

TRANSCRIPTOMICS OF POST-ANTHESIS COLOUR CHANGE IN *LOTUS*

by

Mannfred M.A. Boehm

B.Sc, Nova Scotia Agricultural College (Dalhousie University), 2014

A THESIS SUBMITTED IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS OF
FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

January 2017

© Mannfred M.A. Boehm, 2017

Abstract

Post-anthesis colour change (PACC) is widely thought to be an adaptation to signal floral viability to pollinators, and may be a developmental pre-pattern for bird pollination in *Lotus*. *Lotus filicaulis* and *Lotus sessilifolius* are insect-pollinated herbaceous legumes with flowers that open yellow, shift to orange and finally, red. This thesis examines the ecological and molecular basis for floral colour change in these *Lotus* species.

L. filicaulis was cultivated in a glasshouse from which pollinating insects (bees) were excluded, and the rate of colour change was recorded in both unpollinated and manually pollinated flowers. Unpollinated flowers from both the yellow stage and red stage were sampled for sequencing. The transcriptomes of *L. filicaulis* and *L. sessilifolius* of both colour stages were analyzed for differentially expressed genes and enriched ontologies.

Pollination significantly accelerates PACC in *L. filicaulis*, consistent with the hypothesis that PACC increases pollination efficiency by directing pollinators to unpollinated flowers. RNA-seq results show the synchronized upregulation of the entire cyanidin biosynthesis pathway in the red stage of PACC of both *Lotus* species – possibly at the expense of adjacent pathways competing for the same substrate. This thesis is the foundation for understanding the molecular evolution of PACC in *Lotus* and will be useful in testing the hypothesis that PACC is a preadaptation to bird pollination in *Lotus*.

Preface

This thesis is a draft manuscript that will be submitted for publication: Boehm, M.M.A., Ojeda D.I., and Cronk Q.C.B. 2016. Dissecting the ‘bacon-and-eggs’ phenotype: transcriptomics of floral colour change in *Lotus*. The idea came from discussions with Q.C.B Cronk and D.I. Ojeda. M.M.A. Boehm designed the analytical pipeline, performed all the analyses, conducted the hand pollination experiment and wrote the manuscript. D.I. Ojeda designed the transcriptomics experiment, collected plant material and processed tissue for RNA sequencing. Q.C.B. Cronk contributed to experimental and analytical design, contributed to the analysis of data, and provided feedback on this manuscript.

Table of Contents

Abstract.....	ii
Preface.....	iii
Table of Contents.....	iv
List of Tables.....	vii
List of Figures.....	viii
Acknowledgements.....	ix
1 Introduction.....	1
1.1 What is post-anthesis colour change (PACC)?	1
1.2 Historical record of PACC	1
1.3 Evolutionary ecology of PACC	2
1.4 Insect vision and colours of PACC	2
1.5 Molecular basis of PACC	3
1.6 The genus <i>Lotus</i>	4
1.7 This study	4
2 Materials and Methods.....	8
2.1 Plant material and cultivation	8
2.2 Hand pollination experiment	8
2.3 Transcriptome sequencing	9
2.4 Sequence analysis	9
2.5 TopHat Alignment	9
2.6 STAR Alignment	10
2.7 <i>De novo</i> assembly by Trinity	10
2.8 Nucleotide alignments	11
2.9 Gene family composition	11

3	Results.....	13
3.1	Pollination accelerates colour change in <i>L. filicaulis</i>	13
3.2	Illumina HiSeq Library	13
3.3	<i>De novo</i> assembly of non-model organism RNA reads outperforms alignment of reads to a reference genome of a closely related species.....	14
3.4	The cyanidin pathway is <i>upregulated in the red stage of PACC</i> in <i>L. filicaulis</i> and <i>L. sessilifolius</i>	14
3.5	The carotenoid pathway is upregulated in the yellow stage of PACC in <i>L. filicaulis</i> and <i>L. sessilifolius</i>	14
3.6	Differences between <i>L. filicaulis</i> and <i>L. sessilifolius</i>	15
4	Discussion.....	27
4.1	The significance of pollination as a trigger for PACC	27
4.2	<i>De novo</i> assembly is preferred for RNA-seq data without a reference genome	27
4.3	PACC is associated with the wholesale upregulation of the cyanidin pathway	28
4.4	PACC is associated with highly specific upregulation of members of gene families	28
4.5	Upregulation of flavonoid biosynthesis may be antagonistic to other phenylpropanoid pathways	30
4.6	Carotenoid biosynthesis is downregulated during PACC	30
4.7	Is MAX1 a master regulator of PACC?	30
4.8	The role of FLS in PACC in <i>Lotus sessilifolius</i>	31
4.9	The role of red in PACC	32
5	Conclusion.....	34
	Bibliography.....	35

Appendix 1 Sequences for Trinity contigs listed in Table 5.....	44
Appendix 2 Sequences for Trinity contigs listed in Table 6.....	50
Appendix 3 Supplemental Tables S1-42.....	54
Appendix 4 Supplementary Figures S1 – S10.....	162
Appendix 5 Supplementary Figures S11 – S17.....	176
Appendix 6 Supplementary Figures S18 – S19.....	187
Appendix 7 Supplementary Figures S20 – S28.....	189
Appendix 8 Supplementary Figures S29 – S30.....	201
Appendix 9 Supplementary Figures S31 – S34.....	203
Appendix 10 Supplementary Figure S35.....	208
Appendix 11 Supplementary Figures S36 – S40.....	209
Appendix 12 Supplementary Figures S41 – S47.....	215
Appendix 13 Supplementary Figures S48 – S54.....	224
Appendix 14 Supplementary Figures S55.....	232
Appendix 15 Supplementary Figures S56 – S57.....	233
Appendix 16 Supplementary Figures S58 – S60.....	235
Appendix 17 Supplementary Figure 61.....	238

List of Tables

Table 1	Next-generation library construction for <i>L. filicaulis</i>	12
Table 2	Next-generation library construction for <i>L. sessilifolius</i>	12
Table 3	Comparison of protocols used to detect differential expression of RNA in <i>L. filicaulis</i>	21
Table 4	Comparison of protocols used to detect differential expression of RNA in <i>L. sessilifolius</i>	22
Table 5	Selected genes for enzymes involved in flavonoid biosynthesis in <i>L. filicaulis</i>	23
Table 6	Selected genes for enzymes involved in flavonoid biosynthesis in <i>L. sessilifolius</i>	25

List of Figures

Figure 1	<i>Lotus filicaulis</i> , exhibiting post-anthesis colour change	6
Figure 2	Simplified phylogeny of the <i>Lotus</i> spp. used in this study	7
Figure 3	Timing of stages in the hand pollinated treatment	16
Figure 4	Graphical representation of the differential expression of genes (Trinity contigs) involved in flavonoid and lignin biosynthesis throughout PACC in <i>Lotus filicaulis</i>	17
Figure 5	Graphical representation of the differential expression of genes (Trinity contigs) involved in flavonoid and lignin biosynthesis throughout PACC in <i>Lotus sessilifolius</i>	18
Figure 6	Graphical representation of the differential expression of genes (Trinity contigs) involved in carotenoid biosynthesis throughout PACC in <i>Lotus filicaulis</i>	19
Figure 7	Graphical representation of the differential expression of genes (Trinity contigs) involved in carotenoid biosynthesis throughout PACC in <i>Lotus sessilifolius</i>	20

Acknowledgements

Scripting the RNA-seq pipeline would not have been possible without the patient assistance of Sebastian Bailey, Alistair Blachford, Dan Bock, Melissa Guzman, Andy LeBlanc, Geneviève Leduc-Robert, Andrew MacDonald, Remi MattheyDoret, Diana Rennison, Saemundur Sveinsson, David Tack, and Silu Wang.

Growing healthy *Lotus filicaulis* and *Lotus sessilifolius* was greatly helped by Pierre Fobert, Melina Biron, and Dario Ojeda.

Insightful and often crucial feedback was provided by Brian Ellis, FLORUM, Armando Gerald, Anastasia Kuzmin, the Angert Lab, the Otto Lab, the Samuels Lab, Noriko Okamoto, the Schluter Lab, and Santokh Singh.

I would like to especially acknowledge Prof. Quentin Cronk and Kristen Hudson for continuing to encourage my interest in plants.

1 Introduction

1.1 What is post anthesis colour change (PACC)?

In angiosperms, flowers are typically left on display until pollination or senescence (i.e. until the pollen becomes inviable or the stigma non-receptive), at which time the flower will generally wilt and cease reward production. Often the corolla abscises at this point, thus effectively excluding any further visits from pollinators (Faegri & Vanderpijl, 1979). However, instead of abscission, some species retain the corolla past the cessation of reward production and modify petal colour, a trait I herein refer to as post-anthesis colour change (PACC) (Mohan Ram & Mathur, 1984; Farzad et al., 2003). I use “post-anthesis colour change” instead of the broader “floral colour change” as many flowers change colour between pre-anthesis (bud) and anthesis (flower opening). This early maturation colour change is unrelated to PACC. Several species of *Lotus* L. (Fabaceae) exhibit yellow flowers at anthesis followed by a transition to red (Ojeda et al., 2013) thus these species have sometimes acquired the English vernacular name ‘bacon and eggs’ (Grigson, 1975) (Figure 1). A well-studied example is the North American lotoid legume *Acmispon glaber* (Vogel) Brouillet (*Lotus scoparius* (Nutt.) Ottley) (Jones & Cruzan, 1999).

1.2 Historical record of PACC

The widespread (Weiss, 1995) and striking phenomenon of PACC has long interested botanists. One of the first scientific observations of PACC dates to a letter written from Fritz Müller to Charles Darwin, discussing the pollination ecology of *Lantana camara* L. (Müller, 1877; see also Weiss, 1991; Jones & Cruzan 1999). Müller notes that *L. camara* flowers open as yellow and over a three-day period shift to orange and finally purple. Tellingly, he also writes that butterflies tend to visit yellow flowers most often, occasionally visiting the orange flowers, and never the purple flowers. More recently the phenomenon has been documented in a variety of studies and reviews (Mohan Ram & Mathur, 1984; Weiss, 1991; Ojeda et al., 2013; Brito et al. 2015).

1.3 Evolutionary ecology of PACC

As anticipated by Müller, PACC is likely an adaptation to redirect pollinators at close range while maintaining long-distance appeal of the plant floral display (Weiss, 1991; Jones & Cruzan, 1999), hence the retention of the corolla instead of wilting or abscission. Indeed, in the North American *Acmispon glaber* (formerly known as *Lotus scoparius*) PACC is accelerated by pollination and corresponds temporally with the termination of reward production (Jones & Cruzan 1999). By differentiating between viable and unviable flowers, it is thought that plants are able to signal to pollinators which flowers possess a food reward (Weiss, 1991). This is of potential benefit for both organisms as the pollinator forages with greater efficiency, while the plant may receive a higher rate of successful pollination events. Pollinators could, of course, be directed to young flowers by the wilting or abscission of old flowers, but it is hypothesized that PACC provides a selective advantage by aiding long-distance attractiveness of the overall floral display.

1.4 Insect vision and colours of PACC

The flower colours involved in effective signalling are not random; PACC tends to be a shift from shorter reflective wavelengths (e.g. yellow), to longer ones (e.g. red). Like many floral traits, it is plausible that PACC has evolved through interaction with pollinators, and thus the characteristics of insect vision may be implicated in the colour palette of PACC. Generally, bees (Hymenoptera: Apidae) have three kinds of light receptors with peak sensitivities at 340nm (ultraviolet, S-receptor), 430nm (blue, M-receptor), and 540nm (green, L-receptor) (Chittka, 1996). However, the physiology of hymenopteran light receptors (Chittka, 1996) and colour-reward experiments (Weiss, 1991; Chittka & Waser, 1997; MartínezHarms et al., 2010) suggest that bees are not incapable of detecting red light (620-750nm), and electroretinogram recordings of *Bombus dahlbomii* (Apidae) have demonstrated that red light can induce an L-receptor response (Martínez-Harms et al., 2010). Indeed, in the wild *B. dahlbomii* is known to visit a variety of plant species with flowers that look red to humans, and they can be trained to visit artificial red targets *ex situ* (Martínez-Harms et al., 2010). Results of colour-reward experiments suggest that bees are able to detect red flowers by achromatic contrast (Chittka & Waser, 1997; Martínez-Harms et al., 2010). That is, red flowers likely appear as a dark, hue-less patches

against a background of green plant foliage, providing sufficient contrast to locate such flowers (Martínez-Harms et al., 2010).

Given that a pollinator spends a finite amount of time at any one plant, PACC may be an adaptation to attract insect pollinators at long range by deceptively by retaining flowers, regardless of viability, while honestly directing pollinators at short range to the flowers most likely to benefit from pollen transfer (discussed in Brito et al., 2015). In numerous plant taxa the visual physiology of pollinators has mediated the evolution of flower colour (Rausher, 2008); in the case of PACC, it is conceivable that natural selection would favour a post-anthesis colour that is relatively less attractive than at anthesis, but not so unattractive as to be indistinguishable from its surroundings.

1.5 Molecular basis of PACC

The timing of PACC is potentially tied to a host of environmental and biochemical triggers that influence pigment pathways in conjunction with the cessation of food rewards. In *Lantana camara*, *Gossypium hirsutum* L., and *Viola cornuta* L. PACC has been found to a result from changes in flavonoid biosynthetic processes, specifically those branches that produce anthocyanins and their methylated derivatives (Mohan Ram & Mathur, 1984; Farzad et al., 2003; Tan et al., 2013). The anthocyanin biosynthetic pathway (ABP) is a highly conserved and versatile pathway known to produce the orange, red and purple colouration in floral tissues - in addition to a range of other functions not related to pigmentation (Kong, 2003). The regulation of PACC at the genetic level remains poorly known, however, presuming the anthocyanin pathway is involved, it is possible that PACC is regulated by MYB transcription factors (e.g. PAP1 and PAP2, Feller et al. 2011) as well as WD40s and basic helix-loop-helix proteins (Ramsay & Glover, 2005). These three proteins are known to produce a MYB-bHLH-WD40 transcription complex capable of regulating a range of target genes in numerous plant species (Ramsay & Glover, 2005), including those involved in anthocyanin biosynthesis (Spelt et al., 2002). Additionally, many flavonoid biosynthetic pathway genes are represented as gene families with multiple duplicated copies (Ober, 2005) and whether all or some of these copies are involved in PACC remains to be determined.

1.6 The genus *Lotus*

Lotus is a genus with a diverse range of floral traits to attract pollinators (Ojeda et al., 2013), including varying sugar compositions, corolla shapes, petal micromorphologies, and petal pigmentations (Ojeda et al., 2012; Ojeda et al., 2013).

It is a promising genus for studying the evolution of pollination syndromes, and PACC is particularly interesting because it has been suggested that it is a preadaptation (developmental ‘pre-pattern’), for the evolution of bird pollination in *Lotus* section *Rhyncholotus* (Ojeda et al., 2013). *Lotus* includes the well-studied *L. corniculatus* L., which is of considerable economic importance as a forage crop (Duke, 1981) and exhibits PACC. However, *L. corniculatus* is a polyploid and therefore a complex organism for studying gene expression.

The anthocyanin pathway in *Lotus japonicus* (Regel) K. Larsen (*L. corniculatus* L. var. *japonicus* Regel) has been well studied, but it does not exhibit PACC. Chromatographic analysis of *L. japonicus* aerial tissue has revealed that cyanidin and peonidin make up the flavonoid profile of this species (Suzuki et al., 2008); both of these pigments are known to reflect red and purple light and it is possible these same branches of the anthocyanin pathway are involved in the *Lotus* spp. that exhibit PACC. For this reason I have chosen to leverage the genomic resources available for *L. japonicus* by studying PACC in the closely related *L. filicaulis* Durieu, a diploid perennial herbaceous legume (Ferreira & Pedrosa-Harand, 2014) notable for its post-anthesis transition from yellow to red flowers (Figure 1). For comparison I also investigate another species, *Lotus sessilifolius* DC, which has been the subject of previous studies (Ojeda et al., 2013) and may play a significant role in the evolution of pollination syndromes in *Lotus* (Ojeda et al. 2012). Additionally, an analysis of the pigment composition of *L. sessilifolius* floral tissue suggests that *FLAVONOL SYNTHASE* (*FLS*) may play an important role in PACC of this species (Ojeda et al. 2013).

1.7 This Study

In this study I first investigate whether PACC is triggered by pollination in *Lotus filicaulis*. I then compare RNA expression at anthesis (yellow) and during PACC (red) in *Lotus filicaulis* and *Lotus sessilifolius* to describe the pigment pathways involved, and propose several candidate genes and pathways that I believe play a key role in post-anthesis colour change in these *Lotus*

species. This is the first transcriptome-wide analysis of PACC to date. PACC appears to have evolved independently in *L. filicaulis* and *L. sessilifolius* (Ojeda et al. 2013; Figure 2) as they belong to different clades of *Lotus*. *L. sessilifolius* is in the Canary Island clade of *Lotus* and apparently a tetraploid (Ojeda and Santos-Guerra, 2011, Ojeda et al. 2013). However, *L. sessilifolius* (insect-pollinated) is the sister taxon to the bird pollinated ‘rhyncholotus group’ in *Lotus* (Ojeda et al. 2013), and further studies on this species can provide additional information about the evolution of bird pollination in this group. I discuss the results in the context of observations regarding PACC in other plant species, the role of pollinators in driving the evolution of PACC, the role of gene duplication, and known molecular pathways of pigment biosynthesis.



Figure 1 *Lotus filicaulis*, exhibiting post-anthesis colour change from yellow to red, a trait commonly referred to as the ‘bacon and eggs’ phenotype. Photo taken at the Horticulture Glasshouse, University of British Columbia.

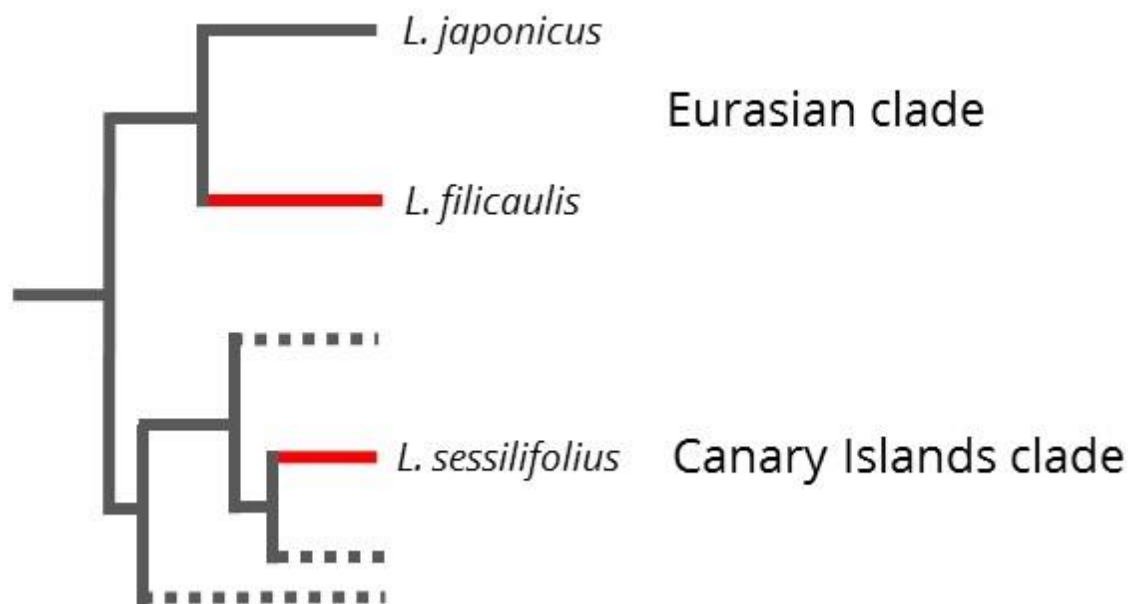


Figure 2 Simplified phylogeny of the *Lotus* spp. used in this study (modified from Ojeda *et al.* 2013, reproduced in Supplementary Figure S54). Red branches show clades where PACC has evolved, and dotted branches indicate the presence of adjacent clades.

2 Materials and Methods

2.1 Plant material and cultivation

The North African species, *Lotus filicaulis* Durieu (voucher: Ojeda 71, herb. UBC) was used for experiments on floral manipulation; *L. filicaulis* is sometimes regarded as synonymous with *L. tenuis* Willd. A second species, *Lotus sessilifolius* DC. (voucher: Ojeda 225, herb. UBC) from the Canary Islands, Spain, was added as a comparison for RNAseq experiments. Both species exhibit strong PACC with a yellow to red colour change. The plants were propagated from seeds in the horticulture glasshouse, University of British Columbia, Vancouver, Canada. All plants were grown in pots of 10-20 cm in diameter at 20-25 °C and were more than 6 weeks old when flowers were collected for analyses.

2.2 Hand pollination experiment

In order to standardize the stages of PACC, five genetic individuals of *L. filicaulis* were tagged for both control and treatment (hand pollination), monitored daily for colour change, and photographed to capture distinct colour stages. 3 flowers were measured per plant for both the control and treatment. Hand pollination was conducted one day after anthesis (flower opening) to ensure the keel petals had separated. Pollen was transferred between flowers using an implement of a Victorinox SwissTool™ (unpointed blade).

The rate of progression through these stages was modelled using the lme4 package (Bates et al. 2015) in R version 3.2.2 (R Development Core Team, 2008). Data were visualized using the ggplot2 package (Wickham, 2009) in R. The R scripts used in this analysis is available at github.com/mannfred/Lotus.

A similar experiment was conducted on 21 individuals of *L. filicaulis*, but rather than tracking each stage of colour change, only the number of days for PACC to initiate was recorded. The results of this experiment can be found in Supplementary Table S42.

2.3 Transcriptome sequencing

The entire corollas (dorsal, lateral and ventral petals) of *L. filicaulis* and *L. sessilifolius* were separated from the rest of the flower at the same developmental stage (Stage 13 of anthesis, Ojeda et al., 2012), but with two different colour stages: 1) flowers at anthesis with a yellow color, and 2) flowers initiating flower color change to red. At least six unpollinated flowers of each type, yellow and red, were collected from one individual. Petals of each colour were removed from the flower and pooled to represent the phenotype across the entire individual. Flower petals were immediately placed on liquid nitrogen after collection and stored at 80°C until RNA extraction. RNA was extracted using the Pure Link™ Plant RNA Reagent from Invitrogen following the manufacturer's protocol. RNA quantity and quality was determined using a QubitR 2.0 Invitrogen (Life Technologies) and with an Agilent 2100 Bioanalyzer, respectively. Samples with a RIN value of 7 or above were used for library construction. Paired end RNA libraries (Tables 2.1 and 2.2) were prepared for each individual sample with the Illumina kit according to the manufacturer's protocol. All samples were sequenced (100 base pair reads) in one lane of an Illumina Hi-Seq 2000 at the NextGen Sequencing Facility at the Biodiversity Research Centre, University of British Columbia (UBC).

2.4 Sequence analysis

Aligning RNA-seq data from *L. filicaulis* and *L. sessilifolius* has not yet been reported. Therefore, I compared two methods to discover differentially expressed genes relevant to PACC. In the first method, RNA-seq reads were aligned to the reference genome of a closely related model organism, *Lotus japonicus*. In the second method, RNA-seq reads were assembled into contigs (putative genes) and aligned to a reference transcriptome that was assembled *de novo* from the reads themselves. All scripts used in the analyses are available at github.com/mannfred/Lotus.

2.5 TopHat Alignment

RNA-seq reads from *L. filicaulis* were analysed by TopHat. FASTQC Galaxy v0.63 (Andrews, 2010) was executed on a public Galaxy server (Goecks et al., 2010) to check the quality of reads and to identify sequencing artifacts. Sequencing reads were mapped to the *Lotus japonicus* genome release 2.5 (Sato et al., 2008) using TopHat2 Galaxy v0.9 (Kim et al., 2013). Sequence

Alignment Map (.sam) files were processed by HTSeq (Anders et al., 2015) to count the number of RNA-seq reads that mapped to each TopHat feature (gene), and differential expression (DE) was calculated using edgeR (Robinson et al., 2010). Raw counts per gene were used instead of FPKM transformed counts because edgeR incorporates its own dispersion-normalization procedure in its DE calculations (Robinson et al., 2010). EdgeR results were processed using a set of custom R scripts to filter significant results (Benjamini-Hochberg FDR <0.05 (Benjamini & Hochberg, 1995)) and differences in fold change.

DE genes were analyzed for ontology term enrichment (Yekutieli FDR<0.05 (Yekutieli & Benjamini, 1999)) using the Singular Enrichment Analysis tool with *Lotus japonicus* 2.5 genome reference in AgriGO (Du et al., 2010). Because the functional annotation of the *L. japonicus* genome is ongoing, genes associated with enriched GO categories were annotated for putative function by BLASTing the *Glycine max* (L.) Merr. CDS library (Williams 82 Assembly 2 version 1, Schmutz et al., 2010) sourced from Phytozome v10.3 (Goodstein et al., 2012).

2.6 STAR Alignment

RNA-seq reads from both *L. filicaulis* and *L. sessilifolius* were mapped to the *Lotus japonicus* genome Kazusa version 2.5 (Sato et al., 2008) using STAR v2.4 (Dobin et al., 2013). Successfully mapped transcripts were assembled and expression was estimated using RSEM v1.2.19 (Li & Dewey, 2011). Differential expression of the yellow and red flower transcriptomes was calculated using the same methods as above.

2.7 De novo Assembly by Trinity

Sequenced Illumina reads from *L. filicaulis* and *L. sessilifolius* were assembled *de novo* into contigs in Trinity v2.0.6 (Grabherr et al., 2011). Using assembled contigs as a reference transcriptome, contigs were measured for expression using Trinity's built-in RSEM tool (Li & Dewey, 2011) and DE was calculated as above. DE contigs were then identified by BLASTn (Altschul et al., 1997) as their top scoring *Lotus japonicus* homolog (expectation value of 1e-4) using the *L. japonicus* v2.5 CDS library sourced from Kazusa (Sato et al., 2008). Ontology term enrichment and annotation of associated genes was analysed as above.

2.8 Nucleotide Alignments

Trinity contigs BLASTing to *L. japonicus* genes of interest were investigated further by BLASTing the Legume IP database (Li et al., 2011) for *Glycine max* (v2.0) (Schmutz et al., 2010), *Medicago truncatula* Gaertn. (Mt4.0v1) (Young et al., 2011), and an outgroup *Arabidopsis thaliana* (L.) Heynh. (TAIR10) (Lamesch et al., 2011) to further identify homologous genes. Trinity contigs were trimmed before and after putative start and stop codons, respectively, which were identified by pairwise sequence alignment of their *L. japonicus* homolog using EMBOSS Needle (Rice et al., 2000). Coding sequences of putative homologs were compiled along with the relevant trimmed Trinity contig into a .fasta file and translated to amino acid sequences using TranslatorX (Abascal et al., 2010). Multiple sequence alignment was performed using the T-Coffee (Notredame et al., 2000) option in TranslatorX. The phylogenetic relationship of a given Trinity contig to its relatives was visualized using RAxML v8.2 (Stamatakis, 2014) in Mesquite (Maddison & Maddison, 2015).

