEGTPHLRITGVHQQKEILDQMAHKLTEEAEKL
DIPFQFNPVLSKLENLDFDKLRVKTGEALAISSIMQLHSLLALDDESGRRKSPL
LSKHSNGIHLQKVLLMNQNTLGDFLKKDMVNSYSPSTDSASSSPVSSTTSMN
AESFLNALWGLSPKVLVVTEQDSNHNGSTLMERLLEALYSYAALFDCLESTV
SRTSLERIKVEKMLFGEEIKNIIACEGAERKERHEKLDKWLQRLDVSGFSNTPL
SYYGMLQARRFLQSYGCEGYRMREENGSVVMCWQDRSLFSTTAWRPRK
> chr3.CM0460.10.r2.d
MLQSLIPRSPLNSSKPTTMKTKRDDGGGGSDSAVDEHSFKRRNFSGEKAIEEG
EAAHEEDQVLQPASSSGEEESTGLKLLGLLLQCAECVAMDNLDFANDLLPEI
AELSSPFGTSPERVGAYFAQALQSRVVSSCLGTYSPLMAKSVTLTQSQRIFNA
FQSYNSVSPLVKFSHFTANQAIFQALDGEDRVHIIDLDIMQGLQWPGLFHILAS
```

```
RARKIRSLRITGFGSSSELLESTGRRLADFASSLGLPFEFHPVEGKIGSVTELSQ
LGVRPGEATVVHWMHHCLYDITGSDLGTLRLLSQLRPKLITTVEQDLSHAGS
FLARFVEALHYYSALFDALGDGLGADSLERHTVEQQLLGCEIRNIVAVGGPK
RTGEVKVERWGDELKRVGFRPVSLRGNPAAQASLLLGMFPWRGYTLVEENG
SLKIGWKDLSLLIASAWQPSDLITYHD
> chr5.CM1667.220.r2.a
MREQMLRQDQRRKDENNGLPLIHLLLTTATSVDETNLDASLENLTDLYQTVS
LTGDSVQRVVAYFADGLVAKLLTKKSPFYDMVMEEPTSDEEFLAFTDLYRVS
PYYQFAHFTANQAILEAFEKEEQKNNKSLHVIDFDVSYGFQWPSLIQSLSEKA
TSNNRISLRITGFGKNMKELQETESRLVSFSKGFHNLVFEFQGLLRGSRVINLR
KKKNETVAVNLVSYLNTLSCFMKVSDTLGFVHSLNPSIVVLVEQEGGRCSRT
FLSRFTDSLHYFAAMFDSLDDCLPLESAERLRIEKKLLGKEIKSMLNYDVDGV
GDCPKYERMETWKLRMENHGFGGMKISSKSMIQAKLLLKMRTNYCPLQFEE
DGGGGGGFRVSERDEGRAISLGWQNRFLLTVSAWQPL
> chr6.CM0139.1440.r2.d partial
LAESLLSCAEKIGYQQYERASKLLSHCKSLTSNRGGPVKRVVHYFAEALQHM
IDKISMFTGVQAIIENVTEAKKIHLIDLEIRKGAQWTTLMHALESRHDCPLELL
KITAIGSGTTSRHEIEDTGEKLKVFAKSLNIPFSFNVVIVSDMVDIITEDVFVIDP
EETVAVYSQFAIRCKIQEPNQLEAIMRMVRTLKPVVMVVAEIEANHNSTSFV
KRFTEALFYFSAFFDCLEACMKHDEQNRIMIETLLGHGIRSIVAERKSRNVKID
VWRAYFSRFGMEETELSRVSLYQADLVAKRFPSGSCCTFHMDGHCLLVGWK
GTPISSVSVWKFT
> chr4.CM0680.320.r2.m - phase: 0
MNVMDYEQFDCSFTPSFMNQLHECTLENVNANAFPIQEDDLLSPNSILDEMF
DHDQSENLSPPNTILDEIFDQQSMEELLQQHNQDFDVMNLDLGNELCHGVSQ
ITQKNVHVSKEADKDSYLNGIQAELMEETSLVDLLLTAAEAVEAQNWYLASE
IIEKLTNDASSLQNGDSLLNRLALFFAQGLFDKSSSPTPKLQYYCDPVSTQTDT
FCVLQILQELSPYVKFAHFTANQAILEATDGVEDVHIIDFDIMEGIQWPPLMV
DLAMRKSTSTTSLRVTAITVDQQSAASAHQTGRRLEEFAASINFPFTFNQVM
MVSEEDLQRIELGHSLIVNCMIHQWMPNRSFSLVKTFLDGVTKLSPKLVVLV
EEELFNFSKLKSMSFVEFFCEALHHYTAISDSLASGMWGSHKVEFALIEKQVL
GLRILDSVRQFPCEREERMLWEEGLFSLKGFKRVGVSSSNISQAKFLVSLFGK
GYWVQFEKCRLALCWKSRPLTVASIWVPMTNLSDKVK
> chr1.CM0122.2130.r2.d partial
EEENGDVELVQFLLAAAERVSCQQFERASRLLLHYQWNSSSYASGNTVQRV
VFHFAEALKERIDRETGRITRVMQFTGVQAMVEHVAGETKIHLIDLDIKCGVQ
CIALMQALAERQNSMVKVKFFKISAIGLTACKTKIEDTGKRLASFAESLNLPFS
YKHILVTDMAEFKDHFEVEEDEAVIIYSAYFLRTMISRPDCLENLMRIIRNIKPS
IMIVLEVEANHNSPSFGNRFIEAMFYYSAFLDCLYTCIEDDECRVFTEAILSAGI
RNIVAMEGRERRVRNVKIDVWRRFFARYRMVEIEFSESSMYQADLVAREFA
VGKFCTVDKNGRSMIVGWKGTPMHSISAWRFL
```

Lunularia cruciata

```
> Lc 6169
```

MINHPLRKPPKEMEKDEGFKLMHRATPYVAFGHYAANSAILNAFQGEDTLHI
VDLGMTHGLQWPYLINRLSSREGGPPRLLRITGLGLATDMAHETGRELTACA