2.9 Gene family composition

Trinity (*de novo* assembly) sequences were BLASTed ($1e-4$) against the *L. japonicus* CDS library at LegumeIP to detect gene duplicates – specifically those not expressed at the time of sampling (this approach assumes no new gene duplications have occurred between *L. japonicus* and our two study species). BLAST results sharing the same annotation (i.e. gene name) as the top alignment were considered to be a part of the same gene family. Certain gene families (e.g. *DFR*, *CHS*) have been especially well-studied in *L. japonicus*, but CDS libraries are not necessarily updated in real time in accordance to ongoing *L. japonicus* research. In these cases, published literature on specific gene families took precedence over LegumeIP results.

Table 1 Next-generation library construction for *L. filicaulis*.

Library no.	Library ID (Barcode)	Yellow / Red	Plant no.	Flower no. (pooled)
1	ATGAGC	Red	1	≥ 6
2	ACTGAT	Yellow	1	≥ 6

Table 2 Next-generation library construction for *L. sessilifolius*.

Library no.	Library ID (Barcode)	Yellow / Red	Plant no.	Flower no. (pooled)
1	CAAAAG	Red	1	≥ 6
2	ATTCCT	Yellow	1	≥ 6

3 Results

3.1 Pollination accelerates colour change in *L. filicaulis*

Four distinct colour stages of PACC were observed; 1) Entirely yellow, 2) blush of orange, typically at the keel, 3) systemic orange colouration across the corolla, and 4) red, onset typically at the keel (Figure 2.1). The general linear model to test the effects of time (days) and treatment on colour was:

$$\text{COLOUR} = \text{CONSTANT} + \text{DAYS} + \text{TREATMENT} + \text{DAYS} * \text{TREATMENT}$$

Among plant standard deviation was effectively zero ($s = 2.664 \times 10^{-5}$) and was omitted as a random effect.

The number of days from anthesis was a significant predictor of colour ($df = 1, 39$; $p = 1.69 \times 10^{-9}$), as well as the effect of treatment through time ($df = 1, 39$; $p = 5.06 \times 10^{-4}$). No significant effect of treatment alone was detected ($df = 1, 39$; $p = 0.851$).

The rate of progression through the four stages of PACC in the hand pollinated treatment is over twice that of the control (see Figure 2.1 caption).

3.2 Illumina HiSeq library

A total of 12.5 million and 11.9 million .fastq reads were sequenced from the yellow-stage and red-stage RNA-seq libraries of *L. filicaulis*, respectively. GC content was 46% for the forward yellow-stage and red-stage reads and 45% for the reverse reads.

Similarly, 6.7 million and 12.0 million .fastq reads were sequenced from the yellow-stage and red-stage RNA-seq libraries of *L. sessilifolius*, respectively. GC content was 45% for the forward yellow-stage and red-stage reads and 44% for the reverse reads.

3.3 *De novo* assembly of non-model organism RNA reads outperforms alignment of reads to a reference genome of a closely related species

Two methods were used to discover differential expression in PACC, 1) direct mapping of reads to the *L. japonicus* genome using two different splice junction mappers, and 2) *de novo* assembly

of reads to contigs followed by BLASTing these contigs to a *L. japonicus* CDS library. The evaluation of the two approaches is summarised in Tables 3.1 and 3.2. Direct alignment by STAR and TopHat achieved low mapping rates due to unsuccessful heterologous mapping from our study species to *L. japonicus*. However, the *de novo* assembly protocol achieved a high rate of gene identification, and produced independent contigs that could be verified for identity and biological relevance by amino acid alignment and phylogenetic methods. Trinity contigs of interest were found to align with homologous genes of related legumes, and their known phylogenetic relationship was successfully predicted by these alignments (Supplementary Figures S1-S53). A comparison of the most differentially expressed contigs of both stages and of both *Lotus* species can be found in Supplementary Tables S38-S41.

3.4 The cyanidin pathway is upregulated in the red stage of PACC in *L. filicaulis* and *L. sessilifolius*

Regardless of the bioinformatics protocol, upregulation of numerous genes in the anthocyanin biosynthetic pathway (ABP) was detected in the red stage of PACC in both *Lotus* species examined (Tables 3.3 and 3.4, Figures 3.2 and 3.3). Except for anthocyanin synthase and flavonoid 3'-hydroxylase (both apparently single copy), gene duplicates or multiple gene families were identified in the *Lotus japonicus* genome for each of the genes of the ABP in both species (Figures 3.2 and 3.3). As a comparison the same analysis was carried out for the lignin biosynthetic pathway (LBP), which is not expected to be associated with PACC. There is no consistent upregulation of LBP genes during PACC. Generally, LBP genes are either not expressed, downregulated during PACC, or not differentially expressed (Figures 3.2 and 3.3).

3.5 The carotenoid pathway is upregulated in the yellow stage of PACC in *L. filicaulis* and *L. sessilifolius*

Genes annotated to the carotenoid biosynthetic process (GO:0016117) were significantly upregulated at the yellow stage (i.e. downregulated at the red stage of PACC) in both *Lotus* species (Supplementary Tables S35-S36). The core carotenoid pathway is represented, including the two major branches producing zeaxanthin and lutein (Figures 3.4 and 3.5).

3.6 Differences between *L. filicaulis* and *L. sessilifolius*

Although the cyanidin pathway is upregulated during PACC in both *Lotus* species, there are differences in the composition of genes expressed, specifically those genes that are a part of a gene family. For example, *CHALCONE SYNTHASE* (*CHS*) is active in both species but the gene family members expressed in *L. filicaulis* and *L. sessilifolius* are disjoint sets. Similar differences are observed in the expression of *DFR*. However, it is notable that of the four *F3H* duplicates known for *L. japonicus*, only one is implicated in PACC for the two *Lotus* species studied. The expression of transcription factors and regulators contribute to numerous other differences between PACC in *L. filicaulis* and *L. sessilifolius* (Supplementary Tables S11-S18, S27-S34). Finally, I bring attention to the presence of *FLAVONOL SYNTHASE* (*FLS*) in *L. sessilifolius* (Table 6, Figure 5). I detect an upregulation of *FLS* during the yellow stage of PACC in this species, and an absence of activity of this gene in *L. filicaulis*. The ecological and evolutionary implications of this are discussed below.

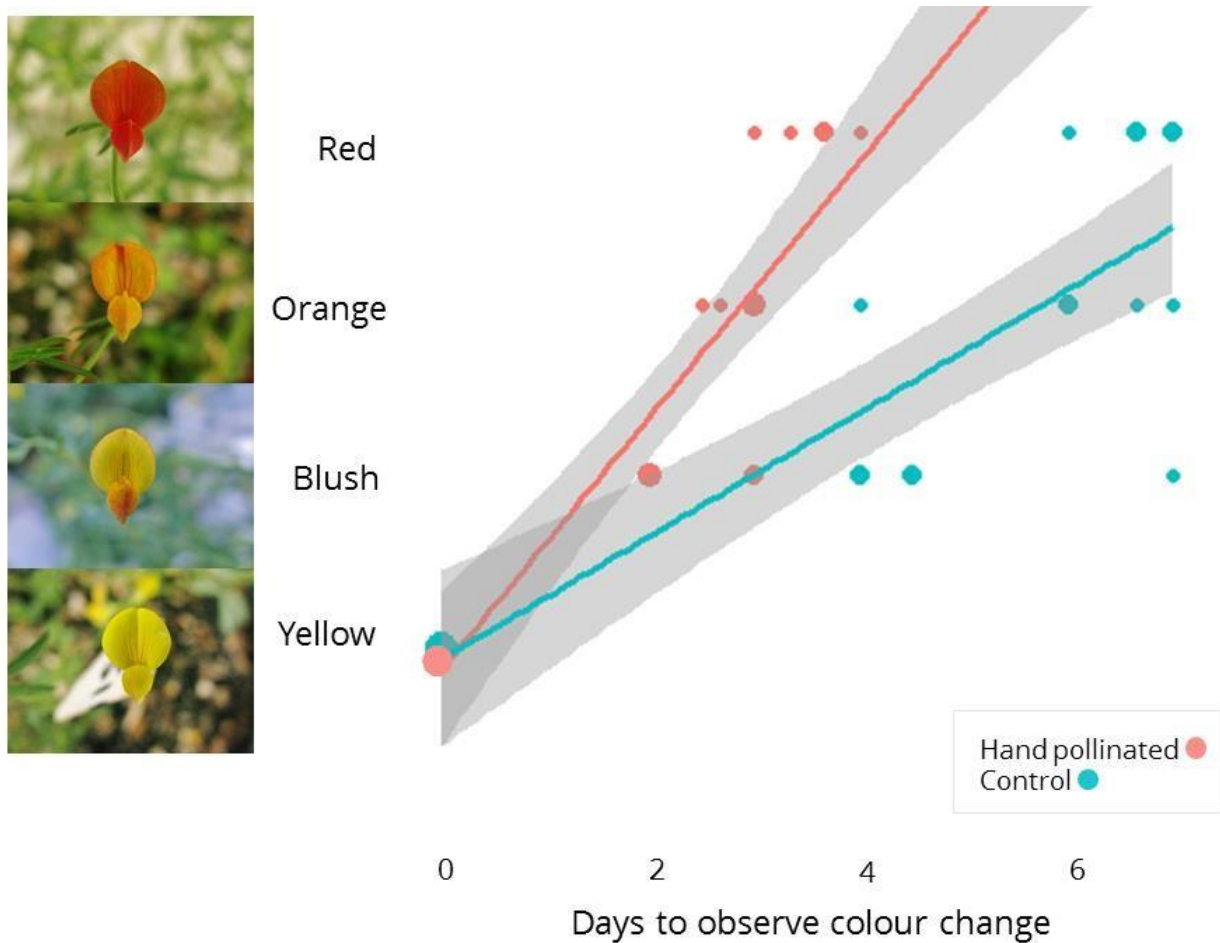


Figure 3 Left panel: Standardization of flower colour using photographs of four distinct stages: anthesis ('Yellow', stage 0), onset of keel colouration ('Blush', stage 1), systemic orange across the corolla ('Orange', stage 2), and full colouration ('Red', stage 3). Graph: timing of stages in the hand pollinated treatment (orange line, colour stage = $0.739(\text{Days}) - 0.121$) and control (unpollinated) (turquoise line, colour stage = $0.358(\text{Days}) - 0.059$). Intercepts were not fixed (anthesis) to maintain zero-mean residuals. Each point represents the mean number of days to one of four stages (mean of three replicates per treatment per individual plant). The size of points is proportional to the amount of over-plotted data.

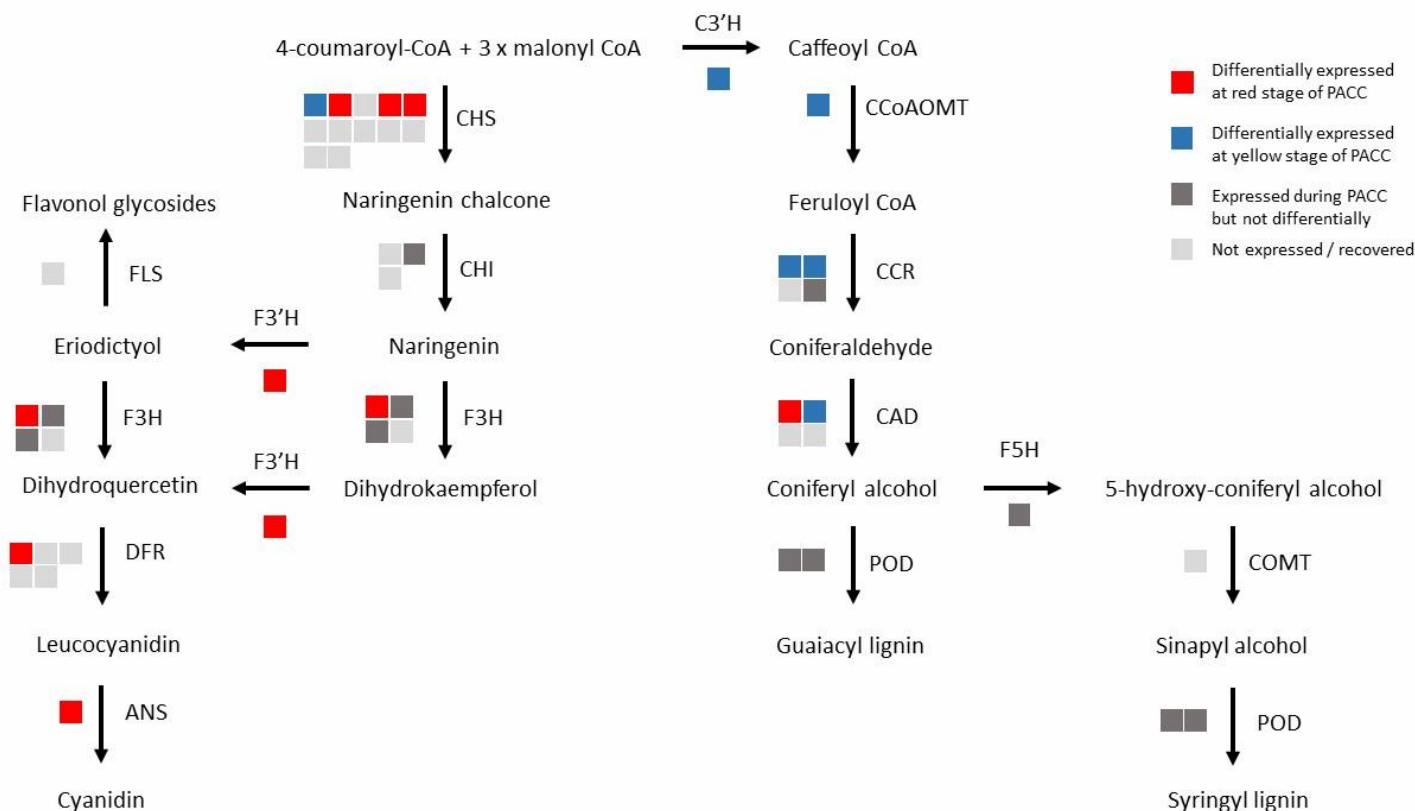


Figure 4 Graphical representation of the differential expression of genes (Trinity contigs) involved in flavonoid and lignin biosynthesis throughout PACC in *Lotus filicaulis*. The boxes represent genes as annotated in the *Lotus japonicus* genome; light grey boxes represent genes that were not found in the *L. filicaulis* transcriptome (gene may be absent or may have been undetected). Expression of putative orthologs in *L. filicaulis* are indicated by colour (red: upregulated at red stage; blue: upregulated at yellow stage; dark grey: not differentially expressed). 4-coumaroyl-CoA and malonyl CoA are derivatives of phenylalanine and acetatepyruvate, respectively. In the flavonoid pathway CHS = chalcone synthase, CHI = chalcone isomerase, F3'H = flavonoid 3'-hydroxylase, F3H = flavanone 3-hydroxylase, DFR is dihydroflavonol 4-reductase, and ANS is anthocyanidin synthase. In the lignin pathway C3'H = coumaroyl shikimate 3'-hydroxylase, CCoAOMT = caffeoyl-CoA O-methyltransferase, CCR = cinnamoyl-CoA reductase, CAD = cinnamyl alcohol dehydrogenase, POD = peroxidase, F5H = ferulate 5-hydroxylase, and COMT = caffeic acid O-methyltransferase.

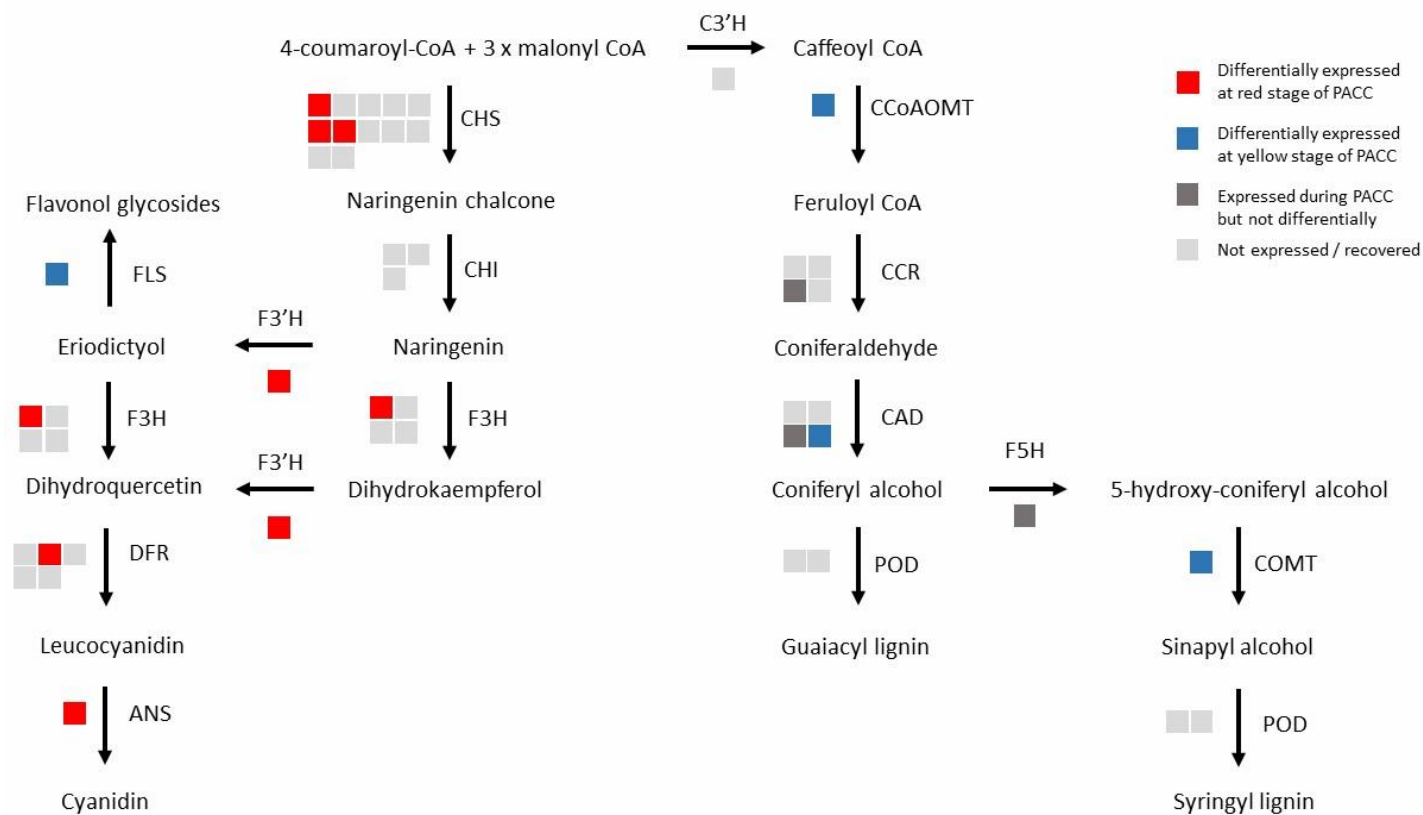


Figure 5 Graphical representation of the differential expression of genes (Trinity contigs) involved in flavonoid and lignin biosynthesis throughout PACC in *Lotus sessilifolius*. Conventions and acronyms are as defined in the legend to Figure 4.

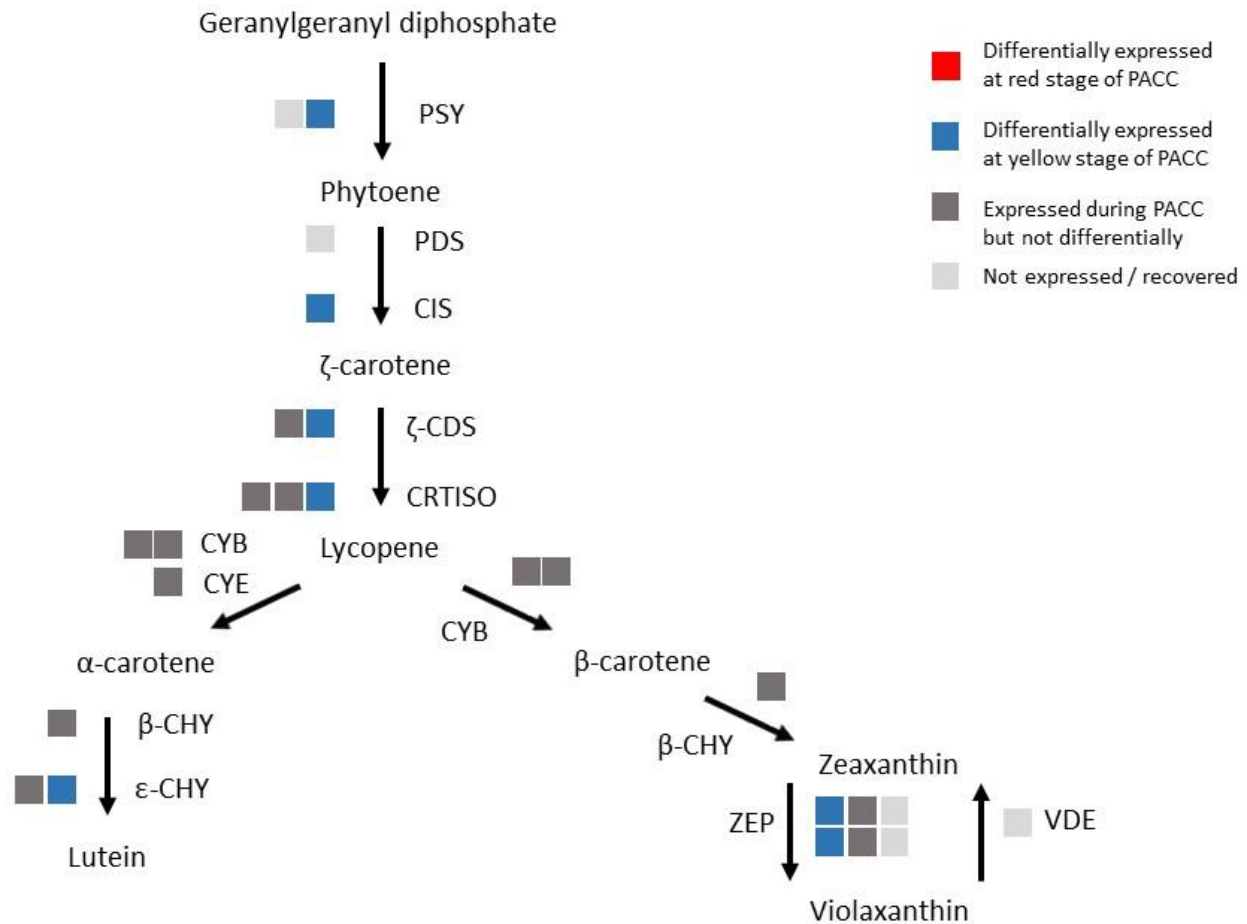


Figure 6 Graphical representation of the differential expression of genes (Trinity contigs) involved in carotenoid biosynthesis throughout PACC in *Lotus filicaulis*. The boxes represent genes as annotated in the *Lotus japonicus* genome; light grey boxes represent genes that were not found in the *L. filicaulis* transcriptome (gene may be absent or may have been undetected). Expression of putative orthologs in *L. filicaulis* are indicated by colour (red: upregulated at red stage; blue: upregulated at yellow stage; dark grey: not differentially expressed). Geranylgeranyl diphosphate is a derivative of pyruvate and glyceraldehyde 3phosphate. In the carotenoid pathway PSY = phytoene synthase, PDS = phytoene desaturase, CIS = carotene isomerase, ζ -CDS = ζ -carotene desaturase, CRTISO = prolycopene isomerase, CYB = β -lycopene cyclase, CYE = ϵ -lycopene cyclase, β -CHY = β -carotene hydroxylase, ϵ -CHY = ϵ -carotene hydroxylase, ZEP = zeaxanthin epoxidase, and VDE = violaxanthin de-epoxidase.

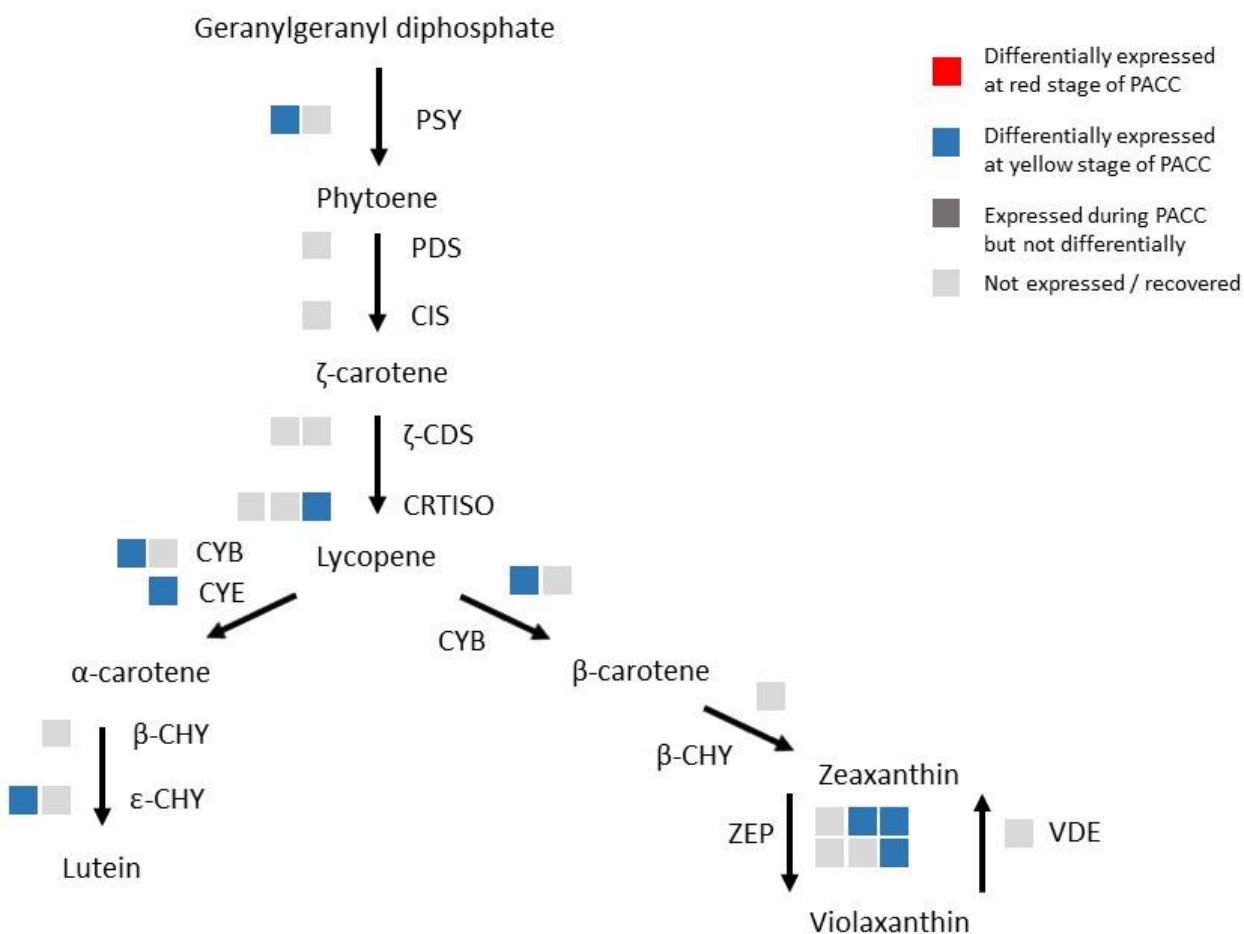


Figure 7 Graphical representation of the differential expression of genes (Trinity contigs) involved in carotenoid biosynthesis throughout PACC in *Lotus sessilifolius*. Conventions and acronyms are as defined in the legend to Figure 6.