KLHRIPFEFRVVAESLENLQPSMLELREGEALAVNSVLQLHCVVKESRGSLNT
VLQHIHELCPRVLTLVEQDASHNGPFFLGRFMEALHYYSAIFDSLDASLPRYS
HQRVKIEQFHFGEEIKNIVSCEGPARVERHERVDQWRRRMSRAGFQPLPSKFL
HDARMWLTLYACDGYTLAEEKGCLVLGWRGKSLVAASSWRC*
> Lc 17330
MAHSPLSWIMKAQLGGLLEDDFDSEPTSVLDLTQCISPPLQTDVWAEKQAAM
LEEQQLLAWQHQQNILYEQQRWEEQQQLCIWEQNESEHGFAELQYLYEHDP
MNTCNDFSYQSDHCGFQDNQQQGLPFESFDDSQFMLDSWISSSGNHDGLNSI
VDDSIPSSYQKAFTIPSPADMQVELDSKFQADVSDVRFVKEEDEAKHSMLTN
LPSNMSNMSNIIPRVNLYQPQDIIDQWYVPSIEDTMCQMQPSVSKLQFQTHPY
PSPLPGNGNQEALPKSSMDAEPHLPSLNPTLEKSASHKPSAGDESCMSDSLNS
VAKQIVPATPVPRKATSNGSNRSVTLAELASGGVKMGGVQLPNVGTQIPSVIR
LNLETQASGMGGVRMVQLLLACAEAVACRDSRQANILLQQLQHMATPYGD
AMQRITVCFVEGLTARLAITGSQPFKFRKAQGPRKPPSEREKLKAFQLVYRAS
PFLAFGHLAANSAILEAFEKEDRLHIVDLGMSHALQWPYLIHDLANRQGGPP
RSLRITGFGLSSAQLKEAGEELVEIAKSRNIPFEFRAVTESLESMQPSMLELRE
GEALAVNSVLQLHCVVKESRGSLNAVLKTIHELSPKILTLVEQDASHNGPFFL
GRFMEALHYYSAIFDSLDVILPRNCPYRVKMEQFHFAEEIKNIVSCEGPARVE
RHERVDQWRRKMSRANFQPLPLKFLNQARMWLSHYNCDGYTLSEEKGCLVI
GWKGKPIIAASSWKG*
Lupinus angustifolius
> Lang AOCW01143302.1
FCPARMDQEQGRGGELQGQESVLDEMIELCSSDAATNAPTCLASSEVDEFVD
SFINMDQYQYDNEDQVSLEKHQSFDHFLVKDEGDTYAFSKKMMNDALGYV
ATAMEGEESEIYENITIAMEEGEVEMHCVEEAGLGIDKGLDLVHMLLACAEA
VSCRDTQQAELLLGKISVLATPSGDSLQRVSYFFATGLKCRLSLLPHNVFANG
TLTTSTIAMNVPLISREDKLEAFQLLYQATPYMAFGLMASNEAICQASQGKSS
IHIIDIGLEHTLQWPSLIRALASRPEGPPKLRITGLIANEDNPKLQISINVLVEEA
RSLGIPLEFHIISEPATTSVLTKEKLNLREGEALFVNSILQLHKYVKESRSYLKTI
LQSIKKLGPTALIVVEQDTNHNGPFFLGRFLESLHYYSAIFDSLEASMPRKSKD
RMKIEKLHFAEEILNIVAYEGSDRMERHERVDQWRRKLGRTGFQVMPLKCTS
QVRMMLSVYDCDGYTLSCEKGTLLLGWKGRPIMMASAWK
Malus domestica
> Mdo MDP0000431628
MALSRNFLPEEYMINHERTNKMSVLDLITHSAASAMPCYPYPHTPLLEEGGT
ASTPPXXDESGNHKRLKRTTSISDSMSSHNSLYSGGGSGDSCITDISLSRRSST
NSLNTLPRLHFRDHIWTYTQRYLAAEAVEEAAXXMSNAEGNEEEEDXTXXG
MRLVQLLIACAEAVACRDKSHASALLSELRANALVFGSSFQRVASCFVQGLA
NRLALVQPLGAVGFIGSAMSTKDFALDKKEEALRLVYEICPHIQFGHFVANSS
ILEAFEGESFVHVVDLGMTLGLPHGDQWXGLIESLATRTGQPPTRLRITGVGL
YGDRLQIIGEELEAYADSLGINLEFSVVXSNLENLRPXDIKMYDGEVLVVNSIL
QLHCVVKESRGALNSVLQMVHELSPKILVLVEQDSSHNGPFFLGRFMEALHY

YSAIFDSLDAMLPKYDTRRAKMEQFYFAEEIKNIISCEGRQGGEAREVDQWR
RRMSRAGFQAAPIKMLAQAKQWLGKIKVCEGYTILEEKGCLVLGWKSKPIV
AASCWKC*
> Mdo MDP0000464809
MPPLSLKLHSLFFKYHLRHQLHNLTQSAQLQNTDPKFGITSRPEEPVVPANPT
FQNGVATKNIHIDPNSSLSLRIFLPDTVLPLKAPNPTSRVGALLSPSPXCSNSDD
GVVYRGYSPDQLVGRRHRKVPIFLQFHGGGFVSGSNDTSWNDAFCRRMAKL
CDAIVVAVGYRLAPESPYPAAFEDGVTVLKWVAKQANLALVQKGRSRIFDSF
GSSMVEPWLAAHGDPSRCVLLGVSCGANLADYVARKAVEAGDLLDPIKVVA
QVLMYPFFIGSTPTRSEIKLANSYLFDKATCMLAWKLFQTEEEFDLDHPAGNP
LMPAGRGPPLKTMPPTLTVVAQHDWMRDRGIAYSEELRKANVDAPLLDYKD
TVHEFATLDVLLETPQAKACAEDITIWYKKEASTYTSHISHIVYISPQGDSIGQ
ASNTLESYTMASDLYVQPEFCGEITSNEVIGFDLNLSSMSCLPHPSLSALEDDS
ANWVSPFVDKTRNHKRLKQEHDANYGFKVGTNYSSFYTGSDSNMCTSGFTS
LPTIQFRDHISAHKRRYLAAEAMEEAFATLMRDKEDESEEGGGSEDEIKLVQQ
LIACAEAVACRDKAHASALLFELRANAQAFGTSFQRVSSCFVQGLAXRLALV
QPLGAGGVIDPNVKSKPFSAEKGPFSAEKDEALHLVYEICPQIQFGHFVANASI
LEAFEGGSSVHVIDLGMTLGLQHGYQWRDLIDRLANRPGQPLHXLRITGVGS
SSERLQAIGDDLKLHAQSMKINXDFSEVESNLENLKPQDFHLVDGEILIINSILQ
LHCLVKESRGTLNSALQTLHELSPKLMVLVEQDTSHNGPFFLGRFMEALHNY
SAIFDSLDAMLPKYDTRRVKMEQFYFGEEIKNIVSCEGPARVERHERVDQWR
RRMRRAGFQPAPLRTMAQAVKWLETNACEGYTVVEDKGCLVLGWKSKPIIA
TSCWK*
Manihot esculenta
> Me cassava4.1_032725m.g
MSPVPYMQSEFCSDNKSYETTFLDLSPFCSVPNFSAVENDSPAWFDLDETRSH
KRIKQEPCSRTRANSLNSVPKLHFRDYIWAYTERYLAIEAMEEAAAAMTVGE
RNEVKDEEGSDGMKLVQQLIACAEAVACRDKKHASALLAELRANALVFGTS
FQRVASCFVQGLSDRLALLQPLGTVGVLAPQANSVNTFKAEKDEALRLVYDI
CPQIQFGHFVANASILEAFEGESSVHIVDLGMTFGLPHGLQWLNLIHSLAKHP
GQQPHRLRITGVGNSAELLQAIGDELDCYARSLGLNFEFLWVESTLENLKPEH
FKLLAGEVVIINSILQLHCVVKESRGALNSFLQILHELSPKLLILVEQDSNHNGP
FFLGRFMEALHYYSAIFDSLDAMLPKYDTRRAKIEQFFFAEEIKNIISCEGPAR
VERHERLDQWRRRMSRAGFQPSPIKTIMQAKQWLEKANFCEGYAVTEEKSC
LVLGWKSKPIIAASCWKCP*
> Me cassava4.1_031685m.g
MAPSLFSCDFSENETMGSISSLDLSLSAMAYYPYLYLPTLENNASICILPFSEDI
REHKRIKRTLSSAESLGSNNNGFYCAGSGNSDCSSISRSNSSNSLNSLPRLHFR
DHIRTYTQRYLAAEAIEEATEAMENTDEGGDEEDGSTDGMRLVQLLIACAEA
VACRDKSHASTLLSELRRNALVFGSSFQRVASCFVQGLADRLSLVQPLGTVG
FTASMMNIMDISSDKKEEALSLVYEICPHIQFGHFVANSSILEAFEGESFVHVV
DLGMTLGLPHGHQWRQLIQSLASRAGKPPRRLRITGVGLCVGRFQTIGDELV
EYAKELGINLEFSVVESNLENLRRDDIKVFDGEVLVVNSILQLHCVVKESRGA
LNSVLQIIHALSPKVLVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDTMLPR

YDTRRAKMEQFYFAEEIKNIVSCEGPARVERHEKMDQWRRRMSRAGFQAAP
IKMMAQAKLWLAKNKVCEGYTVAEEKGCLVLGWKSKPIIAASCWKC*

Medicago truncatula

> MtDELLA1
MKREHQESFGGGVISNNNKTNTNHLNSSKNINFGECSSMQNTNTKQNMWRE
EKETNGGGMDELLAALGYKVRSSDMADVAQKLEQLEMVMGSAQEEGINHL
SSDTVHYDPTDLYSWVQTMLTELNPDSSQINDPLASLGSSSEILNNTFNDDSE
YDLSAIPGMAAYPPQEENTAAKRMKTWSEPESEPAVVMSPPPAVENTRPVVL
VDTQETGVRLVHTLMACAEAIQQKNLKLAEALVKHISLLASLQTGAMRKVA
SYFAQALARRIYGNPEETIDSSFSEILHMHFYESSPYLKFAHFTANQAILEAFA
GAGRVHVIDFGLKQGMQWPALMQALALRPGGPPTFRLTGIGPPQADNTDAL
QQVGWKLAQLAQTIGVQFEFRGFVCNSIADLDPNMLEIRPGEAVAVNSVFEL
HTMLARPGSVEKVLNTVKKINPKIVTIVEQEANHNGPVFVDRFTEALHYYSSL
FDSLEGSNSSSNNSNSNSTGLGSPSQDLLMSEIYLGKQICNVVAYEGVDRVER
HETLTQWRSRMGSAGFEPVHLGSNAFKQASTLLALFAGGDGYRVEENNGCL
MLGWHTRSLIATSAWKLPQNESK
>MtDELLA2
MKREHKLEHEDMSSGSGKSGVCWEDDGGGMDELLAVVGYKVKSSDMAEV
AQKLEQLEQAMMGNNFHDHDESTIAQHLSNDTVHYNPSDISNWLQTMLSNF
DPQPNNPSVNSDDNDLNAIPGKAIYAADEFTSRKRVKRNESVTVTTESTTTRPI
MVVETQEKGIRLVHSMACAEAVEQNNLKMAEALVKQIGYLAVSQEGAMRK
VATYFAEGLARRIYGVFPQHSVSDSLQIHFYETCPNLKFAHFTANQAILEAFQ
GKSSVHVIDFSINQGMQWPALMQALALRPGGPPAFRLTGIGPPASDNSDHLQ
QVGWRLAQFAQTIHVQFEYRGFVANSLADLDASMLELRSPETESVAVNSVFE
LHKLNARPGALEKVFSVIRQIRPEIVTVVEQEANHNGPAFLDRFTESLHYYSTL