Table 3 Comparison of protocols used to detect differential expression of RNA in *L. filicaulis*.

Analysis type (mapping reference)	Reads mapped yellow stage	Reads mapped red stage	DE genes yellow stage ↑	DE genes Red stage ↑	Enriched GO terms yellow stage	Enriched GO terms red stage
TopHat (<i>L. japonicus</i> genome)	8,538,623	7,243,178	1419	962	107	84
STAR (<i>L. japonicus</i> genome)	4,334,187	3,566,554	1398	958	128	97
Trinity (<i>L. filicaulis</i> transcriptome)	14,025,763	13,167,862	3287	1244	112	44

Table 4 Comparison of protocols used to detect differential expression of RNA in *L. sessilifolius*.

Analysis type (mapping reference)	Reads mapped yellow stage	Reads mapped red stage	DE genes yellow stage ↑	DE genes Red stage ↑	Enriched GO terms yellow stage	Enriched GO terms red stage
TopHat (<i>L. japonicus</i> genome)	1,333,258	1,855,091	394	632	122	95
STAR (<i>L. japonicus</i> genome)	1,272,184	1,976,009	610	954	151	120
Trinity (<i>L. sessilifolius</i> transcriptome)	13,085,868	23,373,724	1175	1468	100	57

Table 5 Selected genes for enzymes involved in flavonoid biosynthesis in *L. filicaulis*. The first column lists differentially expressed (FDR<0.05) Trinity contigs. The second column lists the gene names produced by BLASTing and aligning contigs to the *Glycine max* CDS library, followed by the function of the enzyme which it codes for. The fourth and fifth columns list raw counts per gene at each colour stage, followed by the normalized log₂ fold change. Sequences for the Trinity contigs in this table are listed in Appendix 1.

Trinity contig	<i>Glycine max</i> homolog	Function	Counts per gene (Yellow stage)	Counts per gene (Red stage)	Normalized Log ₂ FC (Red/Yellow)
TR19889 c3_g1	Glyma.01g43880.1 <i>CHALCONE SYNTHASE (CHS)</i>	Produces naringenin chalcone from phenylalanine and acetate-pyruvate derivatives.	1385	19064	-4.1392
TR48518 c0_g1	Glyma.11G011500.1 <i>CHALCONE SYNTHASE (CHS)</i>	Produces naringenin chalcone from phenylalanine and acetate-pyruvate derivatives.	77	430	-2.8356
TR19889 c2_g1	Glyma.01G228700.1 <i>CHALCONE SYNTHASE (CHS)</i>	Produces naringenin chalcone from phenylalanine and acetate-pyruvate derivatives.	17	207	-3.9514
TR52191 c0_g1	Glyma.06G202300.1 <i>FLAVONOID 3'-HYDROXYLASE (F3'H)</i>	Catalyses naringenin to eriodictyol	1586	15638	-3.6580
TR38960 c0_g1	Glyma02g05450.1 <i>FLAVANONE 3HYDROXYLASE (F3H)</i>	Catalyses naringenin to dihydrokaempferol and eriodictyol to dihydroquercetin	1830	6123	-2.0988
TR29586 c0_g1	chr5.CM0077.210.r2.m <i>DIHYDROFLAVONOL 4-REDUCTASE (DFR)</i>	Catalyses dihydroquercetin to leucocyanidin and dihydrokaempferol to leucopelargonidin	148	10163	-6.4567

Trinity contig	<i>Glycine max</i> homolog	Function	Counts per gene (Yellow stage)	Counts per gene (Red stage)	Normalized Log ₂ FC (Red/Yellow)
TR50049 c0_g1	<i>ANTHOCYANIDIN SYNTHASE (ANS)</i>	Catalyses leucocyanidin to cyanidin and leucopelargonidin to pelargonidin	1010	14783	-4.2278
TR23981 c0_g1	Glyma17g34530.2 <i>MORE AXILLARY BRANCHES (MAX1)</i>	Positive regulator of several flavonoid biosynthesis enzymes	45	2987	-6.3605
TR11002 c0_g1	Glyma14g06750.1 <i>PRODUCTION OF ANTHOCYANIN PIGMENT (PAP1)</i>	Encodes MYB75 transcription factor involved in anthocyanin metabolism	475	54	2.7778

Table 6 Selected genes for enzymes involved in flavonoid biosynthesis in *L. sessilifolius*. See Table 5 for column descriptions. Sequences for the Trinity contigs in this table are listed in Appendix 2.

Trinity contig	<i>Glycine max</i> homolog	Function	Counts per gene (Yellow stage)	Counts per gene (Red stage)	Normalized Log ₂ FC (Red/Yellow)
TRINITY_DN5 4168_c0_g2	Glyma.11G011500.1 <i>CHALCONE SYNTHASE (CHS)</i>	Produces naringenin chalcone from phenylalanine and acetate-pyruvate derivatives.	65	4382	-6.1743
TRINITY_DN5 4168_c0_g1	Glyma.11G011500.1 <i>CHALCONE SYNTHASE (CHS)</i>	Produces naringenin chalcone from phenylalanine and acetate-pyruvate derivatives.	17	1134	-6.1510
TRINITY_DN5 2652_c1_g1	Glyma.19G105100.1 <i>CHALCONE SYNTHASE (CHS)</i>	Produces naringenin chalcone from phenylalanine and acetate-pyruvate derivatives.	145	660	-2.2875
TRINITY_DN5 2780_c0_g1	Glyma.06G202300.1 <i>FLAVONOID 3'- HYDROXYLASE (F3'H)</i>	Catalyses naringenin to eriodictyol	7	2073	-8.2859
TRINITY_DN5 2390_c0_g1	Glyma.02G048400.1 <i>FLAVANONE 3HYDROXYLASE (F3H)</i>	Catalyses naringenin to dihydrokaempferol and eriodictyol to dihydroquercetin	276	1893	-2.8794

Trinity contig	<i>Glycine max</i> homolog	Function	Counts per gene (Yellow stage)	Counts per gene (Red stage)	Normalized Log ₂ FC (Red/Yellow)
TRINITY_DN5 2789_c0_g1	Glyma.17G252200.1 <i>DIHYDROFLAVONOL 4-REDUCTASE (DFR)</i>	Catalyses dihydroquercetin to leucocyanidin and dihydrokaempferol to leucopelargonidin	19	434	-4.6063
TRINITY_DN5 3094_c0_g1	Glyma.11G027700.1 <i>ANTHOCYANIDIN SYNTHASE (ANS)</i>	Catalyses leucocyanidin to cyanidin and leucopelargonidin to pelargonidin	30	3761	-7.0659
TRINITY_DN3 9800_c0_g1	Glyma.13G082300.1 <i>FLAVONOL SYNTHASE (FLS)</i>	Mediates production of isorhamnetin, quercetin, and kaempferol from dihydroflavonol substrates	1601	112	3.7338

4 Discussion

4.1 The significance of pollination as a trigger for PACC

I present the first experimental evidence that PACC is accelerated by pollination in *L. filicaulis* (Figure 2.1); these measurements support those made of *Acmispon glaber* (Jones & Cruzan, 1999). This is strongly suggestive that PACC is a trait that has evolved via interaction with pollinators rather than an incidental characteristic of the aging process with no fitness consequences. This is unsurprising as many floral traits are indeed driven by pollinators (Schiestl & Johnson, 2013) and it is consistent with this trait being a widespread evolutionary convergence (Weiss, 1995). Viewing PACC as an evolved adaptive trait raises the question of the precise adaptive function that it confers. Brito et al. (2015) suggest that the retention of flowers is to increase attraction of insect pollinators at long range, and then at short range increase foraging efficiency by signalling flower quality to pollinators.

The demonstration of a pollination trigger for PACC also raises the question of what the precise nature of the trigger is. Pollination involves both physical movement of the petals of the flower (thigmostimulus), deposition of pollen on the stigma, subsequent growth of pollen tubes, and fertilization of the ovules. All these processes have physiological consequences and could be transduced into signals leading to PACC. Further experimental work is required to dissect the events of pollination that lead to the acceleration of PACC. Even without pollination, colour change eventually occurs. This suggests that there is an additional, late-acting or ‘failsafe’ trigger for the onset of PACC, one that presumably responds to late acting autogamy or to some aspect of floral aging.

4.2 *De novo* assembly is preferred for RNA-seq data without a reference genome

The two RNA-seq aligners used in our study, STAR and TopHat, performed similarly in aligning sequenced reads to the genome of a sister species. Our results are consistent with a recently published comparison of 11 commonly used alignment protocols (Engström et al., 2013). STAR 1-pass uses gene annotation information to identify exon splice junctions before aligning to the reference genome. TopHat also uses an annotation file to identify junctions, and furthermore, uses this information to align reads directly against annotated transcripts. Therefore, both

alignment protocols require a well-annotated genome, which currently does not exist for *L. filicaulis* or *L. sessilifolius*. Aligning *L. filicaulis* and *L. sessilifolius* reads to the genome of *L. japonicus* led to a substantial loss of data as the majority of reads remained unmapped. However, the number of *L. filicaulis* reads mapped to *L. japonicus* outperformed the number of mapped *L. sessilifolius* reads when considering the same protocol. This underscores the importance of phylogenetic relatedness when using a model genome to study a congener.

Our results strongly support the use of a *de novo* assembly of RNA-seq reads into a transcriptome (then mapping reads back onto the transcriptome) as opposed to attempts at cross-species mapping in *Lotus*. Contigs may then be identified by reciprocal BLAST (with putative genes checked by amino acid alignment) against the most closely related species with an annotated genome.

4.3 PACC is associated with the wholesale upregulation of the flavonoid pathway
Ojeda et al. (2013) previously studied gene expression of some flavonoid biosynthesis genes involved in PACC in *L. filicaulis* and *L. sessilifolius* by PCR. Our results support and extend their findings. With the exception of chalcone isomerase, at least one copy of each gene in the flavonoid biosynthetic pathway leading to cyanidin production is differentially expressed (increased) at the red stage of PACC (Figures 3.2 and 3.3). An alternative scenario might be that most of these genes might already be expressed at the yellow stage with the exception of one critical enzyme without which pigment cannot be made. Under this scheme the upregulation of a single enzyme would be enough to initiate colour change. Our results make it clear that this is not the case. Instead nearly the entire pathway is upregulated. Except for chalcone isomerase, all genes in the pathway have relatively low expression at anthesis and increase dramatically at the PACC stage (Tables 3.3 and 3.4; Figures 3.2 and 3.3). This implies that the genes in the pathway are co-regulated as a unit during PACC, possibly by a master regulator (see discussion below under MAX1).

4.4 PACC is associated with highly specific upregulation of members of gene families
Many of the enzymes involved are coded for by multiple genes in gene families. An example is DFR which has numerous copies (DFR1-DFR5) in *Lotus japonicus* (Shimada et al., 2005). However, during PACC only one paralog is upregulated. Other examples are obvious from

inspection of Figures 3.2 and 3.3. In these cases, the upregulation of specific copies, rather than all copies of a particular enzyme gene, seems to drive PACC. Some gene copies are apparently not expressed at either stage, but may be involved in tissue-specific expression elsewhere. It is worth noting that different gene copies within the *DFR* and *CHS* families are expressed between *L. filicaulis* and *L. sessilifolius*. This is also the case for several gene families of the carotenoid pathway (Figures 3.4 and 3.5). This lends support to the convergent evolution of PACC between *L. filicaulis* and *L. sessilifolius*, as hypothesized by Ojeda et al (2013; Figure 2). However, it remains possible that, regardless of differences in the upregulation of gene copies, some general aspect of the flavonoid pathway (e.g. how it is regulated) is a developmental pre-pattern for PACC that is common to *Lotus*. In this case, the apparent convergent evolution of PACC could be considered a “latent homology” (Nagy et al., 2014).

Because anthocyanins are highly versatile and serve a wide range of functions (Kong, 2003), it may be that gene duplication is key to the evolution of novel adaptations involving anthocyanins, including PACC. This may be especially likely for the putatively tetraploid *L. sessilifolius*, and although *L. filicaulis* is a diploid, it has had numerous whole genome duplication events in its history, like many plants. Neo-functionalization, where directional selection leads to novel function following a gene duplication event, may explain why only certain copies of flavonoid biosynthesis genes were actively transcribed during PACC. Alternatively the association of PACC with duplicated genes might be explained by escape from adaptive conflict (Des Marais & Rausher, 2008). In this scenario a single copy gene is adapted to simultaneously perform its ancestral function while performing some novel function, but is constrained in its ability to improve due to antagonistic pleiotropy. Following gene duplication both ancestral and novel functions have the potential to be better adapted. Marais and Rausher (2008) demonstrated that adaptive changes in the *DFR* gene of pre-duplication and post-duplication plant species are best explained as escape from adaptive conflict. In order to determine if PACC is a result of gene duplication, evidence for the ancestral and novel functions of flavonoid biosynthesis genes would need to be studied in *L. filicaulis*, *L. sessilifolius* and their relatives. In the case of *L. japonicus*, various duplicates of *DFR* differ in their responses to stress and organ specificity (Yoshida et al., 2010), but whether any such gene copies are functionally divergent in our study species, and whether any of these copies are flower-specific remains unstudied.

4.5 Upregulation of flavonoid biosynthesis may be antagonistic to other phenylpropanoid pathways

Expression of genes in the lignin biosynthetic pathway (LBP) suggests that a shift from the lignin branch of the phenylpropanoid pathway to the flavonoid branch may occur over the course of PACC (Figures 3.2 and 3.3). Transcriptome and metabolite profiling of *Fragaria x ananassa* (Duchesne) has demonstrated a metabolic interaction between flavonoid and lignin biosynthesis in strawberry fruit (Ring et al., 2013). Lignin biosynthesis competes for the same upstream substrate (coumaroyl-CoA) as flavonoid biosynthesis and, as would be expected under constant substrate availability, a shift from lignin production to flavonoid production can be measured in anthocyanin-rich tissues (Ring et al., 2013). Therefore it is plausible that in *L. filicaulis* and *L. sessilifolius* a drop in substrate availability may result in the downregulation of LBP genes observed during PACC.

Alternatively, expression of LBP genes may be reduced following flowering, regardless of PACC, due simply to developmental changes associated with aging.

4.6 Carotenoid biosynthesis is downregulated during PACC

Carotenoids are widespread in the yellow to orange flowers of numerous insectpollinated taxa, and likely play a role in the detectability of flowers by pollinators (Tananka et al. 2008, Cronk and Ojeda 2008). Consequently, the biosynthesis of carotenoids is generally well understood (Zhu et al. 2010). I detected the upregulation of the genes controlling the core carotenoid pathway at the yellow stage of PACC in both *Lotus* species – this is expected given the definitively yellow appearance of the flowers of both species at anthesis. To the extent that flavonoid biosynthesis may be crucial to the red stage of PACC, the coincident downregulation of carotenoid biosynthesis genes may be important in accentuating colour change. Alternatively, the downregulation of carotenoids may represent a senescence-related process separate from flavonoid production.

4.7 Is *MAX1* a master regulator of PACC?

PACC in *Lotus filicaulis* and *Lotus sessilifolius* involves a wholesale upregulation of the cyanidin branch of the anthocyanin biosynthetic pathway (ABP). Given the number of genes

involved it is unlikely that each gene is regulated entirely independently. Instead it is possible that triggering of a top-level regulator then initiates a regulatory cascade activating the entire pathway in concert. An obvious such potential regulator would be *PRODUCTION OF ANTHOCYANIN PIGMENT 1* (*PAP1*), a gene known to be a powerful regulator of anthocyanin production in *Arabidopsis* (Borevitz et al., 2000). Perhaps strangely, the *Lotus* homologue of *PAP1* is actually downregulated during PACC in *L. filicaulis* (Table 5) and therefore does not seem to be involved. However, of the other known regulators that are differentially expressed during PACC in our study, one, *MORE AXILLARY BRANCHES* (*MAX1*), is a potential candidate as a top level regulator. *MAX1* is strongly upregulated at the red stage of PACC in *L. filicaulis* (Table 5). The sequence recovered for *MAX1* in *L. filicaulis* lacks a DNA-binding domain and is therefore likely to encode a Cytochrome P450 rather than a transcription factor (Figure S27). Indeed, in *Arabidopsis*, *MAX1* encodes CYP711A1, and is known to be a positive regulator of the flavonoid pathway (Lazar & Goodman, 2006). *Arabidopsis* mutants *max1*, have been demonstrated to have 11 down-regulated flavonoid biosynthesis genes including F3H, F3'H, DFR, ANS, and UFGT (Lazar & Goodman, 2006). However, beyond what is known of Cytochrome P450s in flavonoid biosynthesis in general (Ayabe & Akashi, 2006), the exact biochemical mechanism in ABP of CYP711A1 remains unstudied. Interestingly *MAX1* was first characterized for its effect on axillary bud behaviour. However the link between ABP and axillary bud formation (if any) remains obscure, although regulation of auxin transport through inhibition of PIN-function by flavonoids in the bud is one possibility that has been suggested (Lazar & Goodman, 2006). It is known that fertilized ovules are a source of auxin (Sundberg & Østergaard, 2009), and that auxin can control anthocyanin biosynthesis (Liu et al., 2014). Counter to that, it has been shown that the *MAX* pathway acts as an entirely independent regulator of auxin transport (Bennett et al. 2006). The versatility of *MAX1* is further confirmed by the involvement of *MAX1* in the strigolactone signalling pathway (Challis et al., 2013).

4.8 The role of *FLS* in PACC in *Lotus sessilifolius*

FLS forms flavonols (quercetin, kaempferol, and isorhamnetin) from dihydroflavonols, the same substrate used by *DFR* to produce cyanidin. *FLS* has been shown to be antagonistic to anthocyanin accumulation, which increases in *FLS* knockouts (Lee et al., 2016). Ojeda et al. (2013) found that pre-colour change (yellow flowered) *Lotus* species from the Canary Islands

contain comparatively more flavonols (mostly isorhamnetin) than anthocyanins. Flavonols are generally colourless to the human eye but absorb strongly in the UV. Despite this Ojeda et al. (2013) demonstrated a reduction in UV reflectance of the petals following PACC (Ojeda et al. 2013). Whatever the mechanism, a shift from flavonol to cyanidin production may affect the detectability of the flowers by pollinators. It is of interest that *L. sessilifolius* and *L. filicaulis* differ markedly in the behaviour of *FLS*. The reduction of flavonol production with floral age in *L. sessilifolius* (a Canary Island species) is the basis of the suggestion made by Ojeda et al. (2013) that an *FLS* mediated switch to anthocyanins from flavonols may be occurring.

However, the absence of *FLS* expression in *L. filicaulis* indicates that a different mechanism prevails in that species, although I cannot rule out the possibility that RNA levels were too low to be detected. The differential expression of *FLS* at the yellow stage of *L. sessilifolius* supports the notion that there is an *FLS*-mediated switch from flavonols to anthocyanins (which may therefore be the basis of an ecological function of PACC to signal floral viability to insect pollinators). It also implies that, in this species at least, anthocyanin biosynthesis is antagonistic to the production of flavonols. Furthermore, Ojeda et al. (2013) demonstrated that the transition from bee- to birdpollinated *Lotus* spp. involves the increased biosynthesis of anthocyanins at the expense of flavonol production.

4.9 The role of red in PACC

Red anthocyanin pigmentation of flowers has long been associated with bird pollination. For instance, the scarlet colouration of bird pollinated Lamiaceae species has been attributed to both cyanidin and pelargonidin (Saito & Harborne, 1992). It is thought that in bird pollination syndromes red serves not only as a bird-attractant but also a bee deterrent (Rodríguez-Gironés & Santamaría, 2004), though red-coloured objects should not be considered as invisible to bees (Chittka & Waser, 1997; Martínez-Harms et al., 2010). It is possible that in plants exhibiting ornithophily, red acts as a common signal of low nutritional status to bees. In bird pollinated plants highly dilute nectar is also thought to be an adaptation to deter bees (Bolten & Feinsinger, 1978) as this is difficult for bees to process, and energetically expensive to convert to concentrated storage forms (i.e. honey). There is an intriguing parallel here between PACC and bird pollination; in both cases red pigmentation is potentially signalling aspects of nutritional

status to bees (Chittka & Waser, 1997; Martínez-Harms et al., 2010). Red and related colours that are common in PACC may therefore constitute a preferred signal of foraging unsuitability, so incurring an adaptive advantage to the plant by optimizing pollinator foraging on young, newly opened, yellow flowers.

5 Conclusion

In this thesis I 1) present an analysis of the transcriptomics of post-anthesis colour change in *Lotus filicaulis* and *Lotus sessilifolius* and 2) demonstrate that PACC is mediated by pollination in *L. filicaulis*.

I present the first evidence that PACC is mediated by pollination in *L. filicaulis*. This is the foundation for understanding the ecological function of PACC in *Lotus*. Further work will need to address the physiological basis of the trigger(s) of PACC, e.g. is PACC triggered by the deposition of pollen on the stigma, or some other physical cue?

This is the first transcriptome-wide view of how gene expression is modulated through PACC for any species; the contigs listed in Appendices 1 and 2 are the first coding sequences available for these *Lotus* species. These sequences are hypotheses, which allows for both verification of this thesis and further work – especially if gene expression in other *Lotus* species is to be studied. Thus, this is an important contribution to better understand the evolution of pollination in *Lotus*.

Bibliography

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* 38: W7-13.

Altschul S, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389-3402.

Anders S, Pyl PT, Huber W. 2015. HTSeq – A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-169.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Version 0.63. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.

Ayabe S, Akashi T. 2006. Cytochrome P450s in flavonoid metabolism. *Phytochemistry Reviews* 5: 271-282.

Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1-48

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57: 289–300.

Bennett T, Sieberer T, Willett B, Booker J, Luschnig C, Leyser O. 2006. The Arabidopsis MAX pathway controls shoot branching by regulating auxin transport. *Current Biology* 16(6): 553-563.

Bolten A, Feinsinger P. 1978. Why Do Hummingbird Flowers Secrete Dilute Nectar? *Biotropica* 10: 307.

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., & Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *The Plant Cell* 12: 2383-2393.

Brito V, Weynans K, Sazima M, Lunau K. 2015. Trees as huge flowers and flowers as oversized floral guides: the role of floral color change and retention of old flowers in *Tibouchina pulchra*. *Frontiers in Plant Science* doi: 10.3389/fpls.2015.00362.

Challis RJ, Hepworth J, Mouchel C, Waites R, Leyser O. 2013. A role for more axillary growth1 (MAX1) in evolutionary diversity in strigolactone signaling upstream of MAX2. *Plant Physiology* 161(4): 1885-1902.

Chittka L. 1996. Does Bee Color Vision Predate the Evolution of Flower Color? *Naturwissenschaften* 83: 136-138.

Chittka L, Waser N. 1997. Why red flowers are not invisible to bees. *Israel Journal of Plant Sciences* 45: 169-183.

Cronk QCB, Ojeda ID. 2008. Bird-pollinated flowers in an evolutionary and molecular context. *Journal of Experimental Botany* 59: 715-727.

Des Marais D, Rausher M. 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* 454: 762-765.

Dobin A, Davis C, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras T. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15-21.

Du Z, Zhou X, Ling Y, Zhang Z, Su Z. 2010. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research* 38: W64-W70.

Duke J. 1981. Handbook of legumes of world economic importance. New York, USA: Plenum Press.

Engström P, Steijger T, Sipos B, Grant G, Kahles A, Alioto T, Behr J, Bertone P, Bohnert R, Campagna D et al. 2013. Systematic evaluation of spliced alignment programs for RNA-seq data. *Nature Methods* 10: 1185-1191.

Faegri K, Van Der Pijl L. 1979. Principles of Pollination Ecology. Pennsylvania, USA: Elsevier Science.

Farzad M, Griesbach R, Hammond J, Weiss M, Elmendorf H. 2003. Differential expression of three key anthocyanin biosynthetic genes in a color-changing flower, *Viola cornuta* cv. Yesterday, Today and Tomorrow. *Plant Science* 165: 1333-1342.

Ferreira J, Pedrosa-Harand A. 2014. *Lotus* Cytogenetics. In: Tabata, S, Stougaard, J, eds. The *Lotus japonicus* Genome, New York, USA: Springer Heidelberg, 9-20.

Feller A, Machemer K, Braun EL, Grotewold E. 2011. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal* 66: 94-116.

Goecks J, Nekrutenko A, Taylor J, The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. 2010. *Genome Biology*. 11(8): R86.

Goodstein D, Shu S, Howson R, Neupane R, Hayes R, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N et al. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* 40: D1178-D1186.

Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644-652.

Grigson, G. 1975. The Englishman's Flora. London, UK: Hart-Davis MacGibbon Ltd.

Jones C, Cruzan M. 1999. Floral Morphological Changes and Reproductive Success in Deer Weed (*Lotus scoparius*, Fabaceae). American Journal of Botany 86: 273.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelly R, Salzberg, SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions, and gene fusions. Genome Biology 14: R36.

Kong JM. 2003. Analysis and biological activities of anthocyanins. Phytochemistry 64: 923-933.

Lamesch P, Berardini T, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander D, Garcia-Hernandez M et al. 2011. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research 40: D1202-D1210.

Lazar G, Goodman H. 2006. MAX1, a regulator of the flavonoid pathway, controls vegetative axillary bud outgrowth in Arabidopsis. Proceedings of the National Academy of Sciences 103: 472-476.

Lee WJ, Jeong, C. Y., Kwon, J., Van Kien, V., Lee, D., Hong, S. W., & Lee, H. (2016). Drastic anthocyanin increase in response to PAP1 overexpression in fls1 knockout mutant confers enhanced osmotic stress tolerance in Arabidopsis thaliana. Plant Cell Reports 35: 2369-2379.

Li B, Dewey C. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12: 323.

Li J, Dai X, Liu T, Zhao P. 2011. LegumeIP: an integrative database for comparative genomics and transcriptomics of model legumes. Nucleic Acids Research 40: D1221-D1229.

Liu Z, Shi M, Xie D. 2014. Regulation of anthocyanin biosynthesis in Arabidopsis thaliana red pap1-D cells metabolically programmed by auxins. Planta 239: 765-781.

Maddison WP, Maddison, DR. 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04. <http://mesquiteproject.org>.