FDSLEGSSVEPQDKAMSEVYLGKQICNVVACEGTDRVERHETLNQWRNRFN
SAGFSPVHLGSNAFKQASMLLALFAGGDGYKVEENDGCLMLGWHTRPLIAT
SAWKLAAANSVVVSH
> MtNSP1
MTMEPNPTSDHILDWLEGSVSFFPSFLDDPYNNGYIHEYEIWNQNQDISNQYQ
IDANTNSSNATNSTTNIVAASTTTSTTSLEPNSFNNIPFSDLPKKRNAEDELSLK
KQPQNQKNKRLKSRPMNESDNGDAALEGTVVRKSGGNKKGAAKANGSNSN
NGNNKDGRWAEQLLNPCAVAITGGNLNRVQHLLYVLHELASTTGDANHRL
AAHGLRALTHHLSSSSSSTPSGTITFASTEPRFFQKSLLKFYEFSPWFSFPNNIA
NASILQVLAEEPNNLRTLHILDIGVSHGVQWPTFLEALSRRPGGPPPLVRLTVV
NASSSTENDQNMETPFSIGPCGDTFSSGLLGYAQSLNVNLQIKKLDNHPLQTL
NAKSVDTSSDETLIVCAQFRLHHLNHNNPDERSEFLKVLRGMEPKGVILSENN
MECCCSSCGDFATGFSRRVEYLWRFLDSTSSAFKNRDSDERKMMEGEAAKA
LTNQREMNERREKWCERMKEAGFAGEVFGEDAIDGGRALLRKYDNNWEM
KVEENSTSVELWWKSQPVSFCSLWKLDKQPE
> MtRAM1
MINSLCGSSNSLKEKCLQPNSSNQTNTHSKKNATNSCGDLEQINVLTPQSLNL
PSLKFDLDGDVEVQSPDSSMWEAFFNDHLDNDFMISSPIRNINPNSPQASTYN
NCNYNYAQGMQIQSLSGCSPPRFASQIGSLNSNNQQKGKGLSPLHRVFNSPN

NQYMQHVENLSLPAIEEFLEDFQGDVDHFSSTKVSSECFDMETPISTILDSLTM
QNSSSYGASVNEESTLLHGGNSSSQISQESDIYHQMGSMASASLSQALQQERY
QEKHQKMQAQQQSLTVPIQIGIEQEQDSGLQLVHLLLACAEAVAKGEYMLA
RRYLHQLNRVVTPLGDSMQRVASCFTESLSARLAATLTTKSSSTKKLAPSSLS
SSSSSSCLSTFPSNPMEVLKIYQIVYQACPYIKFAHFTANQAIFEAFEAEERVHV
IDLDILQGYQWPAFMQALAARPGGAPFLRITGVGPCIESVRETGRCLTELAHS
LRIPFEFHPVGEQLEDLKPHMFNRRVGEALAVNTVNRLHRVPGNHLGNLLSM
IRDQAPNIVTLVEQEASHNGPYFLGRFLEALHYYSAIFDSLDATFPVESAPRAK
VEQYIFAPEIRNIVACEGEERIERHERLEKWRKIMEGKGFKGVPLSPNAVTQSR
ILLGLYSCDGYRLTEDKGCLLLGWQDRAIIAASAWRC
> MtNSP2
MDLMDMDAINDLHFSGHSSLTNTPTSDEDYGCTWNHWSPIVNWDTFTGAPD
DFHHLMDTIIEDRTTVLEQLSPSITTTTTTTTTTDEEEEEMETTTTTTTTAIKTH
EVGDDSKGLKLVHLLMAGAEALTGSTKNRDLARVILIRLKELVSQHANGSN
MERLAAHFTEALHGLLEGAGGAHNNHHHHNNNKHYLTTNGPHDNQNDTLA
AFQLLQDMSPYVKFGHFTANQAIIEAVAHERRVHVIDYDIMEGVQWASLIQS
LASNNNGPHLRITALSRTGTGRRSIATVQETGRRLTSFAASLGQPFSFHHCRLD
SDETFRPSALKLVRGEALVFNCMLNLPHLSYRAPESVASFLNGAKTLNPKLVT
LVEEEVGSVIGGFVERFMDSLHHYSAVFDSLEAGFPMQNRARTLVERVFFGP
RIAGSLGRIYRTGGEEERRSWGEWLGEVGFRGVPVSFANHCQAKLLLGLFND
GYRVEEVGVGSNKLVLDWKSRRLLSASLWTCSSSDSDL
> Mt Medtr4g104020.1
MSPALYASTFKCEVDENPSLMGSYYASLYPNLPILENSATSTWILNPFSDHETE
TIRDHKKLKRSTVTIPIWFANFSSHSNSFFNNINSNNNSVNSIPRLHFRDHIRTY
KQRYFASEAVEEAAEDNFNYNNCEAEEDGSCADGMRLVQLLIACAEAVACR
DKSHASVLLSELKSNALVFGSSFQRVASCFVQGLTERLTLIQPIGNNSAGSDTK
SMMNIMDAASEEMEEAFKLVYENCPHIQFGHFVANSIILEAFEGESFLHVVDL
GMSLGLPHGHQWRGLIQSLADRSSHRVRRLRITAIGLCIARIQVIGEELSIYAK
NLGIHLEFSIVEKNLENLKPKDIKVNEKEVLVVNSILQLHCVVKESRGALNAV
LQMIHGLSPKVLVMAEQDSGHNGPFFLGRFMESLHYYSAIFDSLDAMLPKYD
TKRAKMEQFYFAEEIKNIVSCEGPLRMERHEKVDQWRRRMSRAGFQGSPIK
MVVQAKQWLVKNNVCDGYTVVEEKGCLVLGWKSKPIVAVSCWKC*
Mimulus guttatus
> Mgu mgv1a021462m.g
DDENAPIWLPPLYDESSNFKRIKRTVSLGAFTSSNSSSSNVSSHIPRTSSTSSLSS
VPKLQFRDHIMTYTRRYLTVEEEEEEGASVADLVGSEGGQGGGVAGEGNNA
DGMKLVQLLISCAEAVACRDKSHASVLLSELRSNALVFGSAFQRVASCFMQG
LADRLALVQPLGAVGYVVVPPRTTTTTSVASEKREEALRLVYENCPHIQFGH
FVANSSILEAFEGERLVHVVDLGMTLGLPHGHQWRGLVRSLADRPGPPPRLRI
TAVGLCGDEFRAIGDELEAYAVGLGLNLEFSVVESNLENLDHKDIDVRPGEV
LVVNSILQLHCVVKESRGALNSVLRVIHGLSPRVLALVEQDSSHNGPFFLGRF
MEALHYYSAIFDSLDAVLPKYDTKRAKMEQFYFAEEIKNIVSCEGPARVERH
ERVDQWRRRMSRAGFQPSPLKMLAQAKQWLSSINACEGYTIVEEKGCLVLG
WKSKPIVAASCWKC*

Oryyza sativa
> OsDIP1
METMSYPCSLLIPFSTQFEEISSSSLLLWSPQAEENPHENANMYEFDADHSHD
QIHQDHQFLDMMVIQESANEFDGNHSHDQIHQDHEFLETMVIQESANEFDGD
HSHDQIHQDHEFLEMMAIQESANDLLQLQDDFSVPNADPLAASFEFDERLAV
AGHENGNVVATQEESAGDLLLAGAMAVDAGDAVHASAIMSRLDDLLADIA
GRRSCEATSPVDHLAYYFARGLKLRISGAATPASSPPPPAANWSSPAYRMLQE
LTPFVKFAHFTANQAILEATADDLDVHVVDFNVGEGVQWSSLMLKLLLLGTI
TILQPKLVILIEDELSRISKNPPSPSLAAPPPFPEFFSDAVAHFTAVMESTASCLV
SYDDEAWLSLRRVGEEVVGPRVEDAVGRYGSLAGGAQMMEGLRAREVSGF
SVAQGKMLAGLFGGGFGVVHQEKGRLALCWKSRPLISVSLWCPK
> Os LOC_Os01g67650
MGMSPEPCNSISDCSQQQSHHLTLTQQQDSTIICTNQELDYYYRFYDVDEAAF
DGNEVELVSRFSKVTRMDHMISSPYQPTWSPAQAAVDVVGSSETSRVRKKRF
WDVLESCKQKVEAMEAMDTPATATFRVGAGDGGGGGGGGAGGGGGGADG
MRLVQLLVACAEAVACRDRAQAAALLRELQAGAPVHGTAFQRVASCFVQG
LADRLPLAHPPALGPASMAFCIPPSSCAGRDGARGEALALAYELCPYLRFAHF
VANACMLEAFEGESNVHVVDLGMTLGLDRGHQWRGLLDGLAARASGKPAR
VRVTGVGARMDTMRAIGRELEAYAEGLGMYLEFRGINRGLESLHIDDLGVD
ADEAVAINSVLELHSVVKESRGALNSVLQTIRKLSPRAFVLVEQDAGHNGPFF
LGRFMEALHYYAALFDALDAALPRYDARRARVEQFHFGAEIRNVVGCEGAA
RVERHERADQWRRRMSRAGFQSVPIKMAAKAREWLDENAGGGGYTVAEEK
GCLVLGWKGKPVIAASCWKC*
Phaseolus vulgaris
> Pv Phvul.003G208600
MAPTHCSPFNGEIDGENLSHVFDIWATDHYPYFPHQPISKSSSSTLVLPFCDGT
IRDNKRVKRTVGIPFDWIENLIGNSHSFLNTSANNISNRNRDSIPKLHFRDHIRT
YTQRYLAAEPEEEAPEDTNSSEGGGVEEDGCHDGVRLVQLLIACAEAVACRD
KSHASILLSELRANALVFGSSFQRVASCFVQGLTERLNLIQPTGPVSPMPAMM
NIMDTASEEMEEAFRLVYELCPHIQFGHFVTNSIVLEAFVGESFVHVVDLGMT
LGLRHGHQWRGLMQSLANRKADERVRRLRITAVGLCDRLQSIGDELSVYAT
NLGINFEFSVVKKNLENLEPEDIEVREEEVLAVNSILQLHCVVKESRGALNSV
LQMVHGLGPKVLVMMEQDSSHNGPFFLGRFMESLHYYSAIFDSLDVMLPKY
DTKRAKMEQFYFAEEIKNIVSREGPLRMERHERVDQWRRRMSRAGFQAAPIK
MVAQAKQWLQNSKVCEGYTVVEEKGCLVLGWKSKPIVAVSCWKC*
Populus trichocarpa
> Pt Potri.015G091200.1
MAHSFLSGEFNDENVSETSGLDLSLSAMACYPRPPYLPTFDSNAVSWILPFSD
DTRDAKRMRRSSSLVESIRSNNSSLYSGGSSICRSISTNILNSIPKLHFRDHIWT
YTQRYLAAEAVEEAASEAMINADEGGNEEEGNADGMRLVQLLIACAEAVAC
RDKSHASALLSELRSNALVFGSSFQRVASCFVQGLTDRLSLVQPLGAVGFVPT
MNIMDIASDKKEEALRLVYEICPHIRFGHFVANNAILEAFEGESFVHVVDLGM
TLGLSHGHQWRRLIESLAERAGKAPSRLRITGVGLCVDRFRIIGDELKEYAKD
MGINLEFSAVESNLENLRPEDIKINEGEVLVVNSILQLHCVVKESRGALNSVL

QIVHELSPKVLVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTR
RAKMEQFYFAEEIKNIVSCEGPARVERHERVYQWRRRMSRAGFQAAPIKMM
AQAKQWLVKNKVCDGYTVVEEKGCLVLGWKSKPIIAASCWKCLINSSQ*
> Pt Potri.012G093900.1
MAHSFLSREFNDETLSETSGLDLSLAAMACYPHPYLPIFESNVVSRMLPFSDE
TRDVKRIKQSSSMVESIRSNGSSLYSGGSSICRSSSTNSLNNIPKLHFRDHIWTY