Martínez -Harms J, Palacios A, Marquez N, Estay P, Arroyo M, Mpodozis J. 2010. Can red flowers be conspicuous to bees? *Bombus dahlbomii* and South American temperate forest flowers as a case in point. *Journal of Experimental Biology* 213: 564-571.

Mohan Ram HY, Mathur G. 1984. Flower Colour Changes in *Lantana camara*. *Journal of Experimental Botany* 35: 1656-1662.

Nagy LG, Ohm RA, Kovács GM, Floudas D, Riley R, Gácsér A, Sipiczki M, Davis JM, Doty SL, De Hoog GS, Lang BF. (2014) Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. *Nature communications* 5: 4471

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302: 205-217.

Ober D. 2005. Seeing double: gene duplication and diversification in plant secondary metabolism. *Trends in Plant Science* 10: 444-449.

Ojeda, I. Santos-Guerra, A. 2011. The intersection of conservation and horticulture: bird-pollinated *Lotus* species from the Canary Islands (Leguminosae). *Biodiversity and Conservation* 20: 3501-3516.

Ojeda I, Santos-Guerra A, Caujapé-Castells J, Jaén-Molina R, Marrero Á, C. B. Cronk QCB. 2012. Comparative Micromorphology of Petals in Macaronesian *Lotus* (Leguminosae) Reveals a Loss of Papillose Conical Cells during the Evolution of Bird Pollination. *International Journal of Plant Sciences* 173: 365-374.

- Ojeda, I., Santos-Guerra, A., Jaén-Molina, R., Oliva-Tejera, F., Caujapé-Castells, J., & Cronk, Q. 2012. The origin of bird pollination in Macaronesian *Lotus* (Loteae, Leguminosae). *Molecular phylogenetics and evolution* 62: 306-318.
- Ojeda I, Santos-Guerra A, Oliva-Tejera F, Valido A, Xue X, Marrero A, CaujapéCastells J, Cronk QCB. 2013. Bird-pollinated Macaronesian *Lotus* (Leguminosae) evolved within a group of entomophilous ancestors with post-anthesis flower color change. *Perspectives in Plant Ecology, Evolution and Systematics* 15: 193-204.
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rice P, Longden I, Bleasby A. 2000. EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* 16(6): 276-277.
- Ramsay NA, Glover BJ. 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Science* 10: 63-70.
- Ring L, Yeh S, Hucherig S, Hoffmann T, Blanco-Portales R, Fouche M, Villatoro C, Denoyes B, Monfort A, Caballero J et al. 2013. Metabolic Interaction between Anthocyanin and Lignin Biosynthesis Is Associated with Peroxidase FaPRX27 in Strawberry Fruit. *Plant Physiology* 163: 43-60.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1), 139-140.
- Rodríguez-Gironés MA, Santamaría L. 2004. Why Are So Many Bird Flowers Red? *PLoS Biol* 2(10): e350. doi:10.1371/journal.pbio.0020350
- Saito N, Harborne J. 1992. Correlations between anthocyanin type, pollinator and flower colour in the labiatae. *Phytochemistry* 31: 3009-3015.

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K et al. 2008. Genome Structure of the Legume, *Lotus japonicus*. DNA Research 15: 227-239.

Schiestl F, Johnson S. 2013. Pollinator-mediated evolution of floral signals. Trends in Ecology & Evolution 28: 307-315.

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J et al. 2010. Genome sequence of the paleopolyploid soybean. Nature 463: 178-183.

Shimada N, Sasaki R, Sato S, Kaneko T, Tabata S, Aoki T, Ayabe S. 2005. A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. Journal of Experimental Botany 56(419): 2573-2585.

Spelt C, Quattrocchio F, Mol J, Koes R. 2002. ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. Plant Cell 14: 2121-2135.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and postanalysis of large phylogenies. Bioinformatics 30: 1312-1313.

Sundberg E, Østergaard L. 2009. Distinct and dynamic auxin activities during reproductive development. Cold Spring Harbor Perspectives in Biology 1(6): a001628.

Suzuki H, Sasaki R, Ogata Y, Nakamura Y, Sakurai N, Kitajima M, Takayama H, Kanaya S, Aoki K, Shibata D et al. 2008. Metabolic profiling of flavonoids in *Lotus japonicus* using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry. Phytochemistry 69: 99-111.

Tan J, Wang M, Tu L, Nie Y, Lin Y, Zhang X. 2013. The Flavonoid Pathway Regulates the Petal Colors of Cotton Flower. PLoS ONE 8: e72364. Tanaka Y, Sasaki N, and Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. The Plant Journal doi 10.1111/j.1365313X.2008.03447.x

Weiss M. 1991. Floral colour changes as cues for pollinators. Nature 354: 227-229.

Weiss M. 1995. Floral Color Change: A Widespread Functional Convergence. American Journal of Botany 82: 167-185.

Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. New York, USA: Springer-Verlag. <http://www.ggplot2.org>.

Yekutieli D, Benjamini Y. 1999. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. Journal of Statistical Planning and Inference 82: 171-196.

Yoshida K, Iwasaka R, Shimada N, Ayabe S, Aoki T, Sakuta M. 2010. Transcriptional control of the dihydroflavonol 4-reductase multigene family in *Lotus japonicus*. Journal of Plant Research 123: 801-805.

Young N, Debellé F, Oldroyd G, Geurts R, Cannon S, Udvardi M, Benedito V, Mayer K, Gouzy J, Schoof H et al. 2011. The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480: 520-524.

Zhu C, Bai C, Sanahuja G, Yuan D, Farré G, Naqvi S, Shi L, Capell T, Christou P. 2010. The regulation of carotenoid pigmentation in flowers. Archives of Biochemistry and Biophysics 504: 132-141.

Appendix 1

Sequences for Trinity contigs listed in Table 5

>TR19889_c1_g1 [organism=Lotus filicaulis] CHS

CTTGGCTGAGAACAACAAAGGTGCTCGTGTACTCGTTGTTTGTCTGAACTTACTGCAGTTACCTTCC
GTGGCCCTAGTGACACTCACCTAGACAGCCTTGTGGGCAAGCATTGTTTGGAGACGGAGCAGCCGCA
CTCATTGTTGGTTCCGATCCAGTACCAGAAGTTGAGAAGCCTTTGTTTGGAGCTAGTTTGGACTGCACA
GACTATTGCTCCAGATAGTGAAGGAGCCATTGATGGCCACCTTCGTGAAGTTGGATTGACATTTTCATC
TCCTTAAAGATGTTCTGGG

>TR48518_c0_g1 [organism=Lotus filicaulis] CHS

ATGGTGAGTGTAGCTGAGATTCGCAAGGCTCAAAGAGCTGAAGGCCAGCAACCATCTTTGCTATTGG
CACTGCAAATCCTCCTAACTGTGTTGATCAAAGCACCTATCCTGATTTCTACTTCAGAGTCACCAACA
GCGAACACAAAACAGAGCTCAAGGAAAAATTTAGCGCATGTGTGATAAGTCTATGATCAAGAAGAGA
TACATGCACTTAACAGAAGATCTTTTGAAAGAGAACCCTAACATGTGCGCTTATATGGCGCCATCTTT
GGATGCTAGGCAAGACATGGTGGTGGTAGAGGTACCTAGACTAGGCAAAGAAGCTGCTGTCAAGGCTA
TA

>TR19889_c3_g1 [organism=Lotus filicaulis] CHS

ATGGTGAGTGTAGCTGAGATTCGTAAGGCTCAAAGGGCAGAAGGCCAGCAACCATCCTGGCTATTGG
CACTGCAAATCCACCAAACCTGTGTGGACCAAAGCACTTATCCTGATTTCTACTTCAAAATCACAAACA
GCGAGCACATGACAGAGCTTAAGGAGAAGTTTCAGCGCATGTGTGACAAGTCCATGATCAAGAAGAGA
TACATGTACTTGAATGAAGAAATTTTGAAAGAGAACCCCAACCTTTGTGCTTATATGGCACCATCTTT
GGATGCTAGGCAGGACATGGTGGTGGTAGAGGTACCAAGGTTAGGGAAAGAAGCTGCAACAAAGGCTA
TAAAAGAATGGGGCCAGCCAAAATCTAAGATTACTCACTTAATCTTTTGCACCACAAGTGGTGTAGAC
ATGCCTGGTGTGCTGATTATCAACTCACCAAACCTCTGGGTCTTCGCCCATCTGTGAAAAGGTACATGAT
GTACCAACAAGGGTGCTTTGCAGGTGGCAGGTGCTTCGTTTGGCTAAAGATTTGGCTGAGAACAACA
AAGGTGCACGTGTGCTAGTTGTTTGTCTGAGATTACTGCGGTTACATTCCGTGGCCCTAATGACACT
CACCTAGACAGCCTTGTGGGGCAAGCTTTGTTTCGGAGATGGAGCAGCTGCAGTCATTGTTGGTTCTGA
TCCAGTGCCTGAAATTGAGAAACCTTTGTTTGAAGTGTGTTGACTGCACAACTATTGCTCCAGATA
GTGACGGAGCCATCGATGGTCACCTTCGTGAAGTTGGATTGACGTTCCATCTCCTTAAAGATGTTCTCCT
GGGATTGTTTCAAAGAACATTGAGAAAGCACTAATTGAGGCCTTCCAACCATTAGGCATATCTGATTA
CAACTCAATTTTTTGGATTGCACACCCAGGCGGCCAGCAATTCTGGACCAAGTTGAGCAGAAGTTGA
GCTTGAAACCTGAAAAGATGAAAGCCACTAGAGAAGTGCTAAGTGAATATGGGAACATGTCAAGTGCA
TGTGTCCTATTTCATCTTAGATGAAATGAGAAAGAAATCAGCTCAAGATGGACTCAAAACCACTGGTGA
AGGACTTGAATGGGGTGTGCTATTTCGGTTTCGGACCTGGACTTACCATTGAAACTGTTGTTTTGCGTA

GTGTGGCTATTTAATAT

>TR52191|c0_g1 [organism=Lotus filicaulis] F3'H

ATGTTCC

CATGGATGATCATTGGCTTTGCCACCATCACATTCCCTCATCTTCATACACCGTGTCGTGA
AGTTCTCAACACGTCCCTCACTCCCACTCCCTCCGGGCCAAAACCATGGCCATAATAG
GAAACTTTCCCATATGGGCCCCGTGCCGCACCACTCCCTCGCAGCTTTGGCCCCGAGCCC
ACGGCCCCGCTTATGCACCTCAAATTGGGCTTTGTGGACGTTGTGGTGGCGGCGTCAGCCG
CCGTGGCGGAGCAGTTTTTTGAAGGTTTCATGATGCAAATTTTAGTAGCCGGCCGCCTAATG
CCGGAGCTAAGTACATAGCTTATACTATCAGGATCTCGTGTTTCGCTCCGTACGGTGCAC
GGTGGCGATATCTCCGAAAATCACCAACCTCCACCTCTTCTCCGGCAAGGCTTTGGATA
ATTTCAAACACTTGCCTCAGGAAGAGGTATCGAGATTGACACGCAACATATCAAAGTCCA
ACTCAAAGCTGTGAACCTGGGACAACCTACTGAATGTATGCACTACCAACGCATTATCCA
GGGTAATGATCGGAAGAAGAGTGTTCAACGACGGCGACGGTGGGTGTGATCCTAGAGCTG
ATGAGTTTAAGGCCATGGTGGTGGAGTTAATGGTGTAGCGGGTGTTCACATAGGTG
ACTTCATTCTCTCTTGGAGTGGCTAGACCTTCAAGGGGTTCAGCTAAGATGAAGAAAT
TGCACAATAGGTTTTGATGAATTTTTAACCAGCATTATTGAGGAACACAATACTTCTTCCA
AGAGTGAGAATCATAAGGATTTATTGAGTACGTTGCTGTCGCTTAAAGATGTTTCCTGATG
ATGATGGAAACAGACTCAATGATATTGAGATCAAAGCACTACTTCTGAACATGTTTCACAG
CTGGGACTGACACGTCAGCTAGCACACAAGAATGGGCCATTGCAGAACTGATCCGTAGCC
CAAGAATTCTAGCCCAGGTCCAACAAGAATTGGACACAGTTGTGGGCCGAGAAAGAAACG
TGAGAGAAGATGACTTGCCTCATCTCCCATACTTGCAAGCAGTGGTGAAAGAAACCTTCC
GTCTTCACCCATCAACCCCTCTTTCACTGCCACGTGTGCATCTGAGAGCTGTGAGGTAC
TGGGCTACCACATACCAAAGGGGTCCACTCTCTTGGTGAACGTGTGGGCCATAGCCCGTG
ACCCTAAAGAATGGGCTGAGCCATTGGAGTTCAAGCCCGAAAGGTTCTAGAGGGTGATA
AGGTTGATGTTGATGTTAAGGGCAATGACTTTAAGGTGATACCCTTTGGTGTGACGTA
GGATATGTGCTGGGATGAGCCTTGGGCTTCGTATGGTTCAGCTGTAACTGCAACTCTGG
TCCATTCTGTTTAACTGGGAGCTAGAAAATGGGCTTAACCATGAAAAGCTCAACATGGATG
AAGCTTATGGGCTAACATTGCAACGGGCTGTGCCTTTGTGAGTGTACTCTAGGCCCAGGC
TCTCGCCACATGTGTATGCAGCCTCTCATTGA

>TR38960|c0_g1 [organism=Lotus filicaulis] F3H

ATGGCATCATTCA

AACCCAAAACCTCTACCACACTGGCTCAACAAAACACCCTTGAGTCCAGCTTCGTCCGTG
ACGAAGACGAGCGCCCAAAGGTCGCTTACAACAATTTTCAGCAATGAGATCCCGGTAATTT
CCCTCGCCGGAATCGACGAGGTCGATGGCCGGAGATCCGAGATTTGTAACAAGATTGTTG
AAGCTTGTGAGAATTGGGGTATTTTTTCAGGTTGTTGATCATGGGGTCGACACTGAATTGG
TTTCCCACATGACCACTCTTGCTAAAGAGTTCTTTGCTCTTCCACCGGAGGAGAAGCTTC

GGTTTGACATGACCGGTGGCAAAAAGGGTGGTTTTCATCGTCTCCAGCCACCTCCAAGGAG
AATCAGTGCAGGATTGGAGAGAGATAGTGACATACTTTTCATACCCAATCAGGAACAGGG
ACTACTCACGGTGGCCGGACACCCAGCAGGGTGGAAGGCGGTGACGGAGGAGTACAGCG
AGAAGCTGATGGGTTTAGCATGCAAGCTGTTGGAGGTGTTGTCTGAGGCAATGGGGTTAG
AGAAAGAGGCTCTAACAAAGGCGTGTGTGGATATGGAGCAGAAGGTTGTGGTGAATTATT
ACCCAAAATGCCCTCAACCTGACCTCACTCTTGGGCTGAAGCGACACACTGATCCTGGCA
CCATTACTCTGCTTCTTCAAGATCAGGTGGGTGGGCTTCAAGCTACTAGGGATAATGGGA
AGACATGGATCACTGTGCAGCCTCTTGAAGGAGCTTTTGTGTCAATCTTGGTGACCATG
GTCATTATCTTAGCAATGGAAGGTTCAAGAACGCTGACCACCAAGCAGTGGTGAACCTCAA
ACTCAAGCCGTTTATCGATCGCCACATTCCAGAACCCGGCACCAGATGCAACTGTATACC
CTTTGAAGGTTTCGAGAGGGAGAGAAGTCAGTAATGGAAGAACCAATCACATTTGCTGAAA
TGTACAGGAGGAAGATGAGCAAGGACATTGAGCTTGCTCGGATGAAGAAGCTGGCTAAGG
AGAAGAACTGCAGGATTTGGAGAAGGCCAACTTGAACCCAAGCCCATGAATGAGATCT
TTGCTTAAG

>TR29586|c0_g1_i1 [organism=Lotus filicaulis] DFR
ATGGGTTTCAGTACCAGAAACAGTTTGC GTTACCGGAGCTGCAGGTTTCATCGGGTCATGGC
TTGTCATGAGACTCATGGAACGTGGCTATATGGTTTCGCGCCACTGTAAGAGACCCTGCTA
ACATGAAGAAGGTGAAGCATTGCTGGAAGTGCAGAGGCAAAGACCAAGCTGACTCTGT
GGAAAGCTGATCTTGCTGAAGAGGGAAGCTTTGATGAAGCCATCAAAGGGTGCAGTGGAG
TTTTCCATGTGGCCACACCCATGGATTTTGAGTCCAAGGACCCAGAGAATGAAGTGATTA
AACCTACCATAAACGGGTACTAGACATCATGAAGGCATCCCAGAAGGCCAAAAGTGTAC
GGAGGCTGGTTTTCACATCCTCTGCAGGTACCCTCAACGTCATAGAGCACCAAAAACAAA
TGTTTGATGAGAGCTGCTGGAGCGATGTTGAGTTCTGCCGAAGAGTAAAGATGACCGGTT
GGATGTATTTTGTTCAAAGACACTGGCAGAGCAAGAAGCGTGGAATTTGCCAAAGAGC
ATGGCATTGACTTCATTACAATCATTCCACCCCTTGTTGTTGGTTTCCTTTCTAATGCCGA
CAATGCCACCAAGCCTAATCACTGCTCTTTCTCCTATCACCGGAAACGAGGCCCATTTATT
CGATCATAAAGCAAGGCCAATATGTCCACTTGGATGATCTTTGTCTTGCCCATATTTTCC
TGTTTGAGCACCCCTGAATCAGAAGGAAGGTATATCTGTAGTGATCTGAGGCTACTATCC
ATGACATTGCAAAGTTAATCAACTCAAAGTATCCAGAGTACAATATCCCCACTAAGTTCA
AGAATATTCCAGATGAATTGGAGCTTGTCAGATTTTCATCAAAGAAGATCAAAGACATGG
GATTCGAATTTAAATACAGCTTAGAGGATATGTACACTGGAGCAATTGACACGTGCAAGG
AAAAAGGGCTTCTTCCTAAAGCTGCAGAAAATCCAAGCAATGGCAAATAA

>TR50049|c0_g1_i1 [organism=Lotus filicaulis] ANS
ATGGCGCC

TACTGTTGTTGAGAGAGTTGAAAGCTTATCGGGCAGTGGGATACAATCTATCCCAAAGGA
 ATATGTGAGGCCAAAAGAAGAGCTGGCAAACATAGGCGACGTTTTTTGAGGAGGAAAAGAA
 GGTTGGGCCTCAAGTTCCAACGATTGATTTGAAAGAGATAGACTCGCCGGACGAGTTTGT
 TCGGGCCAAGTGCAGGGAGAAGCTTAGGAAGGCAGCCGAGGAGTGGGGGGTGATGCACCT
 GGTGAATCATGGGATCCCGGATGAGCTCCTAAACCAATTGAAGAGTGCTGGGGCAGAGTT
 TTTCTCTTTGCCGGTCGAGGAGAAGGAGAAGTATGCAAATGACCAGGCGGCCGGGAATGT
 GCAAGGGTACGGGAGCAAGCTGGCCAACAATGCCAGTGGTCAGCTTGAATGGGAAGATTA
 CTTCTTCCACCTTATCTTTCCAGAAGACAAGCGTGACCTCTCCATCTGGCCCAAGACACC
 ATCATATTATACGGAAGTTACCAGCGACTATGCAAGGCGACTAAGAGTGCTTGCGAGCAA
 GATACTGGAAGTGCTATCTCTGGAATTGGGTCTCGAAGAAGGAAGGTTAGAGAAGGAAGT
 TGGTGAATGGAAGAGCTTCTACTTCAGATGAAAATCAACTACTACCCAAAATGCCCCCA
 GCCAGAGCTAGCTCTCGGAGTCGAAGCCCACACCGACATAAGCGCACTCACTTTCCTCCT
 CCACAACATGGTCCCAGGCCTGCAACTTTTCTACGAAGGAAAATGGGTACGGCAAAATG
 CGTTCCTGATTCAATCCTCATGCACATTGGTGACACCACTGAGATTTTAAGCAATGGTAA
 GTTCAAGAGCATTCTTCACAGGGGTTTGGTGAACAAGGAAAAAGTTAGAATATCTTGGGC
 AGTGTTCTGTGAACCCCTAAGGAGAAAATCATCCTCAAACCACTGCCTGAGCTTGTGAC
 TGAGACCGAACCAGCGCGGTTTCCGCCACGCACTTTTGCTCAGCATATTCATCACAAACT
 GTTCAGGAAGGACCAGGAAGCTTCAGCCCAATCCAAATGA

>TR23981|c0_g1 [organism=Lotus filicaulis] MAX1

ATGGTG

TTCATGGATTTTGAATGGTTGTTCCAAATTCCAAGTGTCCCTTGGTCTTCAGCAATGTTT
 ACCTTGCTGGCTACAATAGGAGGATTCTCAGTGTATCTGTATGGGCCCTACTGGGGTGTG
 AGGAAAGTTCCAGGTCCACCATCTGTGCCTCTGATAGGACACCTTCCCTTGCTGGCTAAG
 TATGGCCCTGATGTCTTCTCAGTCCTTGCCAAGCAATATGGCCCAATCTACAGATTTTCA
 ATGGGAAGACAACCTCTAATAATCATAGCAGATGCAGAGCTTTGCAAAGAGGCAGGCATC
 AAAAAATTTAAAGATATTACAAACAGAAGCATTCCCTCTCCCATTTCTGCATCCCCCTT
 CACCAAAAGGGTCTCTTCTTACCAAAGATTACAAATGGTCTACCATGAGAAACACCATA
 TTATCACTCTACCAGCCATCACACTTATCCAGATTAGTGCCAACCATGCAGTCATTCTA
 GAATCAGCAACTCAAATCTAGATTCTCAAAAAGAAGACTTTATCTTTTCTAATCTTTTCG
 CTCAGCTTGGCAACTGATGTGATTGGACAAGCAGCATTTGGAGTTGATTTTGGCCTCTCG
 AAGCCTCAACCAGTGTGTGATGAAATCAAGAGTGTTAACAAGGAAGTAAGGGATTCAAGT
 ACTGGTAATGAAGTATCAGATTTTATTAAACCAACACATTTACTCCACAACCCAACTCAAG
 ATGGACCTATCTGGTTCCTTCTCCATCATACTGGGTTTACTTGTCCCAATTCTCCAGGAG
 CCATTTAGGCAGGTTCTCAAGAGAATTCCAGGCACCATGGACTGGAAAATCGAGCGTACT
 AACCAGAAAATAAGTGGCCGTCTTGATGAGATTGTAGAGAAGAGAATGAAAGATAGGGTC
 AGAAGTTCAAAGGACTTCTTGTCACTCATTCTAAATGCAAGAGAATCAAAAATGTTTCA
 GAAAATGTGTTACACCTGACTATATCAGTGCTGTTACTTATGAGCACCTGCTTGCAGGA

TCAGCAACCACATCCTTTACCTTGTCTTCCATTGTCTATTTGGTTGCTGGGCATCCAGAA
GTTGAGAAAAAGATGCTTCAAGAGATTGATGGGTTTGGTCCAGTAGATCAGACACCAACT
TCTCAAGATCTTCAAGAGAAGTTTCCTTACCTTGATCAGGTGATCAAAGAGGCCATGAGA
TATTACACAGTCTCCCCATTGGTTGCAAGAGAAACATCAAATGAAGTAGAGATTGGAGGT
TACCTTCTTCCAAAGGGGACTTGGGTCTGGTTAGCGCTTGGAGTTGTAGCAAAAGATCCC
AGAAACTTTCCGGAGCCAGAAAAGTTCAAACCGGAGAGGTTTGATCCCAAATGTGAAGAA
ATGAAAAGAAGACATCCTTATGCTTTCATACCATTGGAATTGGTCCTCGGGCTTGCATT
GGTCAGAAATTTCCCTGCAAGAAATCAAGCTTCTCTGATTCACTTATACAGGAAATAC
TTGTTTCGGCATTACCTAATATGGAGAATCCCTTAGAACTTGAATATGGCATAGTGCTC
AACTTCAAGCACGGTGTCAAGGTTAGAGCCATAAAAAGAACAGAAAGGAGCTGCTAA

>TR11002|c0_g1_i1[organism=Lotus filicaulis] PAP1
ATTGGCTCTTACCTTTCTACAATATCTCTTTCTTTCTCTTGCTCTTGGCTTCTTTATATT
TTGATATTTTCTTTCTTTCTCAGTGGAGCTCAACTTTGTAGTAGCAGCCTCACTGAAAACAT
TTACCTTTCTCTTCTGATATTTTGGTATTCTCAGAACATCTTTTATTGTTCTAGTTTCAGA
ATTCATCACACTCACATCAAAATTCTGGCACACCACCATCATTTTCCTAAGATTCTCTAC
TTAAAAGGTATGACTTTTATAGATTCTCAACTTTCTCTTGATCACACATAAAAGATTCTA
ACTTATGTGAGAATCAAAGATCTTTGGTTTTCTTCTTTTCTGTGATTCTCTTCTGTAATCA
CAAGAGCACCCCTTTTCTAACTCTTATTTCCATCTCTTGAGGAACTTCATTCAAGTGATA
GGAGAGTTTTGTATGCTTGTGGCTGTGGTTTAAGTTATGGATTTGGAAACACTCTATTCC
CCTTGTTTTATGTCAAATTCAAATTGGTTTTGTCCAAGAGAGTGCTCACAAACACAGAATGG
AGTAGAGAAGATAACAAGAGGTTTGAAAGTGCCCTTGCTATTTATGATAAGGACACACCA
GATAGATGGCTGAATGTGGCTGCAATGATTCTGGGAAGACTGTGCTTGATGTGATCAAG
CAGTACAGGGAATTGGAAGAAGATGTGGGTGAAATTGAAGCAGGGCATGTTCTGTCCCT
GGATATCATTCCTCTTCTTTTACCTTTGAAGTTGTTGAGAACCAGAACTTTGATGGACTC
AAAAGGAAGCCTGGAACAACTCTCAGGGGTTCTGATCATGAGAGGAAGAAAGGAGTACCG
TGGACCGAAGAAGAACACAAACGCTTTCTGATGGGACTTCTAAAGTATGGTAAAGGGGAC
TGGAGAAACATCGCTCGCAATTTTGTATGACAAAGACTCCCACACAAGTAGCTAGCCAT
GCTCAGAAGTACTACATAAGGCAAAAGGTTTCTGGAGGAAAAGATAAGAGGAGACCAAGC
ATCCATGATATAACCACTGTTAATCTTACAGAACTAGTACCACATCAGAAAACAACAAG
CCACTTTCATTCAAAAGTACTAGTAACTACAACAGTGGCTCACTCATGGTTTTCAACCCA
AATTGTGATGACTTGTTGATGATGCCATCTTCATCTGATATCATCACTTCCAAGACCCTT
AACTTCAAGGGCAGGATCTTTATGATTGCTCTTTACATGAGGCTTATGCTAAGTTAAAA
ATTCCTAGTTTCAGGGCAGCACCTAGAACTTCAATAAGGAAGCTGTTTTTGGTATTCAT
GTACTATAAGAATGCCAACTATGACAATGTTCAAATTAAGGCAAACCAATTCCACTGCAT
AGATTATTGTTCTGTAGGATTTTGATAATGTAGATGAAATGAAGCTACTTGTTTCATGAA
CATCCTGAAAGGAAATGGCAATAAAGTTTCAATCCTTAAAACAAAAAAAAAAT