TQRYLAAEAVEEAAAAMINAEEGGNEEEGNSDGMRLVQLLIACAEAVACRD
KSHASALLSELRSNALVFGSAFQRVASCFVQGLIDRLSLVQPLGAVGFVAPTM
NIIDIASDKKEEALRLVYEICPHIRFGHFVANNSILEAFEGESSVHVVDLGMTL
GLPHGHQWRLLIQSLAERAGKPPSRLRITGVGLCVDRFRIIGDELEEYAKDMG
INLEFSVVKSSLENLRPEDIKTSEDEVLVVNSILQLHCVVKESRGALNSVLQIIL
ELSPKVLVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDTMLPKYDTRRAK
MEQFYFAEEIKNIVSCEGPARVERHERVDQWRRRMSRAGFQVAPIKMMAQA
KQWLVQSKVCDGYTVVEEKGCLVLGWKSKPIIAASCWKC*

Prunus persica

> Ppe ppa016418m.g
MSAFDLISHSAMSCYSNPYMPLLEEGNASALTFPYLDESGNHKRLKRTISIAES
MSSHNSLYGGGSNNSCVTNGSISRSGSTNSLNTLPRLHFRDHIWTYTQRYLAA
EAVEEAAAAMINAEGNGAEEDGTADGMRLVQLLIACAEAVACRDKSHASAL
LSELRANALVFGSSFQRVASCFVQGLANRLALVQPLGAVGFIGSPMNAKDFA
LDKKEEALRLVYEICPHIQFGHFVANSSILEAFEGESYVHVVDLGMTLGLPHG
DQWRGLIESLATRAGQPPSRLRITGVGLYGDRMQIIGDELEAYADRLGINLEF
SVVESNLENLRPEDIKLLDGEVLVVNSILQLHCVVKESRGALNSVLQMVHELS
PKILVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTRRAKMEQ
FYFAEEIKNIISCEGPARVERHERVDQWRRRMSRAGFQAAPIKMLVNAKQWL
GKINVCEGYTILEEKGCLVLGWKSKPIVAASCWKC*
> Ppe ppa026722m.g
MASGLYVQAEFSGDNTSDEVIGLDLNLSSMACLPYPSSLSTLEENPAAWVIPFI
DETSSHKRLKQQHSSSYEFNNAGTNYCSLYSGSGSMCARGLDSLPRIHFRDHI
SAYTRRYLAVEAMEEATATLMRGKEGESEEGGRGDATKLVQQLIACAEAVA
CRDKAHASTLLYELRANAKVFGTSFQRVASCFVQGLSDRLALIQPLGAVGLI
GPITKSTAFSAEKDEALHLVYEICPQIQFGHFVANASILEAFEGESSVHVIDLG
MTLGLPHGYQWRNLIDSLANRAGQPLHRLRITGVGNSAERLQAIGNDLKLHA
QSMKLNFEFSAVESSFENLKPQDFNLVDGDVLVVNSILQLHCLVKESRGALN
SVLQTLHQLSPKLMILVEQDTSHNGPFFLGRFMEALHNYSAIFDSLDAMLPKY
DTRRAKMEQFYFGEEIKNIVSCEGPARVERHERVEQWRRRMRRAGFQPAPLK
MIAQAMKWLEINTCEGYTVVEDKGCLVFGWKSKPIIATSCWK*

Ricinus communis

> Rc 29929.t000249
MPSALYMQSESFRDYTRYETTGLDLNIGYSIPHIPVSKNSRPTCFDLDETRSHK
RVKPGGVESIGNGIGCYAVSSNGGSINANCLNSLPRLHFRDYIRAYTERYLAIE
AMEEAAAGLMISKKNEIKEEDIDGMKLVQQLIACAEAVACRDKTHASALLSE
LRANALVFGTSFQRVASCFVQGLSDRLTLLQPLGAVGVLGPAGKTISFTAEKD
EALRLVYEICPQIQFGYFVANATILEAFEGESSIHVVDLGMTLGLPHGEQWRN

LLHCLANRPDKKPRCLRITGVGNSAERLQALGDELDCYARSLGLNFEFLWVE
SSLEKLKSTDFKLLDGEVVIINSILQLHCAVKESRGALNTVLQILHELSPKLLIL
VEQDSGHNGPFFLGRVMEALHYYSAIFDSLDTMLPKYDTKRVKIEQFFYGEEI
KNIVSCEGPARVERHERVDQWRRRMSRAGFQPAQIKMAMQAKQWLGKAKV
CEGYTVTEDKGCLILGWKSKPIIAASCWKCC
> Rc 30147.t000065
MRRRLFPNDLSHDETINETSSLDISLSAMAYYPYPYLPILENNESIFVLDFSDET
REHKKRIKRALSFAESTGSDGIYNTGGSGSGSGSNDTISRSCSTNSLNSLPRLH
FRDHIWTYTQRYLAAEAVEEGAEAMANSEEGENHGEGGNTDGMRLVQLLIA
CAEAVACRDKSHASALLSELRSSALVFGSSFQRVASCFFQGLADRLSLVQPLG
TVSLVTPIMNIMDIASDKKEEALSLVYEICPHIQFGHFVANSSILEAFEGESFVH
VVDLGMTLGLPHGHQWRQLIQSLANRAGKPPCRLRITAVGLCVGRFQTIGDE
LVEYAKDVGINLEFSVVESTLENLQPDDIKVFDGEVLVVNSILQLHCVVKESR
GALNSVLQTIHALSPKILALVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAM
LPRYDTRRAKMEQFYFAEEIKNIVSCEGPARVERHEKVDQWRRRMSRAGFQ
AAPVKMMAQAKQWLGKNKVCDGYTVVEEKGCLVLGWKSKPIVAASCWKC

Selaginella moellendorffii

>Sm 88990
GLQLIHMLLGCGEKIDQEDYIYAGNLLHQLKQLASPTGDSIHRVATHFTDALY
ARLNGTGYRSYTALRAYDPASLEEILGAYHILYQVCPYIKFAHFTSNQAIFEAF
EGEQSVHIIDLEILQGYQWPAFMQALAARQGGAPHLRITGVGMPLEAVQETG
KRLADLAATLRVPFEYHAVGERLEDLQSHMLHRRHGEALAVNCIDRFHRLFT
DDHLVVNPVVRILSMIREQAPRIVTLVEQEANHNTNSFLKRFLEAMHYYSAIF
DSLEATLPQVSPERAKVEQVVFSSEIMNIVACEGSQRIVRHEKVDKWCKIMES
IGFYNVALSPSAVHQSKLLLRLYQTDGYTLVEDKGCLLLGWQDRAIIGASAW
RC*
> Sm 83811
MCSNDVFNPVQREDVLQDKIELKESISALESDGAVGLEFWRGLQHQQEQQEQ
QEQQQQRHAQDQSLFEQEQQQGSRAQPAAAQDHHESGDANVGIRLIQLLLA
CAEAVACRDVNQAATLLSQLQQMASPRGDSMQRVTSCFVEGLTARLAGLQS
ISLSGAAYKPAVAPPAARRSQIPEALRDEGFNLVYEFCPYFSFGHFAANAAILD
AFEGESRVHIVDLGMSSALQWPALLQGLASRPGGPPESIRITGVSCDRSDKLFL
AGEELSRLAESLELQFEFRAVTQAVESLQRGMLDVRDGEAMAINSAFQLHCV
VKESRRSLKSVLQSIHELSPKILTLVEQDACHNGPFFLGRFIEALHYYSAIFDA
VDAILPSDSEERLKIEQYHYAEEIKNIVACEGPDRVERHERADQWRRRMSRA
GFQPKPLKFLGEVKTWLGMYYPSEGYTLVEEKGCIVLGWKGKPIVAASTWR
C*

Setaria italica

> Sit Si000959m.g
MGMPEQPCRSTPNSFTTSFGSSQQMHHLPQHDAALCTEPGLGFPYYYGTDQQ
DAAFDGDEVDLGFRASKVTKVDYYSSPYQPSWPLARADVAAAAAESSRVRK
QRFRDVLESCKQKVEAMEAMESPVAFQEGEDGGVAGDGGGAAAGGGGGG
GGGGGGADGMRLVQLLVACAEAVACRDRAQAAALLRELQVGAPVHGTAF
QRVASCFVQGLADRLALAHPPALGPASMAFCIPPSCAGRDGARGEALALAYE

LCPYLRFAHFVANASILEAFEGESNVHVVDLGMTLGLDRGHQWRGLLDGLA
ARAGAKPKRVRVTGVGAPLDTMRAVGRELEAYAEGLGMRLEFRAVDRSLES
LHADDLGVAADEAVAISSVLELHCVVKESRGALNSVLQTVRKLSPRAFVLVE
QDAGHNGPFFLGRFMEALHYYAAVFDALDAALPRYDARRARVEQFHFGAEI
RNVVGCEGVARVERHERADQWRRRMSRAGFQSVPIRMAARAREWLEENAG
GGGYTVAEEKGCLVLGWKGKPLIAASCWKC*

Solanum lycopersicum

> Sly Solyc03g110950.1
MIHDMANVFLSLEPCNGDQIGYDPMENSLYLTHHKELSYSLNPYTSVLKRNA
PTNNMIISSLSNDSASFKRLRRTPSLGESFGSNTTFYSTESSSTGGSLPRIGSSNS
VNSLSLQPGIHFRDHVWALNQRYLAAEAFEEAAADIINQEEENGEGMKLVQL
LITCAEAVACRDKSRASVLLSELRASALVFGTSFQRVASCFMQGLSDRLALV
QPLGTVGYVATPAMNKTDIALEKKEEALRLLYEICPHIQFGHFVANCLILEAF
EGESFIHVVDLGMSLGLPHGHQWRRLVQSLVNRPGQPPRRLRITAVGQNIEK
LQIIGDELEDYARSLGINLEFSAVESNLENLKPKDIKVYDGEVLVVNSILQLHC
VVKESRGALNSVLQVVHELSPKILVLVEQDSSHNGPFFLGRFMEALHYYSAIF
DSLDVMLPKYDTRRAKIEQFYFAEEIKNIVSCEGPARVERHERVDQWRRRMS
RAGFQAAPIKMVSQAKQWLAKVNGHEGFTITEEKGCLVLGWKSKPIVAASC
WKC*

Solanum tuberosum

>Stu PGSC0003DMP400026755
MIHNMANVFLSLEPCNGDQIGYDPMENGLYLSHHKDLSYSLNPYTSVLKRNA
PTNSMIISTLSNDSGSFKRLRRTPSLGESFGSNSTFYSTDSSSGGSSNCSLPRIGS
TNSVNSLSLQPGIHFRDHVWALNQRYLAAEAIEEAAADIINQEEENGEGMKL
VQLLITCAEAVACRDKSRASVLLSELRASALVFGTSFQRVASCFMQGLADRL
ALVQPLGTVGYVATPAMNKTDIALEKKEEALRLLYETCPHIQFGHFVANCSIL
EAFEGESFIHVVDLGMSLGLPHGHQWRRLIQSLVNRPGQLPHRLRITAVGQNI
EKLQIIGDELEDYARSLGINLEFSAVESNLENLKPKDIKVYDGEILVVNSILQLH
CVVKESRGALNSVLQVVHELSPKILVLVEQDSSHNGPFFLGRFMEALHYYSAI
FDSLDVMLPKYDTRRAKIEQFYFAEEIKNIVSCEGPARVERHERVDQWRRRM
SRAGFQAAPIKMVSQAKQWLAKVNGHEGFTITEEKGCLVLGWKSKPIVAAS