Appendix 2

Sequences for Trinity contigs listed in Table 6

```
>TRINITY_DN54168_c0_g2 [organism=Lotus filicaulis] CHS
ATGGTGAGTGTAGCTGAGATTCGTAAGGCTCAGAGGGCTGAAGGCCAGCAACC
ATCCTCGCTATTGGCACTGCTAATCCACCAAAGTGTGTCGACCAAAGCACTTATCCTGAT
TTTTACTTCAGAATCACTAACAGTGAACACATGACTGAGCTCAAGGAGAAGTTTCAGCGC
ATGTGTGACAAGTCTATGATCAAGAAGAGATATATGCACTTAACTGAAGATCTTTTGAAA
GAGAACCCTAACATGTGCGCTTA
```

```
>TRINITY_DN54168_c0_g1 [organism=Lotus filicaulis] CHS
CATCTCCTTAAAGATGTTCTGGGATTGTTTCAAAGAACATTGATAAAGCACTAGTT
GAGGCTTTCCAACCATTGAACATATCTGACTACAACTCAATTTTTTGGATTGCACACCCA
GGTGGCCCAGCAATTCTTGACCAAGTTGAGCAGAAGTTGAGCTTGAAACCTGAAAAGATG
AGGGCCACTAGGGAAGTGCTAAGTGAATATGGTAACATGTCAAGTGCATGCGTCCTATTC
ATCTTAGACGAAATGAGAAAGAAATCAGCTCAAGATGGACTCAAACCACCGGTGAAGGA
CTTGAATGGGGTGTGTTGTTTGGTTTTGGACCTGGACTTACCATTGAAACAGTTGTTTTG
CGTAGTGTGGCTATTTAA
```

```
>TRINITY_DN52652_c1_g1 [organism=Lotus filicaulis] CHS
ATGGTCACCGTCGAGGAGATCCGCAACGCACAGCGTTCCCATGGCCCCGCCACCATCCTCGCCTTCGG
CACCGCCACCCCTTCCCACTGTGTCTATGCAAGCTGATTACCCTGACTACTACTTCCGCATCACCAACA
GTGAACACATGACTGACCTCAAAGAAAAATTCAAGCGCATGTGTGAAAAGTCCATGATCAGGAAGCGC
TACATGCACCTGACGGAGGAGATTCTGAAGGAGAATCCAAACATGTGCGCGTACATGGCGCCGTCGCT
GGATGCTCGTCAGGACTTGGTGGTGGTGGAGGTGCCGAAGCTTGGTAAAGACGCGGCGGCGAAGGCGA
TCAAGGAGTGGGGTCAACCCAAGTCAAAGATCACTCACCTCGTTTTCTGCACCACCTCCGGCGTCGAC
ATGCCCCGGCGCCGATTACCAGCTCACCAAGCTTCTCGGCCTCAAACCCTCCGTGAAACGCCTCATGAT
GTACCAGCAAGGCTGCTTCGCCGGCGGCACGGTGCTCCGGTTGGCCAAGGATCTCGCCGAGAACAACA
AAGGCGCAAGAGTTCTCGTCGTCTGCTCTGAAATCACCGCCGTCACGTTCCGCGGCCCCCTCCGACACT
CACCTCGACTCGCTCGTCGGGCAAGCCCTCTTCGGCGACGGCGCCGCGCGATGATTGTGGGATCCGA
CCCTGACCTCGCCGTGGAGCGTCCGATTTTCAATTGGTCTCCGCCGCGCAGACGATTCTCCCTGATT
CCGACGGCGCAATCGACGGCCACCTCCGTGAGGTGGGGCTGACATTCCACCTTCTGAAGGACGTTCCG
GGGATCATCTCGAAGCACATTGAGAAGAGCTTGACGGAGGCGTTTGCGCCGATTGGGGTTTCTGATTG
GAACTCGCTGTTCTGGATCGCGCATCCGGGTGGACGGCGATTTTGGACCAGGTTGAAGTGAAACTCC
GGTTGAAAGAGGAGAACTCCGGTCGACCCGGCACGTGCTGAGTGAGTACGGAAACATGTCAAGCGCG
```

TGCGTGCTGTTTTATCTTGGATGAGGTGAGGAGGAGGTCTGAAGGAGGAAGGGAAGGAAACCACCGGAGA
 AGGGTTGGAGTGGGGGGTGCTATTCGGGTTCGGGCCGGGTCTCACCGTCGAGACCGTCGTGTTGCACA
 GTGTTCCC

>TRINITY_DN52780_c0_g1 [organism=Lotus filicaulis] F3'H

ATGTTCCAATGGATGATCATTGGCTTCGCC

ACCATCACAGTCATCATCTTCATACACCGTGTCGTGAAGTTCGCAACACGTCCCTCACTC
 CCACTCCCTCCGGGTCCAAAACCATGGCCCATAATAGGAAACCTCCCCACATGGGTCCG
 GTGCCGCACCACTCACTCGCAGCCTTAGCCCCAACACACGGCCCGCTTATGCACCTTAAA
 TTGGGCTTTGTGGATGTCGTGGTGGCGGCATCAGCCACCGTGGCGGAGCAGTTTTTAAAA
 GTTCATGATGCGAATTTTAGTAGCCGGCCCCCAACGCCGGAGCTAAGTACATAGCTTAT
 AACTATCAGGATCTCGTGTGTTGCTCCCTACGGCGCACGGTGGCGATATCTCCGAAAAATC
 ACCTCTCTTCATCTCTTCTCTGGCAAGGCCTTGGATAATTTTAAGCACTTGCGTCAGGAA
 GAAGTATCAAGATTGACACGCAACATATCAAAGTCCAACCTCAAAAGTTGTCAACTTGGGA
 CAACTACTAAATGTGTGCACTACCAACGCATTATCAAGGGTAATGATTGGAAGGAGAGTG
 TTCAACGACGGCGACGGTGGGTGTGATCCTAGAGCTGATGAGTTTAAGGCCATGGTGGTG
 GAGTTAATGGTGTAGCGGGTGTTCACATAGGTGACTTCATTCTCCTTGGAGTGG
 TTAGACCTTCAAGGGGTGCAAGCTAAGATGAAGAATTTGCACAATAGGTTTGATGAATTT
 TTAACCAGCATAATTGAGGAACACGATACATCTTCAAAGAGTGAGAATCATAAAGATTTG
 TTAAGCACATTATTGTCGCTTAAAGAGGTTGCCGATGATGATGGAAATAAACTCAATGAT
 ATCGAGATCAAAGCACTACTTCTGAACATGTTACAGCTGGGACAGACACGTCAGCAAGC
 ACAACAGAATGGGCCATTGCAGAACTCATCCGCAACCCAAAAATTCTGGCCCAAGTCCAA
 CAAGAATTGGACACAGTCGTGGGCCGAGACAGAAACGTGAGAGAAGATGACTTGCCTCAT
 CTCCCATACTTACAAGCAGTGGTGAAAGAAACCTTTTCGTCTTCATCCATCAACCCCTCTT
 TCACTGCCACGTGTCGCATCTGAGAGCTGTGAGGTACTTGGCTACCACATCCCAAAGGGT
 TCCACACTCTTGGTGAACGTGTGGGCCATAGCCCGTGACCCTAAAGAATGGGCTGACCCA
 TTGGAGTTCAAGCCTGAAAGGTTCTTGACAGGTGGTGGTGATAAGGTTTCATGTTGATGTT
 AAGGGCAATGACTTTGAGGTCATACCCTTTGGTGCTGGACGTAGGATATGTGCTGGGATG
 AGCCTTGGGCTTCGTATGGTTCAGCTGTAACTGCAACTCTGGCTCATTTCGTTTAACTGG
 GAGCTTGAAAATGGGCTTAACCATGAAAAGCTCAACATGGATGAAGCATATGGGTTAACA
 TTGCAACGGGCTGTGCCTTTGTCAGTGTACTCTAGGCCAAGGCTCTCGCCACATGTGTA

>TRINITY_DN52390_c0_g1 [organism=Lotus filicaulis] F3H

CTCACCACTCTAGCCCAACAAAACACCCTTGAGTCCAGCTTCGTCCGTGACGAAGACGAGCGCCAAA
 GGTGCTTACAACAATTTAGCAACGAGATCCCGGTCATTTCCCTCGCCGGAATCGACGAGGTGATG
 GCCGGAGATCCGAGATATGTAACAAGATTGTTGAGGCTTGTGAGAACTGGGGAATCTTCCAGGTTGTC
 AACCACGGCGTCGACACTGAATTGGTTTCCCATATGACCACTCTTGCCAAAGAGTTCTTTGCTCTTCC
 CCCTGAAGAGAAGCTTCGCTTTGACATGTCCGGTGGCAAAAAGGGTGGTTTCATCGTCTCCAGCCACC

TCCAAGGAGAATCAGTGCAGGATTGGAGAGAGATAGTGACATACTTTTCATACCCAATAAGGCAGAGG
 GACTACTCACGGTGGCCAGACACCCCGCGGGGTGGAAGGCGGTGACGGAGGAGTACAGCGAGAAGCT
 GATGGGTTTAGCATGCAAGCTATTGGAGGTGTTATCAGAGGCAATGGGGTTAGAGAAAGAGGCTCTAA
 CAAAGGCATGTGTTGATATGGACCAGAAAGTTGTGGTGAATTATTACCCAAAATGCCCTCAGCCTGAC
 CTCACACTTGGACTGAAGAGACACACTGATCCTGGCACCATTACTCTGCTGCTTCAGGATCAGGTGGG
 TGGGCTTCAAGCAACTAGGGATAATGGGAAGACATGGATCACTGTGCAGCCTGTTGAAGGAGCTTTTCG
 TTGTCAATCTTGGTGACCATGGTCATTATCTGAGCAATGGGAGGTTCAAGAACGCTGATCACCAAGCA
 GTGGTGAACCTCAAACCTCAAGCCGTTTATCGATCGCCACATTCCAGAACCCAGCACCAGATGCAACTGT
 GTACCCATTGAAGGTTTCGAGAGGGAGAGAAGTCGGTCCTAGATGAACCAATCACATTTCGCTGAAATGT
 ACAGGAGGAAGATGAGCAAGGACATTGAGCTTGCTCGGATGAAGAAGCTGGCTAAGGAGAAGAACTG
 CAGGATTTGGAGAAGGCCAAACTTGAGCCTAAGCCCATGAATGAGATCTTTGCTTA

>TRINITY_DN52789_c0_g1 [organism=Lotus filicaulis] DFR
 GAAACGGTTTTGCGTTACAGGGGCTGCTGGTTTTCATCGGGTCATGGCTCGT
 CATGAGGCTCATCGAGCGTGGCTACACGGTTCGAGCCACCGTACGTGACCCAGCAAACAT
 GAAGAAGGTGAAGCATTTGCTGGAAGTCCAAATGCAAAGACCAAGCTGACTCTGTGGAA
 AGCTGATCTTGCTGAAGAGGGAAGCTTTGATGAAGCCATCAAAGGGTGCACTGGAGTTTT
 CCATGTGGCCACACCCATGGATTTTGAAGTCCAAGGACCCTGAGAATGAAGTAATAAAGCC
 AACAATAAATGGATTGTTAGACATCTTGAAAGCATGTGAGAAGGCCAAAAGTGTAAAGAAG
 GTTGGTATTACATCCTCAGCAGGGACTGTGGATGTTACAGAGCAACAAAAGCCTGTTAT
 TGATGAAACATGCTGGAGTGACATTGAATTCTGCCTCAGAGTCAAGATGACTGGCTGGAT
 GTATTTTCGTTTCCAAGACACGAGCAGAACAAGAAGCGTGGAAATATGCAAAAGAGCACAA
 CATAGACTTCGTCTCAGTCATTCCACCTCTTGTGGTTGGCCCATTTCTCATGCCAACAAT
 GCCACCTAGCCTAATCACTGCTCTCTCACTCATCACAGGAAATGAGGCCCATTAATCAAT
 CATAAAGCAAGGCCAATATGTCCACTTAGATGACCTTTGTCTAGCTCATATATTCTTGTT
 TGAGAACCCAAAAGCCCAAGGGAGATACATGTGCTCTGCATATGAGGCAACCATTTCATGA
 AGTTGCAAGAATGATCAACAAGAAGTACCCAGAGTTCAATGTCCCCACAAAGTTCAAGGA
 TATTCGGATGAATTGGACATCATTAATTTTCTTCAAAGAAGATCACAGACTTGGGGTT
 CAAATTTAAGTACAGCTTAGAGGACATGTACACAGGAGCCATTGAAACCTGCAGAGAAAA
 GGGGCTTCTTCCTAAAACCTTCTGAAACTCCAGTTACTAATGGCACAACCTCAGAAATAA

>TRINITY_DN53094_c0_g1 [organism=Lotus filicaulis] ANS
 ATGGCGCCTACTGTTGTTGAGAGGGTTGAAAGCTTATCAACCAGTGGGATACAATCTATCCCAAAGGA
 ATATGTGAGGCCGAAAGAGGAGCTGGAAAACATAGGCAACGTTTTTGGAGGAGAAAAGAAGGATGGGC
 CTCAAGTTCCAACAATTGACTTGAAAGAGATAGACTCTCCCGACGAGTTTGTTCGGGCCAATTGCAGG
 GAGAAGCTGAGGAAGGCGGCGGAGGAGTGGGGGGTGATGCACCTGGTGAACCATGGGATCCCGGACGA
 GCTTCTGAACCAATTGAAGACAGCTGGAGCCGAGTTTTTCTCTTTGCCTGTGAGGAGAAGGAGAAGT
 ACGCGAATGATCAGACGACCGGAAATGTGCAAGGGTACGGGAGCAAGCTGGCTAACAATGCCAGTGGC

CAGCTTGAATGGGAAGATTACTTCTTCCATCTTATCTTTCCAGAAGACAAACGTGACCTCTCTATCTG
 GCCCCAAGACACCATCATATTATACAGAGGTTACCAGCGACTATGCAAGGCGACTGAGAGTCCTTGCGA
 GCAAGGTGCTGGAGGTGCTATCTCTGGAATTGGGTCTGGAAGAAGGAAGGTTAGAGAAGGAAGTTGGT
 GGAATGGAAGAGCTTCTACTTCAGATGAAGATCAACTACTACCCAAAATGCCCCAGCCAGAGCTTGC
 TCTAGGAGTTGAAGCCACACCGATGTAAGCTCACTCACTTTCTCGTCCACAACATGGTCCCAGGCC
 TGCAGCTTTTCTACCAAGATAAATGGATCACAGCAAAATGTGTCCCTGATTCAATTCTCATGCACATT
 GGTGACACCATTGAGATCCTGAGCAATGGCAAGTTCAAGAGCATTCTTCACAGGGGGTTGGTGAACAA
 GGAAAAAGTTAGAATATCTTGGGCTGTGTTCTGTGAACCCCTAAGGAGAAAATCATCCTCAAGCCAC
 TTCCTGAGCTTGTGACTGAGACCGAACGGCGCTGTTTCCGCCGCGCACTTTTGCTCAGCATCTTCAT
 CACAAATTGTTTCAGGAAGGACCAGGAAGCTGCTGCTCAATCCAAATGA

>TRINITY_DN39800_c0_g1 [organism=Lotus filicaulis] FLS

TGGAG

GTGCTAAGGGTGCAAACCGTGGCAACCCAGTCAAAGACGCTTCCATCCCCTCGATGTTT
 GTGAGGCCGAAGACAGAGCAGCCTGGCCTCACGACCGTTTCGCGGGGTGGAAGTGGAGGTG
 CCGATCATTGATTTGAGCGGCTCCGATGAAGAAAAGGTGCTGCGTGAGATCGTGGAGGCG
 AGTAAGGAGTGGGGGATGTTTCAAGTTGTGAATCATGAGATACCTGTTGAAGTTATAGCC
 AAATTGCAGAGTGTGGGGAACATTTCTTTGAGTTGCCACAAGGGGAAAAAGAGGTGTGT
 GGTAAAGATTGATGGGTCTGATTCTGTGGAAGGTTATGGGACTAAACTTCAGAAGGAGGTG
 AATGGGAAGAAAGGTTGGGTGGATCATTTGTTTCATATTATTTGGCCAACTTCTGCCATT
 AACTACCGATTCTGGCCTAAGAATCCTCCTTCTTACAGGGAGGTGAATGAGGAATATGGC
 AAGTACCTGCGTAGTGTGGGGGACAACTTTTCAAAGCTTGTCAATTGGGTTGGGCCTT
 GAAGAAAATGAGCTAAAAGAAGCTGCAGGTGGAGATGACATGATTACCTATTAAAGATC
 AACTACTACCCACCATGTCCATGTCCTGAACTGGTTCTGGGTGTGCCACCACACACAGAC
 ATGTCATTTCATGACCATTCTAGTTCCCAACGAGGTGCAGGGCCTCCAAGCCTTCAGGGAT
 GGTCACTGGTACGATGTAAAGTATGTCCCCAATGCCCTCGTCATTACATTGGAGACCAA
 ATGGAGATAGTGAGCAATGGGAAATATAAGGCTGTCTGCACAGAACAACAGTGAACAAA
 GAGGAGACAAGAATGTCATGGCCAGTGTTTCATAGAGCCCCAAGGACACCATGAAGTTGGT
 CCTCACTCCAAGTTGGTTAACCACGACAATCCACCAAAATACAAGACCAAGAAATATAAG
 GACTATGCCTACTGCAAGCTTAATAAGATCCCTCAGTAA

Appendix 3

Supplementary Tables S1-S54

Table S1. Tophat Results: Significantly enriched GO terms corresponding to the yellow stage of a *L. filicaulis* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0015979photosynthesis	42	1.70E-05
GO:0009628response to abiotic stimulus	179	1.70E-05
GO:0042221response to chemical stimulus	243	2.50E-05
GO:0050896response to stimulus	376	0.00014
GO:0010033response to organic substance	159	0.00014
GO:0009725response to hormone stimulus	124	0.00014
GO:0044281small molecule metabolic process	207	0.00018
GO:0044283small molecule biosynthetic process	107	0.00024
GO:0009719response to endogenous stimulus	131	0.00024
GO:0019684photosynthesis, light reaction	26	0.00024
GO:0046394carboxylic acid biosynthetic process	69	0.0024
GO:0016053organic acid biosynthetic process	69	0.0024
GO:0006091generation of precursor metabolites and energy	44	0.0037
GO:0009416response to light stimulus	81	0.0041
GO:0008152metabolic process	796	0.0041
GO:0005976polysaccharide metabolic process	40	0.0041
GO:0009314response to radiation	81	0.0044
GO:0010218response to far red light	15	0.0044
GO:0043436oxoacid metabolic process	106	0.0044
GO:0006950response to stress	233	0.0044
GO:0010608posttranscriptional regulation of gene expression	25	0.0044
GO:0006082organic acid metabolic process	106	0.0044
GO:0019752carboxylic acid metabolic process	106	0.0044
GO:0009409response to cold	52	0.0047
GO:0009607response to biotic stimulus	102	0.0056
GO:0042180cellular ketone metabolic process	108	0.0056
GO:0044264cellular polysaccharide metabolic process	35	0.0064
GO:0009637response to blue light	16	0.008
GO:0006519cellular amino acid and derivative metabolic process	87	0.008

Enriched GO Term: Description	Number of associated genes	FDR
GO:0032501multicellular organismal process	207	0.0086
GO:0016072rRNA metabolic process	16	0.011
GO:0009639response to red or far red light	29	0.011
GO:0051707response to other organism	96	0.011
GO:0008610lipid biosynthetic process	58	0.011
GO:0006073cellular glucan metabolic process	29	0.011
GO:0019748secondary metabolic process	58	0.011
GO:0016441posttranscriptional gene silencing	17	0.012
GO:0009733response to auxin stimulus	44	0.012
GO:0033692cellular polysaccharide biosynthetic process	24	0.012
GO:0044042glucan metabolic process	29	0.012
GO:0051704multi-organism process	115	0.012
GO:0043648dicarboxylic acid metabolic process	22	0.014
GO:0007275multicellular organismal development	197	0.014
GO:0000271polysaccharide biosynthetic process	24	0.014
GO:0055114oxidation reduction	133	0.015
GO:0005975carbohydrate metabolic process	114	0.016
GO:0008652cellular amino acid biosynthetic process	33	0.018
GO:0009791post-embryonic development	119	0.02
GO:0070887cellular response to chemical stimulus	80	0.023
GO:0006720isoprenoid metabolic process	25	0.024
GO:0009617response to bacterium	55	0.024
GO:0006520cellular amino acid metabolic process	55	0.024
GO:0009250glucan biosynthetic process	18	0.024
GO:0006364rRNA processing	14	0.024
GO:0042214terpene metabolic process	13	0.026
GO:0009415response to water	41	0.028
GO:0016246RNA interference	12	0.028
GO:0009765photosynthesis, light harvesting	10	0.029
GO:0009414response to water deprivation	40	0.031
GO:0009308amine metabolic process	64	0.037
GO:0032502developmental process	204	0.038
GO:0006629lipid metabolic process	91	0.044
GO:0006721terpenoid metabolic process	20	0.044
GO:0030422production of siRNA involved in RNA interference	11	0.046
GO:0009309amine biosynthetic process	34	0.048

Enriched GO Term: Description	Number of associated genes	FDR
GO:0044106cellular amine metabolic process	57	0.048
GO:0003824catalytic activity	688	0.002
GO:0009507chloroplast	311	3.60E-35
GO:0044435plastid part	232	3.60E-35
GO:0044434chloroplast part	228	1.00E-34
GO:0009536plastid	319	1.90E-33
GO:0009532plastid stroma	145	7.50E-26
GO:0009570chloroplast stroma	139	2.80E-24
GO:0044464cell part	910	2.10E-19
GO:0043227membrane-bounded organelle	630	2.10E-19
GO:0044444cytoplasmic part	570	2.10E-19
GO:0005623cell	910	2.10E-19
GO:0043231intracellular membrane-bounded organelle	628	2.10E-19
GO:0009526plastid envelope	123	6.00E-19
GO:0043226organelle	656	1.00E-18
GO:0043229intracellular organelle	655	1.20E-18
GO:0005737cytoplasm	598	1.80E-18
GO:0044422organelle part	370	4.00E-18
GO:0044446intracellular organelle part	368	4.50E-18
GO:0009941chloroplast envelope	118	5.70E-18
GO:0009579thylakoid	106	6.70E-17
GO:0031976plastid thylakoid	92	1.30E-16
GO:0009534chloroplast thylakoid	92	1.30E-16
GO:0031984organelle subcompartment	92	1.50E-16
GO:0044436thylakoid part	86	2.70E-16
GO:0044424intracellular part	720	3.50E-16
GO:0055035plastid thylakoid membrane	78	2.10E-15
GO:0042651thylakoid membrane	80	2.80E-15
GO:0005622intracellular	727	3.50E-15
GO:0009535chloroplast thylakoid membrane	77	4.20E-15
GO:0031975envelope	136	5.50E-15
GO:0031967organelle envelope	134	1.90E-14
GO:0034357photosynthetic membrane	80	3.50E-14
GO:0048046apoplast	71	3.00E-08
GO:0005576extracellular region	92	1.30E-07
GO:0010287plastoglobule	20	6.20E-06

Enriched GO Term: Description	Number of associated genes	FDR
GO:0030312external encapsulating structure	87	6.30E-06
GO:0016020membrane	441	8.20E-06
GO:0005618cell wall	85	1.40E-05
GO:0031090organelle membrane	148	2.70E-05
GO:0031977thylakoid lumen	20	4.60E-05
GO:0005886plasma membrane	234	7.00E-05
GO:0043232intracellular non-membrane-bounded organelle	121	7.10E-05
GO:0043228non-membrane-bounded organelle	121	7.10E-05
GO:0010319stromule	15	0.00015
GO:0009505plant-type cell wall	40	0.00029
GO:0031974membrane-enclosed lumen	80	0.00042
GO:0070013intracellular organelle lumen	79	0.00042
GO:0043233organelle lumen	79	0.00044
GO:0005730nucleolus	48	0.0026
GO:0009295nucleoid	10	0.01
GO:0044459plasma membrane part	104	0.012
GO:0031978plastid thylakoid lumen	11	0.015
GO:0009543chloroplast thylakoid lumen	11	0.015
GO:0009521photosystem	13	0.016
GO:0042646plastid nucleoid	8	0.025
GO:0009506plasmodesma	91	0.036
GO:0055044sympplast	91	0.036
GO:0030054cell junction	91	0.038
GO:0005911cell-cell junction	91	0.038
GO:0000229cytoplasmic chromosome	6	0.043
GO:0009508plastid chromosome	6	0.043
GO:0046658anchored to plasma membrane	12	0.044

Table S2. Tophat Results: Significantly enriched GO terms corresponding to the red stage of a *L. filicaulis* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0050896response to stimulus	266	0.00056
GO:0055085transmembrane transport	70	0.00056
GO:0005975carbohydrate metabolic process	92	0.0012
GO:0051179localization	152	0.0014
GO:0009605response to external stimulus	56	0.0015
GO:0042221response to chemical stimulus	164	0.0015
GO:0009991response to extracellular stimulus	27	0.0016
GO:0051234establishment of localization	145	0.0016
GO:0006810transport	144	0.0016
GO:0009932cell tip growth	23	0.0017
GO:0071704organic substance metabolic process	139	0.0017
GO:0009813flavonoid biosynthetic process	19	0.0029
GO:0035295tube development	24	0.0029
GO:0048868pollen tube development	24	0.0029
GO:0031667response to nutrient levels	23	0.0029
GO:0031668cellular response to extracellular stimulus	25	0.0029
GO:0071496cellular response to external stimulus	25	0.0031
GO:0006575cellular amino acid derivative metabolic process	41	0.0048
GO:0042398cellular amino acid derivative biosynthetic process	32	0.0049
GO:0009725response to hormone stimulus	82	0.0051
GO:0010033response to organic substance	105	0.0059
GO:0009812flavonoid metabolic process	20	0.0074
GO:0006811ion transport	53	0.0074
GO:0009719response to endogenous stimulus	87	0.0074
GO:0031669cellular response to nutrient levels	21	0.0077
GO:0009699phenylpropanoid biosynthetic process	22	0.01
GO:0009860pollen tube growth	18	0.011
GO:0042594response to starvation	19	0.011
GO:0009826unidimensional cell growth	28	0.011