CWKC*

Sorghum bicolor

>Sb Sb03g043030
MGTSEQPCTSNLFTASSYGTSQQIHHLLPQHDSVICTEPGMGFPYYYGTDQQD
AAFDGDEVELGFQASKATRVDYYSSPYQPSWPLARAAATESSRVRKQRFRD
VLESCKQKVEAMEAMESPVAFQEGEDGLAVGDGGGAGAAAGGGAGAGGG
NGGGADGMRLVQLLVACAEAVACRDRAQAALLRELQAGAPVHGTAFQRV
ASCFVQGLADRLALAHPPALGPASMAFCIPPSCTGRDGARGEALALAYELCP
YLRFAHFVANASILEAFEGESNVHVLDLGMTLGLDRAHQWRGLLDGLAARA
GAKPARVRVTAVGAPAETMRAVGRELEAYAEGLGLCLEFRAIDRSLESLHM
DDLGIAADEAVAISSILELHCVVKESRGALNSVLQTIRKLSPKAFVLVEQDAG
HNGPFFLGRFMEALHYYAAVFDALDAALPRYDARRARVEQFHFGAEIRNVV

GCEGAARVERHERADQWRRRMSRAGFQSVPIRMAARAREWLEENAGGGGY
TVAEEKGCLVLGWKGKPVIAASCWKS*

Theobroma cacao

> Tc Thecc1EG014574
MGHGFLANEFQKTDRIDEVIGLDLELSAMAFCYQPFMPIMGDNACGWSLPFS
GEIRDTKRLRRTISIPESIGSSGSLSSGGNSDSSLSRSGSTSSLNSFSRLHFRDHV
LTYNQRYLAAEAVEEAAAAMISSEESGGEEDETADGMRLVQLLIACAEAVA
CRDKSHASALLSELRANALVFGSSFQRVASCFVQGLADRLALVQPLGTVGLV
APVMNIMDISSDKKEEALRLVYEICPHIQFGHFVANSSILEAFEGESFVHVVDL
GMTLGLPHGHQWRHLIQSLANRAGKAPSRLRITAVGLSDHRFHIIGQELEAY
AKDLGMNLEFSVVKSNLENLRPEDIKVFDGEVLVVNSILQLHCVVKESRGAL
NSVLQMIHELSPKVLVLVEQDSSHNGPFFLGRFMEALHYYSAIFDSLDAMLPK
YDTRRAKMEQFYFAEEIKNIVSCEGPGRVERHERVDQWRRRMSRAGFQAAP
LRMMTQAKQWLGKNKVCEGYTVVEDKGCLVLGWKSKPIVAASCWKC*

Vitis vinifera

> Vv GSVIVG01007532001
MASDLLSEERSDEVSGLDTSLSAQAYYSRSYLPIFQNGSATNWFHYSDEARN
HKRLKRTQSIAESIGSNSSLYSGGKSYSNSSSSFINRSSSTNSLNSLPRLHFRDHI
WTYTQRYLAAEAVEEAAAAMISAAEGEVEEDGSGDGMRLVQLLIACAEAVA
CRDKTHASSLLSELRANALVFGSSFQRVASCFVQGLADRLSLVQPLGAVGFIA
PSINPLDTAWEKKEEALRLVYEICPHIKFGHFVANASILEAFEGENFAHVVDL
GMTLGLAHGQQWRQLIHSLANRAGRPPRRLRITGVGLCVDRFKIIGEELEAY
AQDLDINLDILQLHCVVKESRGALNSVLQKINELSPKVLVLVEQDSSHNGPFF
LGRFMEALHYYSAIFDSLEAMLPKYDTRRAKIEQFYFGEEIKNIVSCEGPARV
ERHERVDQWRRRMSRAGFQAAPIKMMAQAKQWLGKVKACEGYNIMEEKG
CLVLGWKSKPIVAASCWKC*

> Vv GSVIVG01010305001
MKLVHQLITCAKVVAFRDKSHASALLSELRANALVFGTSFQRVASCFVQGLS
DRLSLIQSLGAVGVGGCTVKTMDITPEKEEAFRLFFEICPQIQFGHLAANASIL
EAFEGESSVHVVDLGMNLGSPQGQQWRSLMHSLANRAGKPPSSLQITGVGT
AAECLKDIIDELEVYAESLGMNFQFSMLHCVVKESRGALNSVLQKIRELSPKA
VVLVEQDASHNGPFFLGRFMEALHYYSAIFDSLDAMLPKYDTRRAKMEQFY
FAEEIKNIISCEGSARVERHQRLDQWRRRMSRAGFQSSPMKMITEAKQWLEK
VKLCDGYTIVDEKGCLVLGWKSKPIIAASCWKCS*

Zea mays

> Zm GRMZM2G013016
MGTSEQPCTSTPNSFTTTSSSYVTSQQIHHLPQHDSVVCTEPGLGFPYYYVTD
RQQDAAFDGDEVELGFQASKATRVDYYSSPYQLSWPLARAAAAESSRVRKK
RFWDVLESCKQKVEAMEAMESPLVAFQEAEDGGAVVGDGGGGGGGRGSG
GGADGMRLVQLLVACAEAVACRDRAQAAALLRELQAGAPVHGTAFQRVAS
CFVQGLADRLALAHPPALGPASMAFCIPPSCAGRDGGARAEALALAYDLCPY
LRFAHFVANASILEAFEGETNVHVLDLGMTLGLDRAHQWRALLDGLAARAG
AAARPARVRVTAVGAPADAMRAVGRELLAYAEGLGMCLEFRAVDRSLESL
HIDDLGIAADEAVAINSVLELHCVVKESRGALNSVLQTIRKLSPKAFVLVEQD

AGHNGPFFLGRFMEALHYYAAVFDALDAALPRYDARRARVEQFHFGAEIRN
VVGCEGAARVERHERADQWRRRMSRAGFQSMPIRMAARAREWLEENAGG
GGYTVAEEKGCLVLGWKGKPVIAASCWKC*