Enriched GO Term: Description	Number of associated genes	FDR
GO:0060560developmental growth involved in morphogenesis	28	0.011
GO:0016036cellular response to phosphate starvation	14	0.011
GO:0009737response to abscisic acid stimulus	43	0.014
GO:0032989cellular component morphogenesis	36	0.016
GO:0048588developmental cell growth	21	0.016
GO:0006725cellular aromatic compound metabolic process	93	0.016
GO:0015698inorganic anion transport	14	0.016
GO:0044282small molecule catabolic process	37	0.016
GO:0006576cellular biogenic amine metabolic process	13	0.017
GO:0044262cellular carbohydrate metabolic process	59	0.018
GO:0009267cellular response to starvation	18	0.019
GO:0019438aromatic compound biosynthetic process	70	0.019
GO:0051704multi-organism process	81	0.022
GO:0048610reproductive cellular process	18	0.026
GO:0009856pollination	25	0.03
GO:0016137glycoside metabolic process	16	0.033
GO:0009718anthocyanin biosynthetic process	11	0.034
GO:0008361regulation of cell size	34	0.037
GO:0010200response to chitin	17	0.037
GO:0032502developmental process	144	0.038
GO:0006519cellular amino acid and derivative metabolic process	59	0.038
GO:0019748secondary metabolic process	40	0.038
GO:0000902cell morphogenesis	31	0.043
GO:0032501multicellular organismal process	140	0.044
GO:0006595polyamine metabolic process	7	0.046
GO:0044248cellular catabolic process	61	0.046
GO:0022892substrate-specific transporter activity	72	0.00021
GO:0005215transporter activity	90	0.00021
GO:0022857transmembrane transporter activity	77	0.00021
GO:0015291secondary active transmembrane transporter activity	30	0.0005

Enriched GO Term: Description	Number of associated genes	FDR
GO:0022891substrate-specific transmembrane transporter activity	65	0.00072
GO:0022804active transmembrane transporter activity	45	0.00072
GO:0016791phosphatase activity	27	0.0017
GO:0016798hydrolase activity, acting on glycosyl bonds	44	0.0017
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	42	0.0017
GO:0015075ion transmembrane transporter activity	50	0.0017
GO:0003779actin binding	13	0.0047
GO:0015103inorganic anion transmembrane transporter activity	15	0.0047
GO:0015297antiporter activity	19	0.0053
GO:0042578phosphoric ester hydrolase activity	29	0.0058
GO:0008509anion transmembrane transporter activity	18	0.024
GO:0016020membrane	305	0.0014
GO:0005886plasma membrane	164	0.0033

Table S3. Yellow Stage TopHat Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=1). The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0018.1300.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase	143	11	2.87090
chr5.CM1077.590.r2.m	Glyma.03G187000.1	UDP-glucosyl transferase 73C	44	1	4.52042
chr1.CM0122.1600.r2.m	Glyma.03G222100.1	Nitrogen regulatory protein P-II	134	18	2.07132
chr1.CM0012.990.r2.m	Glyma.04G254200.1	B3 DNA binding domain // Auxin response factor	119	12	2.48162
chr1.CM0064.910.r2.m	Glyma.05G076300.1	NA	1067	169	1.83858
chr3.LjT40P18.80.r2.m	Glyma.07G021600.1	Shikimate O-hydroxycinnamoyltransferase	350	25	2.98347
chr3.CM0452.240.r2.d	Glyma.07G021600.1	Shikimate O-hydroxycinnamoyltransferase	8669	1189	2.04670
chr3.LjT10E18.60.r2.m	Glyma.10G210600.1	B3 DNA binding domain // Auxin response factor	230	26	2.32171
chr3.CM0136.10.r2.m	Glyma.11G145500.1	AUX/IAA family // B3 DNA binding domain // Auxin response factor	28	2	2.93242
chr4.CM0432.2880.r2.m	Glyma.13G173300.1	O-METHYLTRANSFERASE	2443	223	2.63371
chr3.CM0106.330.r2.m	Glyma.13G344700.1	B-box zinc finger	712	37	3.44375
chr1.LjT35H06.80.r2.m	NA	NA	51	0	8.32466

Table S4. Red Stage TopHat Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR = 0.0029) . The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0608.560.r2.m	Glyma.01G048200	RIBONUCLEASE T2	872	8349	-4.07828
chr2.CM0124.30.r2.m	Glyma.01G088300	STEROL REGULATORY ELEMENT-BINDING PROTEIN	8	26	-2.49628
chr3.LjB14006.120.r2.a	Glyma.01G166200	naringenin 3-dioxygenase	228	542	-2.06782
chr4.CM0119.240.r2.m	Glyma.01G166200	naringenin 3-dioxygenase	1043	3498	-2.56494
chr1.CM0284.250.r2.m	Glyma.01G228700	Naringenin-chalcone synthase	7	144	-5.15029
chr2.CM0018.760.r2.m	Glyma.01G228700	Naringenin-chalcone synthase	115	905	-3.79376
chr2.CM0018.730.r2.m	Glyma.01G228700	Naringenin-chalcone synthase	48	170	-2.63974
chr2.CM0018.1200.r2.m	Glyma.01G228700	Naringenin-chalcone synthase	3	61	-5.09247
chr5.CM0077.110.r2.m	Glyma.02G158700	bifunctional dihydroflavonol 4-reductase/flavanone 4- reductase	32	929	-5.67183
chr5.CM0077.210.r2.m	Glyma.02G158700	bifunctional dihydroflavonol 4-reductase/flavanone 4- reductase	150	10354	-6.92689
chr1.CM0105.1870.r2.m	Glyma.03G113400	Proton-exporting ATPase	304	1049	-2.60557
chr1.CM0133.560.r2.m	Glyma.04G052100	Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies	40	2793	-6.9393
chr1.LjT39K18.30.r2.m	Glyma.04G147700	HEME OXYGENASE	636	1935	-2.42427
chr2.CM0191.680.r2.m	Glyma.04G227700	Quercetin 3-O- methyltransferase	4	21	-3.16141
chr1.CM0410.430.r2.m	Glyma.08G247100	Transferase family	2403	7930	-2.54173
chr6.CM0013.1540.r2.a	Glyma.09G038900	MYB-LIKE DNA-BINDING PROTEIN MYB	7	125	-4.94628

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0021.2820.r2.m	Glyma.11G027700	Leucocyanidin oxygenase	1001	15407	-4.76318
	.1				
chr5.CM0077.120.r2.m	Glyma.14G072700	bifunctional dihydroflavonol	11	61	-3.27199
	.1	4-reductase/flavanone 4-			
		reductase			
chr5.CM0180.690.r2.m	Glyma.20G241700	chalcone isomerase	529	1254	-2.06419
	.1				

Table S5. Yellow Stage TopHat Results: Genes for enzymes annotated to GO:0009809 (Lignin biosynthetic process, FDR=0.24). The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr6.LjT35D18.30.r2.m	Glyma.03G223000.1	Methionine adenosyltransferase	17391	2734	1.849893 404
chr3.LjT40P18.80.r2.m	Glyma.07G021600.1	Shikimate O-hydroxycinnamoyltransferase	350	25	2.983475 586
chr3.CM0452.240.r2.d	Glyma.07G021600.1	Shikimate O-hydroxycinnamoyltransferase	8669	1189	2.046705 844
chr3.LjT45M09.50.r2.d	Glyma.07G023700.1	cinnamoyl-CoA reductase	1692	266	1.849549 234
chr4.CM0387.960.r2.m	Glyma.07G258700.1	beta-glucosidase	131	1	6.090954 877
chr6.CM0385.240.r2.d	Glyma.09G038500.1	CHITINASE	1204	8	6.398266 711
chr5.CM0200.3250.r2.m	Glyma.10G262400.1	cinnamyl-alcohol dehydrogenase	4676	705	1.910116 099

*Lignin biosynthetic process was not found to be enriched at the red stage of floral colour change in *L. filicaulis*, given the TopHat results.

Table S6. STAR Results: Significantly enriched GO terms corresponding to the yellow stage of a *L. filicaulis* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0009628response to abiotic stimulus	184	1.00E-06
GO:0050896response to stimulus	379	1.80E-05
GO:0042221response to chemical stimulus	241	1.80E-05
GO:0044281small molecule metabolic process	206	0.00017
GO:0009416response to light stimulus	87	0.00024
GO:0015979photosynthesis	37	0.00024
GO:0009314response to radiation	87	0.00032
GO:0010033response to organic substance	153	0.00059
GO:0044283small molecule biosynthetic process	103	0.00086
GO:0006950response to stress	237	0.00086
GO:0005976polysaccharide metabolic process	42	0.00089
GO:0008610lipid biosynthetic process	63	0.00097
GO:0009725response to hormone stimulus	117	0.00097
GO:0008152metabolic process	789	0.0018
GO:0044264cellular polysaccharide metabolic process	37	0.0018
GO:0009607response to biotic stimulus	104	0.0018
GO:0046394carboxylic acid biosynthetic process	68	0.0018
GO:0016053organic acid biosynthetic process	68	0.0018
GO:0009719response to endogenous stimulus	123	0.0021
GO:0019684photosynthesis, light reaction	23	0.0023
GO:0009637response to blue light	17	0.0027
GO:0033692cellular polysaccharide biosynthetic process	26	0.0027
GO:0000271polysaccharide biosynthetic process	26	0.0028
GO:0051707response to other organism	98	0.0028
GO:0010218response to far red light	15	0.0028
GO:0009409response to cold	52	0.0028
GO:0043436oxoacid metabolic process	105	0.0028
GO:0032501multicellular organismal process	208	0.0028
GO:0010608posttranscriptional regulation of gene expression	25	0.0028
GO:0006082organic acid metabolic process	105	0.0028
GO:0019752carboxylic acid metabolic process	105	0.0028
GO:0007275multicellular organismal development	200	0.0033
GO:0042214terpene metabolic process	15	0.0035
GO:0005975carbohydrate metabolic process	117	0.0037
GO:0019748secondary metabolic process	59	0.0045

Enriched GO Term: Description	Number of associated genes	FDR
GO:0009791post-embryonic development	122	0.0047
GO:0006720isoprenoid metabolic process	27	0.0051
GO:0042180cellular ketone metabolic process	106	0.0051
GO:0055114oxidation reduction	134	0.0067
GO:0006073cellular glucan metabolic process	29	0.0075
GO:0032502developmental process	208	0.0085
GO:0044042glucan metabolic process	29	0.0095
GO:0071554cell wall organization or biogenesis	45	0.011
GO:0006091generation of precursor metabolites and energy	40	0.012
GO:0051704multi-organism process	113	0.014
GO:0009639response to red or far red light	28	0.014
GO:0044255cellular lipid metabolic process	70	0.015
GO:0006629lipid metabolic process	93	0.016
GO:0071555cell wall organization	32	0.017
GO:0006519cellular amino acid and derivative metabolic process	83	0.017
GO:0048856anatomical structure development	168	0.018
GO:0044262cellular carbohydrate metabolic process	80	0.018
GO:0006721terpenoid metabolic process	21	0.018
GO:0009266response to temperature stimulus	62	0.019
GO:0009250glucan biosynthetic process	18	0.019
GO:0016441posttranscriptional gene silencing	16	0.019
GO:0005982starch metabolic process	14	0.025
GO:0008299isoprenoid biosynthetic process	22	0.025
GO:0009617response to bacterium	54	0.025
GO:0048731system development	105	0.027
GO:0048513organ development	105	0.027
GO:0034637cellular carbohydrate biosynthetic process	37	0.03
GO:0016051carbohydrate biosynthetic process	38	0.031
GO:0071669plant-type cell wall organization or biogenesis	24	0.031
GO:0009414response to water deprivation	39	0.038
GO:0010154fruit development	53	0.038
GO:0006633fatty acid biosynthetic process	31	0.039
GO:0009658chloroplast organization	20	0.04
GO:0016072rRNA metabolic process	14	0.041
GO:0010035response to inorganic substance	75	0.044
GO:0032787monocarboxylic acid metabolic process	56	0.048

Enriched GO Term: Description	Number of associated genes	FDR
GO:0003824catalytic activity	689	0.00013
GO:0009507chloroplast	312	1.80E-36
GO:0044435plastid part	232	2.90E-36
GO:0044434chloroplast part	228	1.10E-35
GO:0009536plastid	321	2.20E-35
GO:0009532plastid stroma	148	7.90E-28
GO:0009570chloroplast stroma	142	3.30E-26
GO:0044464cell part	920	2.20E-24
GO:0005623cell	920	2.20E-24
GO:0043227membrane-bounded organelle	639	3.00E-23
GO:0044444cytoplasmic part	579	3.00E-23
GO:0043231intracellular membrane-bounded organelle	637	3.50E-23
GO:0005737cytoplasm	608	3.10E-22
GO:0043226organelle	662	7.90E-22
GO:0043229intracellular organelle	661	9.00E-22
GO:0044424intracellular part	732	2.60E-20
GO:0044422organelle part	371	1.90E-19
GO:0009579thylakoid	111	2.10E-19
GO:0044446intracellular organelle part	369	2.10E-19
GO:0009526plastid envelope	122	3.10E-19
GO:0005622intracellular	738	4.60E-19
GO:0031976plastid thylakoid	96	8.40E-19
GO:0009534chloroplast thylakoid	96	8.40E-19
GO:0031984organelle subcompartment	96	1.00E-18
GO:0009941chloroplast envelope	117	3.60E-18
GO:0044436thylakoid part	89	4.60E-18
GO:0055035plastid thylakoid membrane	82	1.30E-17
GO:0042651thylakoid membrane	84	1.80E-17
GO:0009535chloroplast thylakoid membrane	81	2.80E-17
GO:0034357photosynthetic membrane	84	3.10E-16
GO:0031975envelope	136	1.70E-15
GO:0031967organelle envelope	134	6.10E-15
GO:0048046apoplast	75	6.60E-10
GO:0005576extracellular region	97	2.20E-09
GO:0016020membrane	443	8.30E-07
GO:0030312external encapsulating structure	88	1.90E-06

Enriched GO Term: Description	Number of associated genes	FDR
GO:0031090organelle membrane	152	2.10E-06
GO:0031977thylakoid lumen	22	3.50E-06
GO:0005618cell wall	86	4.40E-06
GO:0010287plastoglobule	19	1.60E-05
GO:0010319stromule	16	3.60E-05
GO:0005886plasma membrane	232	4.80E-05
GO:0009505plant-type cell wall	41	0.0001
GO:0043232intracellular non-membrane-bounded organelle	118	0.00013
GO:0043228non-membrane-bounded organelle	118	0.00013
GO:0031974membrane-enclosed lumen	81	0.00016
GO:0070013intracellular organelle lumen	80	0.00016
GO:0043233organelle lumen	80	0.00016
GO:0031978plastid thylakoid lumen	13	0.0014
GO:0009543chloroplast thylakoid lumen	13	0.0014
GO:0005730nucleolus	46	0.0056
GO:0009295nucleoid	10	0.0089
GO:0044459plasma membrane part	103	0.01
GO:0042646plastid nucleoid	8	0.023
GO:0009506plasmodesma	90	0.035
GO:0055044symplast	90	0.035
GO:0030054cell junction	90	0.037
GO:0005911cell-cell junction	90	0.037
GO:0000229cytoplasmic chromosome	6	0.041
GO:0009508plastid chromosome	6	0.041

Table S7. STAR Results: Significantly enriched GO terms corresponding to the red stage of a *L. filicaulis* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0050896response to stimulus	278	3.30E-06
GO:0042221response to chemical stimulus	175	2.40E-05
GO:0009605response to external stimulus	60	0.00014
GO:0005975carbohydrate metabolic process	95	0.00014
GO:0009991response to extracellular stimulus	29	0.00025
GO:0009725response to hormone stimulus	89	0.00033
GO:0009813flavonoid biosynthetic process	21	0.00042
GO:0010033response to organic substance	111	0.00054
GO:0031668cellular response to extracellular stimulus	27	0.00054
GO:0051179localization	151	0.00054
GO:0009719response to endogenous stimulus	93	0.00054
GO:0071496cellular response to external stimulus	27	0.00054
GO:0071704organic substance metabolic process	141	0.00054
GO:0031667response to nutrient levels	24	0.0009
GO:0055085transmembrane transport	66	0.0009
GO:0051234establishment of localization	144	0.00093
GO:0042398cellular amino acid derivative biosynthetic process	34	0.00093
GO:0006575cellular amino acid derivative metabolic process	43	0.00094
GO:0006810transport	143	0.00094
GO:0009812flavonoid metabolic process	22	0.00099
GO:0009699phenylpropanoid biosynthetic process	24	0.0017
GO:0009932cell tip growth	22	0.0023
GO:0031669cellular response to nutrient levels	22	0.0029
GO:0006725cellular aromatic compound metabolic process	97	0.0035
GO:0042594response to starvation	20	0.0041
GO:0035295tube development	23	0.0041
GO:0048868pollen tube development	23	0.0041
GO:0009718anthocyanin biosynthetic process	13	0.0041
GO:0044282small molecule catabolic process	39	0.0048

Enriched GO Term: Description	Number of associated genes	FDR
GO:0019438aromatic compound biosynthetic process	73	0.0051
GO:0009737response to abscisic acid stimulus	44	0.0067
GO:0044262cellular carbohydrate metabolic process	61	0.0067
GO:0006519cellular amino acid and derivative metabolic process	63	0.0067
GO:0032989cellular component morphogenesis	37	0.0071
GO:0009267cellular response to starvation	19	0.0077
GO:0019748secondary metabolic process	43	0.0077
GO:0006950response to stress	161	0.0079
GO:0016036cellular response to phosphate starvation	14	0.0086
GO:0046283anthocyanin metabolic process	14	0.012
GO:0016137glycoside metabolic process	17	0.013
GO:0006811ion transport	51	0.013
GO:0015698inorganic anion transport	14	0.014
GO:0009698phenylpropanoid metabolic process	27	0.014
GO:0006576cellular biogenic amine metabolic process	13	0.014
GO:0009826unidimensional cell growth	27	0.015
GO:0060560developmental growth involved in morphogenesis	27	0.015
GO:0009628response to abiotic stimulus	109	0.015
GO:0009415response to water	31	0.017
GO:0009860pollen tube growth	17	0.018
GO:0000902cell morphogenesis	32	0.022
GO:0048588developmental cell growth	20	0.024
GO:0009714chalcone metabolic process	7	0.024
GO:0042181ketone biosynthetic process	7	0.024
GO:0009715chalcone biosynthetic process	7	0.024
GO:0044281small molecule metabolic process	131	0.025
GO:0044248cellular catabolic process	62	0.026
GO:0006012galactose metabolic process	6	0.027
GO:0006796phosphate metabolic process	93	0.029
GO:0006793phosphorus metabolic process	93	0.032

Enriched GO Term: Description	Number of associated genes	FDR
GO:0006066alcohol metabolic process	38	0.032
GO:0034656nucleobase, nucleoside and nucleotide catabolic process	15	0.032
GO:0034655nucleobase, nucleoside, nucleotide and nucleic acid catabolic process	15	0.032
GO:0006595polyamine metabolic process	7	0.04
GO:0044283small molecule biosynthetic process	65	0.04
GO:0006855multidrug transport	11	0.042
GO:0009856pollination	24	0.042
GO:0048610reproductive cellular process	17	0.042
GO:0051704multi-organism process	78	0.045
GO:0008361regulation of cell size	33	0.046
GO:0009611response to wounding	26	0.046
GO:0015893drug transport	11	0.046
GO:0046395carboxylic acid catabolic process	15	0.047
GO:0016054organic acid catabolic process	15	0.047
GO:0000904cell morphogenesis involved in differentiation	21	0.048
GO:0044270cellular nitrogen compound catabolic process	17	0.048
GO:0005215transporter activity	89	0.00069
GO:0022857transmembrane transporter activity	75	0.00092
GO:0022892substrate-specific transporter activity	69	0.0011
GO:0015291secondary active transmembrane transporter activity	29	0.0012
GO:0016798hydrolase activity, acting on glycosyl bonds	44	0.0026
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	42	0.0026
GO:0022891substrate-specific transmembrane transporter activity	62	0.0034
GO:0016791phosphatase activity	26	0.0035
GO:0022804active transmembrane transporter activity	42	0.0035
GO:0003824catalytic activity	466	0.0035
GO:0015103inorganic anion transmembrane transporter activity	15	0.0046

Enriched GO Term: Description	Number of associated genes	FDR
GO:0042578phosphoric ester hydrolase activity	29	0.0065
GO:0015075ion transmembrane transporter activity	46	0.012
GO:0015297antiporter activity	18	0.013
GO:0003779actin binding	12	0.013
GO:0016210naringenin-chalcone synthase activity	7	0.02
GO:0008509anion transmembrane transporter activity	18	0.02
GO:0015238drug transmembrane transporter activity	11	0.044
GO:0016787hydrolase activity	183	0.047
GO:0016020membrane	302	0.0026
GO:0005886plasma membrane	159	0.016

Table S8. Yellow Stage STAR Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=1). The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0018.1300.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase.	237.49	21.67	2.60689
chr5.CM1077.590.r2.m	Glyma.03G187000.1	UDP-glucosyl transferase 73C	81.84	1.67	4.479036
chr1.CM0122.1600.r2.m	Glyma.03G222100.1	Nitrogen regulatory protein P-II	118	15	2.151554
chr1.CM0012.990.r2.m	Glyma.04G254200.1	B3 DNA binding domain // Auxin response factor	111	12	2.383216
chr1.CM0064.910.r2.m	Glyma.05G076300.1	NA	725	113	1.863352
chr3.LjT40P18.80.r2.m	Glyma.07G021600.1	Shikimate O-hydroxycinnamoyltransferase.	329	23	3.015946
chr3.CM0452.240.r2.d	Glyma.07G021600.1	Shikimate O-hydroxycinnamoyltransferase.	8006	1066	2.091311
chr3.LjT10E18.60.r2.m	Glyma.10G210600.1	B3 DNA binding domain // Auxin response factor	192	19	2.513908
chr4.CM0432.2880.r2.m	Glyma.13G173300.1	O-METHYLTRANSFERASE	1884	169	2.660585
chr3.CM0106.330.r2.m	Glyma.13G344700.1	B-box zinc finger	603	34	3.327629
chr1.LjT35H06.80.r2.m	NA	NA	40	0	7.976076

Table S9. Red Stage STAR Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR = 0.00042) . The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0608.560.r2.m	Glyma.01G048200.1	RIBONUCLEASE T2	792	7451	-4.05106
chr3.LjB14006.120.r2.a	Glyma.01G166200.1	naringenin 3-dioxygenase	263.79	764.34	-2.3498
chr4.CM0119.240.r2.m	Glyma.01G166200.1	naringenin 3-dioxygenase	2001.21	6829.66	-2.58854
chr3.CM0590.770.r2.d	Glyma.01G228700.1	Naringenin-chalcone synthase	21.17	687.21	-5.83859
chr1.CM0284.250.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase	11.33	274.05	-5.43578
chr2.CM0018.760.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase	327.66	1550.24	-3.05735
chr2.CM0018.730.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase	86.27	323.5	-2.72878
chr3.CM0590.840.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase	21.17	687.21	-5.83859
chr2.CM0018.1190.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase	176.4	432.25	-2.11192
chr2.CM0018.1200.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase.	5.93	97.27	-4.79591
chr5.CM0077.110.r2.m	Glyma.02G158700.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase	29	799	-5.59379
chr5.CM0077.210.r2.m	Glyma.02G158700.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase	140	9266	-6.86429
chr1.CM0105.1870.r2.m	Glyma.03G113400.1	Proton-exporting ATPase.	285	949	-2.55225
chr1.CM0133.560.r2.m	Glyma.04G052100.1	Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies	40	2680	-6.87787
chr1.LjT39K18.30.r2.m	Glyma.04G147700.1	HEME OXYGENASE	546	1688	-2.44547
chr2.CM0191.680.r2.m	Glyma.04G227700.1	Quercetin 3-O-methyltransferase.	3	20	-3.48622

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr1.CM0410.430.r2. m	Glyma.08G247100.1	Transferase family	2201	7151	-2.51737
chr6.CM0013.1540.r2 .a	Glyma.09G038900.1	MYB-LIKE DNA-BINDING PROTEIN MYB	7	114	-4.81165
chr2.CM0021.2820.r2 .m	Glyma.11G027700.1	Leucocyanidin oxygenase.	929	14061	-4.73711
chr5.CM0077.120.r2. m	Glyma.14G072700.1	bifunctional dihydroflavonol 4- reductase/flavanone 4- reductase	14	60	-2.90284
chr5.CM0180.690.r2. m	Glyma.20G241700.1	chalcone isomerase	484	1165	-2.08436

Table S10. Yellow Stage STAR Results: Genes for enzymes annotated to GO:0009809 (Lignin biosynthetic process, FDR=0.41). The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr6.LjT35D18.30.r2.m	Glyma.03G223000.1	Methionine adenosyltransferase	16493	2557	1.87182
chr3.LjT40P18.80.r2.m	Glyma.07G021600.1	cinnamoyl-CoA reductase	329	23	3.01594
chr3.CM0452.240.r2.d	Glyma.07G021600.1	cinnamoyl-CoA reductase	8006	1066	2.09131
chr3.LjT45M09.50.r2.d	Glyma.07G023700.1	Shikimate O-hydroxycinnamoyltransferase.	1426	220	1.87849
chr4.CM0387.960.r2.m	Glyma.07G258700.1	beta-glucosidase	106	1	5.78762
chr6.CM0385.240.r2.d	Glyma.09G038500.1	CHITINASE	1058	11	5.75861
chr5.CM0200.3250.r2.m	Glyma.10G262400.1	cinnamyl-alcohol dehydrogenase	3805	583	1.88869

*Lignin biosynthetic process was not found to be expressed at the red stage of floral colour change in *L. filicaulis*, given the STAR results.

Table S11. Trinity Results: Significantly enriched GO terms corresponding to the yellow stage of a *L. filicaulis* flower.

Enriched GO category:Description	Number of associated genes	FDR
GO:0009790embryonic development	162	1.60E-16
GO:0032501multicellular organismal process	557	6.60E-16
GO:0032502developmental process	566	1.30E-15
GO:0007275multicellular organismal development	532	5.60E-15
GO:0009628response to abiotic stimulus	399	1.90E-10
GO:0009791post-embryonic development	314	2.80E-10
GO:0009987cellular process	1908	1.10E-09
GO:0044237cellular metabolic process	1498	1.10E-09
GO:0048856anatomical structure development	434	1.20E-09
GO:0000003reproduction	319	1.30E-09
GO:0050896response to stimulus	859	1.70E-09
GO:0008152metabolic process	1871	1.80E-09
GO:0006396RNA processing	140	2.40E-09
GO:0044281small molecule metabolic process	462	5.30E-08
GO:0009793embryonic development ending in seed dormancy	101	7.00E-08
GO:0009058biosynthetic process	821	4.60E-07
GO:0022414reproductive process	296	1.00E-06
GO:0010467gene expression	548	1.30E-06
GO:0044238primary metabolic process	1517	1.40E-06
GO:0003006reproductive developmental process	260	2.30E-06
GO:0010154fruit development	136	2.30E-06
GO:0040007growth	142	2.70E-06
GO:0044085cellular component biogenesis	161	4.10E-06
GO:0042254ribosome biogenesis	46	4.40E-06
GO:0009653anatomical structure morphogenesis	203	5.10E-06
GO:0048316seed development	126	6.00E-06
GO:0065007biological regulation	739	6.10E-06
GO:0022613ribonucleoprotein complex biogenesis	49	7.00E-06
GO:0044249cellular biosynthetic process	783	7.90E-06
GO:0009416response to light stimulus	178	1.10E-05
GO:0048731system development	254	1.10E-05
GO:0009314response to radiation	180	1.30E-05
GO:0008610lipid biosynthetic process	133	1.40E-05
GO:0034470ncRNA processing	58	2.60E-05
GO:0034660ncRNA metabolic process	73	3.50E-05
GO:0048608reproductive structure development	226	3.60E-05
GO:0050789regulation of biological process	641	4.00E-05
GO:0006950response to stress	519	4.90E-05

Enriched GO category:Description	Number of associated genes	FDR
GO:0048518positive regulation of biological process	128	4.90E-05
GO:0042221response to chemical stimulus	496	5.30E-05
GO:0009607response to biotic stimulus	224	6.20E-05
GO:0048513organ development	247	8.00E-05
GO:0009887organ morphogenesis	73	8.80E-05
GO:0015979photosynthesis	67	9.30E-05
GO:0009409response to cold	110	9.60E-05
GO:0044283small molecule biosynthetic process	211	0.00023
GO:0005976polysaccharide metabolic process	80	0.00025
GO:0051707response to other organism	210	0.00026
GO:0051704multi-organism process	255	0.00026
GO:0006412translation	130	0.00026
GO:0008219cell death	64	0.00029
GO:0016265death	64	0.00029
GO:0009266response to temperature stimulus	140	0.0003
GO:0043436oxoacid metabolic process	225	0.0003
GO:0019752carboxylic acid metabolic process	225	0.0003
GO:0006082organic acid metabolic process	225	0.0003
GO:0010033response to organic substance	317	0.00031
GO:0016043cellular component organization	331	0.00035
GO:0040008regulation of growth	46	0.00038
GO:0007166cell surface receptor linked signaling pathway	33	0.00043
GO:0016072rRNA metabolic process	32	0.00044
GO:0048869cellular developmental process	165	0.0006
GO:0042180cellular ketone metabolic process	228	0.00078
GO:0009451RNA modification	33	0.00082
GO:0000271polysaccharide biosynthetic process	49	0.00093
GO:0044264cellular polysaccharide metabolic process	69	0.00095
GO:0006364rRNA processing	29	0.00097
GO:0006629lipid metabolic process	205	0.00099
GO:0033692cellular polysaccharide biosynthetic process	48	0.00099
GO:0009725response to hormone stimulus	237	0.0011
GO:0009719response to endogenous stimulus	256	0.0011
GO:0019748secondary metabolic process	120	0.0016
GO:0023033signaling pathway	67	0.002
GO:0044255cellular lipid metabolic process	150	0.002
GO:0009657plastid organization	55	0.0029
GO:0008380RNA splicing	37	0.0032

Enriched GO category:Description	Number of associated genes	FDR
GO:0009640photomorphogenesis	28	0.0037
GO:0046394carboxylic acid biosynthetic process	131	0.0037
GO:0016053organic acid biosynthetic process	131	0.0037
GO:0010608posttranscriptional regulation of gene expression	44	0.0037
GO:0019684photosynthesis, light reaction	39	0.0041
GO:0006996organelle organization	204	0.0042
GO:0006457protein folding	51	0.0045
GO:0010025wax biosynthetic process	20	0.0046
GO:0044267cellular protein metabolic process	474	0.0053
GO:0050794regulation of cellular process	534	0.0058
GO:0006519cellular amino acid and derivative metabolic process	176	0.0062
GO:0019538protein metabolic process	556	0.0062
GO:0048646anatomical structure formation involved in morphogenesis	57	0.0062
GO:0023052signaling	255	0.0067
GO:0009733response to auxin stimulus	86	0.0069
GO:0009637response to blue light	27	0.0069
GO:0012501programmed cell death	51	0.0074
GO:0006779porphyrin biosynthetic process	29	0.0078
GO:0006631fatty acid metabolic process	80	0.0096
GO:0005982starch metabolic process	26	0.01
GO:0065008regulation of biological quality	214	0.011
GO:0032787monocarboxylic acid metabolic process	120	0.011
GO:0009639response to red or far red light	53	0.011
GO:0046148pigment biosynthetic process	51	0.012
GO:0019953sexual reproduction	38	0.013
GO:0023034intracellular signaling pathway	33	0.014
GO:0030154cell differentiation	120	0.015
GO:0033014tetrapyrrole biosynthetic process	30	0.015
GO:0000375RNA splicing, via transesterification reactions	23	0.016
GO:0007389pattern specification process	53	0.018
GO:0003002regionalization	47	0.02
GO:0008283cell proliferation	32	0.02
GO:0009606tropism	27	0.02
GO:0006633fatty acid biosynthetic process	62	0.02
GO:0023060signal transmission	200	0.023
GO:0006778porphyrin metabolic process	34	0.024

Enriched GO category:Description	Number of associated genes	FDR
GO:0023046signaling process	200	0.025
GO:0043067regulation of programmed cell death	24	0.026
GO:0042440pigment metabolic process	57	0.027
GO:0016071mRNA metabolic process	40	0.031
GO:0000377RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	22	0.031
GO:0033013tetrapyrrole metabolic process	35	0.032
GO:0006073cellular glucan metabolic process	50	0.041
GO:0022621shoot system development	94	0.043
GO:0048367shoot development	94	0.043
GO:0032989cellular component morphogenesis	91	0.044
GO:0006875cellular metal ion homeostasis	16	0.044
GO:0048569post-embryonic organ development	58	0.048
GO:0003824catalytic activity	1566	6.40E-06
GO:0009055electron carrier activity	60	2.10E-05
GO:0005198structural molecule activity	108	0.0096
GO:0016740transferase activity	596	0.0096
GO:0004872receptor activity	38	0.019
GO:0005623cell	2324	2.10E-90
GO:0044464cell part	2324	2.10E-90
GO:0043226organelle	1659	5.60E-66
GO:0043229intracellular organelle	1656	9.90E-66
GO:0044424intracellular part	1838	1.20E-65
GO:0005622intracellular	1866	2.20E-65
GO:0043227membrane-bounded organelle	1581	1.60E-64
GO:0043231intracellular membrane-bounded organelle	1576	3.10E-64
GO:0005737cytoplasm	1511	1.00E-62
GO:0044444cytoplasmic part	1406	5.50E-57
GO:0009536plastid	696	8.80E-55
GO:0009507chloroplast	659	4.80E-53
GO:0044435plastid part	414	1.10E-32
GO:0044434chloroplast part	398	9.00E-30
GO:0044446intracellular organelle part	814	1.90E-28
GO:0044422organelle part	817	2.40E-28
GO:0012505endomembrane system	148	1.80E-23
GO:0031976plastid thylakoid	188	2.40E-23
GO:0009534chloroplast thylakoid	188	2.40E-23
GO:0031984organelle subcompartment	188	3.10E-23

Enriched GO category:Description	Number of associated genes	FDR
GO:0044436thylakoid part	176	4.20E-23
GO:0009579thylakoid	212	3.90E-22
GO:0042651thylakoid membrane	164	1.00E-21
GO:0055035plastid thylakoid membrane	159	1.00E-21
GO:0009535chloroplast thylakoid membrane	158	1.50E-21
GO:0034357photosynthetic membrane	165	3.20E-20
GO:0005739mitochondrion	261	4.30E-13
GO:0009532plastid stroma	206	2.00E-12
GO:0031975envelope	243	9.20E-12
GO:0009526plastid envelope	193	1.60E-11
GO:0031967organelle envelope	240	2.40E-11
GO:0009570chloroplast stroma	192	2.40E-10
GO:0009941chloroplast envelope	183	2.90E-10
GO:0031974membrane-enclosed lumen	197	1.00E-09
GO:0070013intracellular organelle lumen	194	1.20E-09
GO:0043233organelle lumen	194	1.40E-09
GO:0043232intracellular non-membrane-bounded organelle	280	3.80E-09
GO:0043228non-membrane-bounded organelle	280	3.80E-09
GO:0031090organelle membrane	332	1.60E-08
GO:0009505plant-type cell wall	92	1.00E-07
GO:0010287plastoglobule	36	1.10E-06
GO:0016020membrane	971	1.40E-06
GO:0030529ribonucleoprotein complex	142	2.30E-06
GO:0044428nuclear part	183	4.20E-06
GO:0031977thylakoid lumen	37	8.10E-06
GO:0005634nucleus	524	9.50E-06
GO:0032991macromolecular complex	360	5.30E-05
GO:0031978plastid thylakoid lumen	26	0.0001
GO:0009543chloroplast thylakoid lumen	26	0.0001
GO:0030312external encapsulating structure	164	0.00011
GO:0031224intrinsic to membrane	250	0.00016
GO:0005618cell wall	159	0.00039
GO:0005730nucleolus	100	0.00051
GO:0005576extracellular region	155	0.00074
GO:0005840ribosome	104	0.00097
GO:0031981nuclear lumen	134	0.0017
GO:0031225anchored to membrane	44	0.0063
GO:0009706chloroplast inner membrane	22	0.013

Enriched GO category:Description	Number of associated genes	FDR
GO:0009295nucleoid	17	0.015
GO:0042170plastid membrane	34	0.016
GO:0005794Golgi apparatus	162	0.017
GO:0016021integral to membrane	204	0.02
GO:0009528plastid inner membrane	22	0.022
GO:0010319stromule	19	0.022
GO:0042646plastid nucleoid	14	0.023
GO:0031969chloroplast membrane	30	0.03
GO:0005887integral to plasma membrane	13	0.035
GO:0005874microtubule	24	0.043
GO:0009521photosystem	22	0.043
GO:0044425membrane part	429	0.048

Table 12. Trinity Results: Significantly enriched GO terms corresponding to the red stage of a *L. filicaulis* flower.

Enriched GO category:Description	Number of associated genes	FDR
GO:0050896response to stimulus	426	2.60E-23
GO:0006950response to stress	280	1.20E-17
GO:0023052signaling	168	1.60E-17
GO:0065007biological regulation	368	1.60E-17
GO:0032502developmental process	256	2.80E-16
GO:0050794regulation of cellular process	289	1.80E-15
GO:0007242intracellular signaling cascade	111	8.50E-15
GO:0032501multicellular organismal process	245	8.50E-15
GO:0050789regulation of biological process	318	1.10E-14
GO:0009607response to biotic stimulus	134	1.40E-14
GO:0023060signal transmission	133	2.90E-14
GO:0023046signaling process	133	3.00E-14
GO:0051707response to other organism	127	9.80E-14
GO:0007275multicellular organismal development	232	1.70E-13
GO:0007165signal transduction	128	2.30E-13
GO:0009987cellular process	780	9.60E-13
GO:0048856anatomical structure development	202	2.50E-12
GO:0042221response to chemical stimulus	246	2.90E-12
GO:0051179localization	226	8.80E-12
GO:0006575cellular amino acid derivative metabolic process	74	1.00E-11
GO:0023034intracellular signaling pathway	34	1.00E-11
GO:0051704multi-organism process	141	1.10E-11
GO:0042398cellular amino acid derivative biosynthetic process	60	1.10E-11
GO:0009699phenylpropanoid biosynthetic process	45	2.50E-11
GO:0023033signaling pathway	51	2.70E-11
GO:0010033response to organic substance	167	4.30E-11
GO:0009719response to endogenous stimulus	142	4.60E-11
GO:0009790embryonic development	69	4.70E-11
GO:0009751response to salicylic acid stimulus	42	5.60E-11
GO:0009653anatomical structure morphogenesis	106	3.60E-10
GO:0051234establishment of localization	211	4.40E-10
GO:0009605response to external stimulus	86	8.00E-10
GO:0006810transport	208	1.20E-09
GO:0019748secondary metabolic process	74	1.30E-09
GO:0044238primary metabolic process	626	2.30E-09
GO:0009698phenylpropanoid metabolic process	50	3.50E-09
GO:0007166cell surface receptor linked signaling pathway	26	4.60E-09

Enriched GO category:Description	Number of associated genes	FDR
GO:0012501programmed cell death	39	5.00E-09
GO:0006519cellular amino acid and derivative metabolic process	101	5.00E-09
GO:0008219cell death	42	6.30E-09
GO:0016265death	42	6.30E-09
GO:0009813flavonoid biosynthetic process	32	1.00E-08
GO:0048518positive regulation of biological process	70	1.40E-08
GO:0048731system development	123	2.10E-08
GO:0009725response to hormone stimulus	122	4.40E-08
GO:0009056catabolic process	109	5.80E-08
GO:0009624response to nematode	32	6.70E-08
GO:0000003reproduction	136	7.10E-08
GO:0048513organ development	120	1.20E-07
GO:0009628response to abiotic stimulus	163	2.00E-07
GO:0009812flavonoid metabolic process	33	2.30E-07
GO:0044237cellular metabolic process	589	2.80E-07
GO:0009723response to ethylene stimulus	34	4.50E-07
GO:0009753response to jasmonic acid stimulus	43	4.50E-07
GO:0022414reproductive process	131	5.00E-07
GO:0009620response to fungus	48	7.50E-07
GO:0040007growth	67	1.10E-06
GO:0031323regulation of cellular metabolic process	169	2.40E-06
GO:0007610behavior	15	2.60E-06
GO:0002376immune system process	55	2.90E-06
GO:0009058biosynthetic process	332	4.50E-06
GO:0042493response to drug	21	4.90E-06
GO:0044283small molecule biosynthetic process	99	6.50E-06
GO:0061024membrane organization	27	7.00E-06
GO:0016044cellular membrane organization	27	7.00E-06
GO:0003006reproductive developmental process	111	1.40E-05
GO:0080090regulation of primary metabolic process	159	1.40E-05
GO:0006955immune response	52	2.70E-05
GO:0006952defense response	100	2.80E-05
GO:0031326regulation of cellular biosynthetic process	152	2.90E-05
GO:0050832defense response to fungus	36	3.10E-05
GO:0009889regulation of biosynthetic process	152	3.20E-05
GO:0048869cellular developmental process	78	3.70E-05
GO:0044281small molecule metabolic process	183	3.80E-05
GO:0015893drug transport	19	4.20E-05

Enriched GO category:Description	Number of associated genes	FDR
GO:0044248cellular catabolic process	89	5.60E-05
GO:0044249cellular biosynthetic process	314	7.90E-05
GO:0006873cellular ion homeostasis	20	0.0001
GO:0055082cellular chemical homeostasis	20	0.0001
GO:0048608reproductive structure development	97	0.00011
GO:0008152metabolic process	710	0.00011
GO:0051171regulation of nitrogen compound metabolic process	145	0.00011
GO:0019222regulation of metabolic process	173	0.00012
GO:0009308amine metabolic process	67	0.00012
GO:0045449regulation of transcription	138	0.00014
GO:0010556regulation of macromolecule biosynthetic process	142	0.00014
GO:0009793embryonic development ending in seed dormancy	40	0.00018
GO:0019219regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	142	0.00018
GO:0009887organ morphogenesis	34	0.0002
GO:0048646anatomical structure formation involved in morphogenesis	31	0.0002
GO:0009611response to wounding	39	0.0002
GO:0009737response to abscisic acid stimulus	58	0.00021
GO:0009408response to heat	32	0.00023
GO:0016042lipid catabolic process	19	0.00023
GO:0016043cellular component organization	141	0.00023
GO:0031667response to nutrient levels	28	0.00024
GO:0045087innate immune response	45	0.00027
GO:0043067regulation of programmed cell death	16	0.00027
GO:0044242cellular lipid catabolic process	16	0.00027
GO:0007568aging	23	0.00031
GO:0009991response to extracellular stimulus	31	0.00033
GO:0019953sexual reproduction	22	0.00039
GO:0009266response to temperature stimulus	62	0.00042
GO:0006350transcription	145	0.00053
GO:0009791post-embryonic development	113	0.00054
GO:0043069negative regulation of programmed cell death	11	0.00056
GO:0000278mitotic cell cycle	17	0.00062
GO:0060255regulation of macromolecule metabolic process	151	0.00069
GO:0009863salicylic acid mediated signaling pathway	14	0.00076
GO:0040008regulation of growth	22	0.00078
GO:0016192vesicle-mediated transport	36	0.00084
GO:0048522positive regulation of cellular process	42	0.00085

Enriched GO category:Description	Number of associated genes	FDR
GO:0042594response to starvation	24	0.00088
GO:0046942carboxylic acid transport	20	0.00088
GO:0010468regulation of gene expression	146	0.00089
GO:0006855multidrug transport	16	0.0009
GO:0010646regulation of cell communication	25	0.0009
GO:0006928cellular component movement	15	0.0011
GO:0060548negative regulation of cell death	11	0.0011
GO:0071554cell wall organization or biogenesis	43	0.0011
GO:0006022aminoglycan metabolic process	8	0.0013
GO:0009966regulation of signal transduction	24	0.0013
GO:0023051regulation of signaling process	24	0.0013
GO:0010941regulation of cell death	16	0.0014
GO:0015849organic acid transport	20	0.0014
GO:0006979response to oxidative stress	50	0.0015
GO:0071446cellular response to salicylic acid stimulus	14	0.0015
GO:0071495cellular response to endogenous stimulus	54	0.0019
GO:0015837amine transport	15	0.0021
GO:0006629lipid metabolic process	86	0.0027
GO:0009733response to auxin stimulus	40	0.0029
GO:0019932second-messenger-mediated signaling	10	0.003
GO:0048878chemical homeostasis	29	0.003
GO:0009415response to water	39	0.0031
GO:0009414response to water deprivation	38	0.0037
GO:0048583regulation of response to stimulus	38	0.0037
GO:0006865amino acid transport	14	0.004
GO:0009814defense response, incompatible interaction	30	0.004
GO:0006796phosphate metabolic process	120	0.004
GO:0009739response to gibberellin stimulus	20	0.0044
GO:0006793phosphorus metabolic process	120	0.0045
GO:0000272polysaccharide catabolic process	13	0.0047
GO:0044106cellular amine metabolic process	54	0.0047
GO:0065008regulation of biological quality	92	0.0049
GO:0071310cellular response to organic substance	63	0.0049
GO:0009888tissue development	50	0.0049
GO:0009409response to cold	44	0.0052
GO:0022603regulation of anatomical structure morphogenesis	14	0.0058
GO:0042439ethanolamine and derivative metabolic process	7	0.006
GO:0007267cell-cell signaling	7	0.006

Enriched GO category:Description	Number of associated genes	FDR
GO:0048523negative regulation of cellular process	44	0.006
GO:0048316seed development	47	0.0065
GO:0007010cytoskeleton organization	26	0.0065
GO:0043687post-translational protein modification	127	0.0066
GO:0006576cellular biogenic amine metabolic process	15	0.0066
GO:0006464protein modification process	137	0.0066
GO:0030036actin cytoskeleton organization	18	0.0071
GO:0009057macromolecule catabolic process	48	0.0071
GO:0030029actin filament-based process	19	0.0072
GO:0010118stomatal movement	16	0.0072
GO:0009642response to light intensity	23	0.008
GO:0030154cell differentiation	53	0.0083
GO:0005975carbohydrate metabolic process	100	0.009
GO:0070887cellular response to chemical stimulus	71	0.0093
GO:0042538hyperosmotic salinity response	13	0.01
GO:0009873ethylene mediated signaling pathway	13	0.01
GO:0009416response to light stimulus	66	0.011
GO:0009755hormone-mediated signaling pathway	43	0.012
GO:0030003cellular cation homeostasis	13	0.013
GO:0009612response to mechanical stimulus	7	0.014
GO:0000902cell morphogenesis	38	0.014
GO:0007626locomotory behavior	8	0.014
GO:0006972hyperosmotic response	14	0.014
GO:0006811ion transport	60	0.015
GO:0010154fruit development	48	0.016
GO:0016567protein ubiquitination	20	0.016
GO:0009314response to radiation	66	0.016
GO:0043436oxoacid metabolic process	87	0.016
GO:0010054trichoblast differentiation	14	0.016
GO:0006082organic acid metabolic process	87	0.016
GO:0019752carboxylic acid metabolic process	87	0.016
GO:0016310phosphorylation	102	0.017
GO:0048585negative regulation of response to stimulus	16	0.018
GO:0032870cellular response to hormone stimulus	45	0.018
GO:0022610biological adhesion	7	0.019
GO:0007155cell adhesion	7	0.019
GO:0051094positive regulation of developmental process	14	0.019
GO:0046395carboxylic acid catabolic process	18	0.02

Enriched GO category:Description	Number of associated genes	FDR
GO:0016054organic acid catabolic process	18	0.02
GO:0044255cellular lipid metabolic process	60	0.021
GO:0048468cell development	35	0.022
GO:0043412macromolecule modification	140	0.023
GO:0071555cell wall organization	27	0.023
GO:0006026aminoglycan catabolic process	6	0.024
GO:0006030chitin metabolic process	6	0.024
GO:0006032chitin catabolic process	6	0.024
GO:0019722calcium-mediated signaling	7	0.026
GO:0032446protein modification by small protein conjugation	20	0.027
GO:0010119regulation of stomatal movement	11	0.027
GO:0015692lead ion transport	5	0.028
GO:0009963positive regulation of flavonoid biosynthetic process	5	0.028
GO:0007584response to nutrient	5	0.028
GO:0048645organ formation	10	0.028
GO:0032989cellular component morphogenesis	40	0.028
GO:0070882cellular cell wall organization or biogenesis	17	0.029
GO:0010053root epidermal cell differentiation	16	0.03
GO:0050801ion homeostasis	21	0.031
GO:0016049cell growth	38	0.032
GO:0009867jasmonic acid mediated signaling pathway	13	0.032
GO:0007264small GTPase mediated signal transduction	17	0.033
GO:0006754ATP biosynthetic process	9	0.033
GO:0019725cellular homeostasis	24	0.033
GO:0048519negative regulation of biological process	55	0.034
GO:0032535regulation of cellular component size	42	0.035
GO:0090066regulation of anatomical structure size	42	0.035
GO:0042180cellular ketone metabolic process	87	0.035
GO:0006644phospholipid metabolic process	18	0.035
GO:0048532anatomical structure arrangement	12	0.035
GO:0007389pattern specification process	23	0.035
GO:0071395cellular response to jasmonic acid stimulus	13	0.037
GO:0023057negative regulation of signaling process	11	0.037
GO:0009968negative regulation of signal transduction	11	0.037
GO:0010648negative regulation of cell communication	11	0.037
GO:0008361regulation of cell size	39	0.038
GO:0023056positive regulation of signaling process	8	0.038
GO:0009967positive regulation of signal transduction	8	0.038

Enriched GO category:Description	Number of associated genes	FDR
GO:0002218activation of innate immune response	10	0.039
GO:0002253activation of immune response	10	0.039
GO:0006839mitochondrial transport	9	0.039
GO:0048584positive regulation of response to stimulus	16	0.041
GO:0070647protein modification by small protein conjugation or removal	21	0.044
GO:0022403cell cycle phase	17	0.045
GO:0006970response to osmotic stress	53	0.046
GO:0000087M phase of mitotic cell cycle	10	0.046
GO:0007067mitosis	10	0.046
GO:0009055electron carrier activity	56	1.10E-19
GO:0003824catalytic activity	688	7.80E-17
GO:0030528transcription regulator activity	119	1.40E-14
GO:0022804active transmembrane transporter activity	76	2.40E-11
GO:0005215transporter activity	135	3.30E-11
GO:0003700transcription factor activity	97	4.90E-10
GO:0022857transmembrane transporter activity	107	3.90E-08
GO:0015291secondary active transmembrane transporter activity	45	3.90E-08
GO:0015294solute:cation symporter activity	19	1.30E-07
GO:0043169cation binding	268	1.60E-07
GO:0022892substrate-specific transporter activity	97	2.30E-07
GO:0042578phosphoric ester hydrolase activity	47	3.20E-07
GO:0015293symporter activity	19	3.70E-07
GO:0004091carboxylesterase activity	36	1.40E-06
GO:0016787hydrolase activity	266	1.80E-06
GO:0015295solute:hydrogen symporter activity	14	7.40E-06
GO:0016298lipase activity	24	9.40E-06
GO:0022891substrate-specific transmembrane transporter activity	85	1.20E-05
GO:0004497monooxygenase activity	26	2.30E-05
GO:0005509calcium ion binding	40	2.50E-05
GO:0000287magnesium ion binding	25	2.70E-05
GO:0016791phosphatase activity	36	2.90E-05
GO:0004722protein serine/threonine phosphatase activity	18	5.20E-05
GO:0008324cation transmembrane transporter activity	51	6.40E-05
GO:0016563transcription activator activity	11	8.70E-05
GO:0019787small conjugating protein ligase activity	31	8.80E-05
GO:0015075ion transmembrane transporter activity	64	8.80E-05
GO:0016788hydrolase activity, acting on ester bonds	105	0.00011
GO:0046872metal ion binding	247	0.00012

Enriched GO category:Description	Number of associated genes	FDR
GO:0051119sugar transmembrane transporter activity	14	0.00013
GO:0016874ligase activity	51	0.00014
GO:0016301kinase activity	122	0.00015
GO:0016879ligase activity, forming carbon-nitrogen bonds	41	0.00018
GO:0016638oxidoreductase activity, acting on the CH-NH2 group of donors	10	0.0002
GO:0016881acid-amino acid ligase activity	34	0.00023
GO:0046943carboxylic acid transmembrane transporter activity	19	0.00024
GO:0004842ubiquitin-protein ligase activity	29	0.00034
GO:0005342organic acid transmembrane transporter activity	19	0.00044
GO:0005506iron ion binding	59	0.00044
GO:0008134transcription factor binding	15	0.0006
GO:0016798hydrolase activity, acting on glycosyl bonds	53	0.0006
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	50	0.00087
GO:0015297antiporter activity	23	0.0013
GO:0004672protein kinase activity	100	0.0014
GO:0015144carbohydrate transmembrane transporter activity	14	0.0014
GO:0016740transferase activity	242	0.0014
GO:0016773phosphotransferase activity, alcohol group as acceptor	113	0.0014
GO:0004721phosphoprotein phosphatase activity	19	0.0018
GO:0015662ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism	9	0.0028
GO:0008194UDP-glycosyltransferase activity	24	0.0047
GO:0005275amine transmembrane transporter activity	14	0.0052
GO:0015171amino acid transmembrane transporter activity	13	0.0065
GO:0015179L-amino acid transmembrane transporter activity	6	0.0078
GO:0015197peptide transporter activity	6	0.0078
GO:0004674protein serine/threonine kinase activity	80	0.0092
GO:0008081phosphoric diester hydrolase activity	11	0.0095
GO:0015300solute:solute antiporter activity	13	0.011
GO:0015238drug transmembrane transporter activity	13	0.017
GO:0030246carbohydrate binding	27	0.017
GO:0004568chitinase activity	6	0.019
GO:0030234enzyme regulator activity	27	0.029
GO:0008415acyltransferase activity	22	0.03
GO:0015399primary active transmembrane transporter activity	22	0.033
GO:0005337nucleoside transmembrane transporter activity	5	0.037
GO:0012505endomembrane system	126	5.00E-52
GO:0044464cell part	882	5.80E-40

Enriched GO category:Description	Number of associated genes	FDR
GO:0005623cell	882	5.80E-40
GO:0031224intrinsic to membrane	166	3.80E-20
GO:0016021integral to membrane	141	7.10E-16
GO:0009505plant-type cell wall	47	1.70E-07
GO:0005622intracellular	600	5.40E-07
GO:0044424intracellular part	578	1.30E-05
GO:0005887integral to plasma membrane	13	1.80E-05
GO:0044425membrane part	201	4.00E-05
GO:0005739mitochondrion	94	0.00012
GO:0031410cytoplasmic vesicle	23	0.00012
GO:0008287protein serine/threonine phosphatase complex	10	0.00013
GO:0031982vesicle	25	0.00014
GO:0031988membrane-bounded vesicle	23	0.00078
GO:0016023cytoplasmic membrane-bounded vesicle	21	0.00078
GO:0043226organelle	482	0.0073
GO:0043229intracellular organelle	481	0.0077
GO:0043227membrane-bounded organelle	450	0.023
GO:0043231intracellular membrane-bounded organelle	449	0.023
GO:0005634nucleus	197	0.023
GO:0045177apical part of cell	7	0.036
GO:0005856cytoskeleton	30	0.043

Table S13. Red Stage Trinity Results: Genes for enzymes annotated to the over-represented GO:0009813 (Flavonoid biosynthetic process, FDR=1.0e-08). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR29586 c1_g3	chr5.CM0077.790.nc	Glyma.17G252200.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR1)	0	78.33	-9.4766
TR29586 c0_g1	chr5.CM0077.690.nc	Glyma.14G072700.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR3)	148	10163	-6.4567
TR17543 c0_g1	chr5.CM0077.790.nc	Glyma.14G072700.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR5)	0	28	-8.0022
TR29586 c1_g4	chr5.CM0077.790.nc	Glyma.17G252200.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR5)	0	65.1	-9.2140
TR29586 c2_g1	chr5.CM0077.790.nc	Glyma.17G252200.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR5)	9	195	-4.7724
TR29586 c1_g1	chr5.CM0077.790.nc	Glyma.17G252200.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR5)	6	38.56	-3.0276
TR52576 c0_g1	chr5.CM0077.780.nc	Glyma.17G252200.1	bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase (DFR5)	7	33	- 2.5697 2
TR43444 c0_g2	chr1.CM0133.170.nc	Glyma.17G227500.1	Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies	6	98.36	-4.3545
TR43444 c0_g1	chr1.CM0133.170.nc	Glyma.17G227500.1	Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies	0	13.64	-7.0078

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR23263 c0_g1	LjSGA_060583.1	Glyma.17G227500.1	Cytochrome P450 CYP3/CYP5/CYP6/CYP 9 subfamilies	13	182	-4.1492
TR23981 c0_g1	chr1.CM0133.170.nc	Glyma.04G052100.1	Cytochrome P450 CYP3/CYP5/CYP6/CYP 9 subfamilies (putative MAX1)	45	2897	-6.3605
TR38960 c0_g1	chr4.CM0119.240.nc	Glyma.02G048400.1	Flavanone 3- dioxygenase (F3H1)	1830	6123	- 2.0988 4
TR61355 c0_g2	LjSGA_043399.2	Glyma.20G009300.2	Flavanone 3- dioxygenase (F3H1)	117.49	453.22	-2.3081
TR13112 c0_g1	chr2.CM0124.280.nd	Glyma.02G124700.1	Flavanone 3- dioxygenase (F3H2)	6	33	-2.7873
TR53383 c0_g1	chr4.CM0387.340.nc	Glyma.08G092800.1	Flavanone 3- dioxygenase (F3H5)	1	50	-5.8140
TR50359 c0_g2	LjSGA_025000.1	Glyma.12G235300.1	Flavanone 3- dioxygenase (F3H8)	4	40	-3.6326
TR46116 c0_g1*	chr2.CM0124.160.nd	Glyma.02G124700.1	Flavanone 3- dioxygenase *	1	21	-4.5668
TR52191 c0_g1	chr4.CM0429.360.nc	Glyma.06G202300.1	Flavonoid 3'- monooxygenase (F3'H1)	1586	15638	-3.6580
TR38633 c0_g1	chr6.CM0055.290.nc	Glyma.09G146300.1	Kinesin motor domain // PPR repeat // Trehalose- phosphatase	1219	3586	-1.9130
TR50049 c0_g2	chr2.CM0304.350.nc	Glyma.11G027700.1	Leucocyanidin oxygenase (ANS1)	0	165.97	-10.565
TR50049 c0_g1	chr2.CM0304.350.nc	Glyma.11G027700.1	Leucocyanidin oxygenase (ANS1)	1010	14783. 03	-4.2278
TR42413 c0_g1	chr2.LjT36E17.150.n d	Glyma.14G032400.1	MATE EFFLUX FAMILY PROTEIN	0	53	-8.9201
TR32220 c0_g1	LjT02E24.110.nc	Glyma.09G038900.1	MYB-LIKE DNA- BINDING PROTEIN MYB	7	111	- 4.3163 8

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR19889 c3_g1	chr1.CM0591.390.nd	Glyma.11G011500.1	Naringenin-chalcone synthase (CHS1)	1385	19064	-4.1392
TR48518 c0_g1	chr2.CM0018.700.nc	Glyma.11G011500.1	Naringenin-chalcone synthase (CHS2)	77	430	-2.8356
TR19889 c2_g1	chr2.CM0018.710.nc	Glyma.01G228700.1	Naringenin-chalcone synthase (CHS9)	17	207	-3.9514
TR51814 c0_g1	chr1.CM0105.880.nc	Glyma.03G262600.1	Proton-exporting ATPase.	272	951	-2.1617
TR1381 c0_g1	chr1.CM0105.600.nc	Glyma.07G048700.1	Quercetin 3-O-methyltransferase (3-OMT)	4	126	-5.2853
TR57984 c0_g1	chr2.CM0641.550.nc	Glyma.01G048400.1	Ribonuclease T(2)	383	12922	-5.4323
TR57984 c0_g2	chr2.CM0641.560.nc	Glyma.01G048200.1	RIBONUCLEASE T2	850	8077	-3.6046
TR45786 c0_g1	chr3.LjT46L11.20.nc	Glyma.12G238900.1	Thaumatococcus family	1	261	-8.1955
TR14130 c0_g1	chr1.CM0063.200.nc	Glyma.08G010500.1	Transferase family	1	29	-5.0304
TR45798 c0_g1	chr1.CM0371.40.nc	Glyma.16G039500.1	Transferase family	21	432	-4.7098
TR33284 c0_g1	chr1.CM0410.130.nc	Glyma.18G268100.1	Transferase family	2499	8136	-
						2.0594
						4
TR34067 c0_g2	LjSGA_053613.2	Glyma.10G062200.1	UDP-glucosyl transferase 73C (UGT)	0	16.8	-7.2859
TR48968 c0_g1	chr1.CM0012.680.nd	Glyma.19G187000.1	UDP-glucosyl transferase 73C (UGT)	164	5940	-5.5340

Table S14. Yellow Stage Trinity Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=1). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST equivalent, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR27395 c0_g1	chr1.CM0122.1600.r2.m	Glyma.03G222100.2	Nitrogen regulatory protein P-II	119	14	2.7213
TR4932 c0_g1	chr1.CM0476.410.r2.a	Glyma.04G227700.1	Quercetin 3-O-methyltransferase (3-OMT)	5616	1315	1.7378
TR4801 c6_g1	LjSGA_035363.1	Glyma.05G052000.1	Cytochrome P450	57	0	8.6691
TR31484 c0_g1	chr1.CM0064.910.r2.m	Glyma.05G076300.1	NA	1020	163	2.2883
TR35194 c0_g1	chr3.CM0452.240.r2.d	Glyma.08G220200.1	Shikimate O-hydroxycinnamoyltransferase (HCT)	8894	1190	2.5452
TR34244 c0_g1	LjSGA_050386.1	Glyma.09G186300.1	Cytochrome P450 CYP2 subfamily	74	14	2.0370
TR32422 c0_g1	chr5.CM1125.790.r2.a	Glyma.10G026000.1	EGL3 (ENHANCER OF GLABRA3), DNA BINDING / TRANSCRIPTION FACTOR (EGL3)	131.18	0	9.8676
TR66030 c0_g1	chr1.CM0012.990.r2.m	Glyma.10G053500.1	AUX/IAA family // B3 DNA binding domain // Auxin response factor	169	23	2.5151
TR34067 c0_g1	chr5.CM1077.590.r2.m	Glyma.10G062200.1	UDP-glucosyl transferase 73C (UGT73C)	170	0.2	10.2433
TR34449 c0_g1	chr3.LjT10E18.60.r2.m	Glyma.12G076200.1	AUX/IAA family // B3 DNA binding domain // Auxin response factor	265	29	2.8306
TR46051 c0_g2	LjSGA_013806.1	Glyma.12G109800.1	Quercetin 3-O-methyltransferase (3-OMT)	4923.93	244.53	3.9718

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log2FC
TR28709 c0_g1	chr4.CM0432.2880.r2.m	Glyma.13G173300.1	O-METHYLTRANSFERASE (OMT)	2026.03	168.18	3.2347
TR28709 c0_g2	chr4.CM0432.2880.r2.m	Glyma.13G173300.1	O-METHYLTRANSFERASE (OMT)	231.97	25.82	2.7958
TR39181 c0_g1	chr3.CM0106.330.r2.m	Glyma.13G344700.1	B-box zinc finger	625	32	3.9265
TR30560 c0_g1	chr3.LjT40P18.80.r2.m	Glyma.13G371000.1	Shikimate O-hydroxycinnamoyltransferase (HCT)	259.39	14.83	3.7436
TR30560 c0_g2	chr3.LjT40P18.80.r2.m	Glyma.13G371000.1	Shikimate O-hydroxycinnamoyltransferase (HCT)	80.61	7.17	3.1560
TR3240 c0_g1	LjSGA_097658.1	Glyma.15G154100.1	Flavanone 3-hydroxylase (F3H)	1520	83	3.8365
TR11418 c0_g1	LjSGA_027597.1	Glyma.15G221300.1	Flavonol 3-O-glucosyltransferase (UFGT)	1126	23	5.2502
TR24510 c0_g2	LjSGA_071841.2	Glyma.15G221300.1	Flavonol 3-O-glucosyltransferase (UFGT)	134	0	9.9003
TR20072 c0_g1	chr6.CM0013.1540.r2.a	Glyma.16G073000.1	myb proto-oncogene protein, plant	315	0	11.1325
TR4801 c4_g1	LjSGA_035363.1	Glyma.17G134100.1	Cytochrome P450 CYP2 subfamily	77	0	9.1021
TR26322 c0_g1	chr1.CM0410.460.r2.m	Glyma.18G268100.1	Transferase family	534	104	2.0025
TR50011 c0_g1	chr1.CM0104.1140.r2.m	Glyma.19G105100.1	Naringenin-chalcone synthase (CHS)	24259	5817	1.7036
TR19685 c0_g1	LjSGA_045968.2	Glyma.20G180000.1	B3 DNA binding domain // Auxin response factor	26.01	0	7.5409

Table S15. Similarly Expressed Trinity Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=1). The first two columns list Trinity contigs expresand gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR43354 c0_g1	chr5.CM0180.690.r2.m	Glyma.20G241700.1	chalcone isomerase (<i>CHI</i>)	61.37	168.88	-1.6533
TR20726 c0_g1	LjSGA_018608.1	Glyma.18G048300.1	Flavanone 3-hydroxylase (<i>F3H2</i>)	14.78	10.7	0.2720
TR20726 c0_g2	LjSGA_018608.1	Glyma.18G048300.1	Flavanone 3-hydroxylase (<i>F3H2</i>)	4.18	2.31	0.6411
TR9870 c1_g1	LjSGA_041422.1	Glyma.08G149800.1	Flavanone 3-hydroxylase (<i>F3H3</i>)	2.73	0	4.4836
TR31785 c0_g1	LjSGA_067411.1	Glyma.02G136200.1	Flavanone 3-dioxygenase (<i>F3H4</i>)	1218.06	1201.12	0.7287
TR123 c0_g1	chr2.CM0191.680.r2.m	Glyma.04G227700.1	Quercetin 3-O-methyltransferase (<i>3-OMT</i>)	0.27	1.84	-2.7781
TR31193 c0_g1	chr4.CM0227.690.r2.m	Glyma.02G104700.1	Flavonol 3-O-glucosyltransferase (<i>UFGT</i>)	3.44	1.08	1.5431
TR45767 c0_g1	chr4.CM1616.680.r2.m	Glyma.08G062000.1	Anthocyanidin reductase (<i>ANR</i>)	0.46	2.7	-2.5202
TR3325 c1_g1	chr2.CM0021.2950.r2.m	Glyma.01G215400.1	3,5-epimerase/4-reductase (<i>GME</i>)	9.01	12.07	-0.6159
TR18411 c0_g1	LjSGA_047295.1	Glyma.01G096600.1	transcription factor MYC2	24.74	7.74	1.4663
TR38249 c0_g1	chr4.CM0007.1140.r2.m	Glyma.05G127300.2	26S proteasome regulatory subunit N12	62.23	63.42	-0.2245

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log2FC
TR34052 c0_g1	chr2.CM0608.560.r2.m	Glyma.02G107900.1	RIBONUCLEASE T2	0	5.01	-5.5424
TR17827 c0_g1	LjSGA_052002.1	Glyma.19G187500.1	UDP-glucosyl transferase 73C (<i>UGT73C</i>)	3.21	3.48	-0.3123
TR62356 c1_g1	chr5.CM0052.810.r2.m	Glyma.02G081000.1	Sterol 3-beta-glucosyltransferase	13.47	31.35	-1.4669
TR62356 c1_g3	chr5.CM0052.810.r2.m	Glyma.10G224000.1	SPHINGOMYELIN SYNTHETASE // SUBFAMILY NOT NAMED	13.29	14.18	-0.3049
TR35590 c0_g1	chr2.CM0124.30.r2.m	Glyma.02G100700.1	STEROL REGULATORY ELEMENT-BINDING PROTEIN	1.18	1.53	-0.3171
TR5549 c0_g1	chr2.CM0124.30.r2.m	Glyma.02G100700.1	STEROL REGULATORY ELEMENT-BINDING PROTEIN	0.42	2.39	-2.5963
TR26623 c0_g1	chr2.CM0177.810.r2.m	Glyma.09G204500.1	transcription factor MYC2	24.75	22.14	-0.0484
TR61787 c0_g1	LjSGA_091313.0.1	Glyma.02G136000.1	ATP-dependent Clp protease adaptor protein ClpS	5.78	16.45	-1.5641
TR32422 c0_g2	chr5.CM1125.790.r2.a	Glyma.10G026000.1	EGL3 (ENHANCER OF GLABRA3), DNA BINDING / TRANSCRIPTION FACTOR	49.62	66.36	-0.6287

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR40782 c0_g1	chr3.CM0136.10.r2.m	Glyma.11G145500.1	AUX/IAA family // B3 DNA binding domain // Auxin response factor	5.23	0	4.8824
TR7528 c0_g1	chr3.CM0136.10.r2.m	Glyma.11G145500.1	AUX/IAA family // B3 DNA binding domain // Auxin response factor	4.6	0	4.4836
TR20842 c1_g1	LjSGA_014979.1	Glyma.18G104100.1	Transferase family	1731.14	1737.77	1.4098
TR19685 c0_g2	LjSGA_045968.2	Glyma.20G180000.1	B3 DNA binding domain // Auxin response factor	6.87	3.12	0.9419
TR14793 c0_g1	LjSGA_111580.1	Glyma.13G082700.1	Thaumatococcus family	3.54	0.6	2.3293
TR14793 c1_g1	LjSGA_111580.1	Glyma.14G163700.1	Thaumatococcus family	4.99	0	3.9305
TR47026 c0_g1	chr1.CM1911.100.r2.m	Glyma.03G122000.1	coumaroylquinate(coumaroylshikimate) 3'-monooxygenase	38.65	16.98	0.9787
TR47026 c0_g2	chr1.CM1911.100.r2.m	Glyma.03G122000.1	coumaroylquinate(coumaroylshikimate) 3'-monooxygenase	182.35	58.95	1.4244
TR38719 c0_g1	LjSGA_024944.1	Glyma.10G210600.1	B3 DNA binding domain // Auxin response factor	1.24	0	4.4836
TR38719 c0_g2	LjSGA_024944.1	Glyma.10G210600.1	B3 DNA binding domain // Auxin response factor	13.28	5.31	1.1199
TR52576 c1_g1	chr5.CM0077.120.r2.m	NA	NA	2.08	1.34	0.5912
TR65913 c1_g1	chr2.CM0191.680.r2.m	NA	NA	1.2	9.4	-2.7781

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log2FC
TR38780 c0_g2	chr5.CM0180.670.r2.m	NA	NA	4.1	8.72	-1.2662
TR55653 c1_g1	LjSGA_022904.1	NA	NA	4.66	2.09	1.1443
TR62356 c0_g1	chr5.CM0052.810.r2.m	NA	NA	13.82	12.13	0.0737
TR42787 c0_g1	chr1.CM0104.1140.r2.m	NA	NA	2.72	2.5	0.0564
TR17090 c0_g1	LjSGA_020776.1.1	NA	NA	2.29	0	4.4836
TR41909 c0_g1	LjSGA_020776.1.1	NA	NA	6.82	0	3.9305
TR31193 c1_g1	chr4.CM0227.690.r2.m	NA	NA	2.89	0	3.9305
TR37424 c0_g1	chr4.CM0227.690.r2.m	NA	NA	2.8	1.11	1.1847
TR13846 c0_g1	chr3.CM0136.10.r2.m	NA	NA	7.76	0	6.1747
TR27061 c0_g1	chr6.LjT09C09.100.r2.a	NA	NA	1.64	1.03	0.5912
TR30767 c0_g1	chr6.LjT09C09.100.r2.a	Glyma.20G207700.1	NA	4.24	0	4.4836
TR40275 c0_g1	chr6.LjT09C09.100.r2.a	Glyma.02G004200.1	NA	3.84	0	3.93057
TR19889 c1_g1	chr2.CM0018.760.r2.m	NA	NA	52.86	99.33	-0.8652
TR25290 c0_g1	chr5.CM0180.690.r2.m	NA	NA	1.11	3.35	-1.6119
TR12068 c0_g1	chr5.CM0052.810.r2.m	NA	NA	0.6	4.36	-2.7781
TR48553 c0_g1	chr5.CM1125.790.r2.a	NA	NA	1.45	3.57	-1.2976
TR46080 c0_g1	LjSGA_040150.1	NA	NA	1.33	3.88	-1.7486
TR46080 c0_g2	LjSGA_040150.1	NA	NA	5.24	5.38	-0.2436

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR39677 c1_g1	LjSGA_092302.1	NA	NA	1.44	1.4	-0.1316
TR28458 c1_g1	LjSGA_054002.1	NA	NA	1.82	0	3.9305
TR51898 c0_g1	chr2.CM0124.30.r2. m	NA	NA	0.46	2.38	-2.4460
TR3728 c0_g1	chr1.CM0476.410.r2. a	NA	NA	2.45	3.5	-0.31712
TR61647 c0_g1	chr5.CM0200.3220.r 2.m	NA	NA	68.47	112.77	-0.9246
TR5553 c0_g1	chr3.CM0452.240.r2. d	NA	NA	2.23	3.14	-0.3171
TR60161 c0_g1	chr1.CM0104.1140.r 2.m	NA	NA	2.64	5.23	-0.8954

Table S16. Yellow Stage Trinity Results: Genes for enzymes annotated to the over-represented GO:000 (Lignin biosynthetic process, FDR=). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR4932 c0_g1	chr1.CM0476.410.r2. a	Glyma.04G227700.1	Quercetin 3-O-methyltransferase (-3-OMT)	5616	1315	1.7378
TR4801 c6_g1	LjSGA_035363.1	Glyma.05G052000.1	Cytochrome P450	57	0	8.6691
TR37863 c1_g1	LjSGA_024849.1	Glyma.07G023700.1	cinnamoyl-CoA reductase (CCR)	24	1	4.0867
TR47015 c0_g1	LjSGA_040039.1	Glyma.07G026300.1	NAD DEPENDENT EPIMERASE/DEHYD RATASE // SUBFAMILY NOT NAMED	239	22	3.0785
TR25023 c0_g1	LjSGA_027787.1	Glyma.07G050600.1	No apical meristem (NAM) protein	69	0	8.9441
TR8127 c0_g1	chr4.CM0387.960.r2. m	Glyma.07G258700.1	beta-glucosidase	118	1	6.3777
TR35194 c0_g1	chr3.CM0452.240.r2. d	Glyma.08G220200.1	Shikimate O-hydroxycinnamoyltransferase (HCT)	8894	1190	2.5452
TR5179 c0_g1	chr1.CM0017.470.r2. a	Glyma.08G362900.1	Transaldolase.	4763	939	1.9860
TR46051 c0_g2	LjSGA_013806.1	Glyma.12G109800.1	caffeoyl-CoA O-methyltransferase (CCoAOMT)	536	10	5.3722
TR30560 c0_g1	chr3.LjT40P18.80.r2. m	Glyma.13G371000.1	Quercetin 3-O-methyltransferase (-3-OMT)	4923.93	244.53	3.9718
TR30560 c0_g2	chr3.LjT40P18.80.r2. m	Glyma.13G371000.1	Shikimate O-hydroxycinnamoyltransferase (HCT)	259.39	14.83	3.7436
TR34012 c0_g1	chr3.LjT45M09.50.r 2.d	Glyma.15G003600.1	Shikimate O-hydroxycinnamoyltransferase (HCT)	80.61	7.17	3.1560

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR49520 c0_g1	chr6.CM0385.240.r2.d	Glyma.15G143600.1	cinnamoyl-CoA reductase (CCR)	2562	398	2.3295
TR36532 c0_g1	chr6.LjT35D18.30.r2.m	Glyma.15G190500.1	CHITINASE	1112	7	6.9328
TR4801 c4_g1	LjSGA_035363.1	Glyma.17G134100.1	Methionine adenosyltransferase (MAT)	19644	3001	2.3539
TR38742 c0_g1	LjSGA_021886.2	Glyma.17G171100.1	Cytochrome P450 CYP2 subfamily	77	0	9.1021
TR61337 c0_g1	chr5.CM0200.3250.r2.m	Glyma.20G128600.1	cinnamyl-alcohol dehydrogenase (CAD)	138.5	20.92	2.3536
TR61337 c0_g3	chr5.CM0200.3250.r2.m	Glyma.20G128600.1	cinnamyl-alcohol dehydrogenase (CAD)	47.87	0	8.4218
TR61337 c0_g2	chr5.CM0200.3250.r2.m	Glyma.20G128600.1	cinnamyl-alcohol dehydrogenase (CAD)	4056.51	588.62	2.4273

Table S17. Red Stage Trinity Results: Genes for enzymes annotated to the over-represented GO:000 (Lignin biosynthetic process, FDR=). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR24303 c0_g1	LjSGA_128268.1	Glyma.01G177900.1	coniferyl-alcohol glucosyltransferase (CAGT)	60	761	-4.0182
TR13393 c0_g1	LjSGA_030253.1	Glyma.03G184000.1	Methionine adenosyltransferase (MAT)	1894	5353	-1.8553
TR1381 c0_g1	chr1.CM0105.600.nc	Glyma.07G048700.1	Quercetin 3-O-methyltransferase (3-OMT)	4	126	-5.2853
TR14130 c0_g1	chr1.CM0063.200.nc	Glyma.08G010500.1	Transferase family	1	29	-5.0304
TR4013 c1_g3	LjSGA_011490.1	Glyma.12G019700.1	Cinnamyl-alcohol dehydrogenase (CAD)	8	172	-4.7586
TR45798 c0_g1	chr1.CM0371.40.nc	Glyma.16G039500.1	Transferase family	21	432	-4.7098

Table S18. Similarly Expressed Trinity Results: Genes for enzymes annotated to GO:0009813 (Lignin process, FDR=1). The first two columns list Trinity contigs expresand gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log2FC
TR4638 c1_g1	chr2.CM0249.1510.r2.m	Glyma.11G069600.1	LACCASE	0.49	0.28	0.5912
TR13756 c0_g1	chr6.CM0084.30.r2.m	Glyma.18G065100.1	LACCASE	8.82	0	10.251
TR2439 c0_g1	chr1.CM2049.30.r2.m	Glyma.18G193400.1	LACCASE	3.37	0.6	2.2824
TR35330 c1_g1	chr1.CM2049.30.r2.m	Glyma.18G193200.1	LACCASE	3.99	0.59	2.5184
TR48670 c0_g2	chr1.CM2049.30.r2.m	Glyma.07G142400.1	LACCASE	2.73	0	5.6687
TR34461 c0_g1	chr1.CM0147.370.r2.m	Glyma.08G359100.1	LACCASE	7.9	0	6.4344
TR34461 c1_g1	chr1.CM0147.370.r2.m	Glyma.18G177400.1	LACCASE	4.28	0	8.2967
TR34461 c2_g1	chr1.CM0147.370.r2.m	Glyma.07G133900.1	LACCASE	12.03	0	5.6687
TR13756 c0_g2	chr6.CM0084.30.r2.m	Glyma.18G065100.1	LACCASE	0	1.77	-6.5269
TR65076 c0_g1	chr6.CM0084.30.r2.m	Glyma.02G224800.1	LACCASE	0.59	6.62	-3.6477
TR62643 c0_g1	chr3.CM0243.390.r2.m	Glyma.07G054200.1	LACCASE	0.59	2.87	-2.2060
TR36532 c0_g1	chr6.LjT35D18.30.r2.m	Glyma.15G190500.1	Methionine adenosyltransferase (<i>MAT</i>)	1483.98	252.36	2.3539
TR45709 c0_g1	chr6.LjT35D18.30.r2.m	Glyma.17G039000.1	Methionine adenosyltransferase (<i>MAT</i>)	112.08	58.48	0.7446
TR64116 c0_g1	LjSGA_030253.1	Glyma.13G141600.1	Methionine adenosyltransferase (<i>MAT</i>)	40.05	69.47	-0.6148
TR13393 c0_g1	LjSGA_030253.1	Glyma.03G184000.1	Methionine adenosyltransferase (<i>MAT</i>)	101.29	317.51	-1.8553

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log2FC
TR24590 c0_g1	LjSGA_140570.1	Glyma.13G141600.1	Methionine adenosyltransferase (<i>MAT</i>)	21.61	58.23	-1.6092
TR13516 c0_g1	LjSGA_035751.3	Glyma.07G089700.1	Cytochrome P450 CYP2 subfamily	25.96	16.48	0.4589
TR15440 c0_g1	LjSGA_035751.3	Glyma.09G186400.1	Cytochrome P450 CYP2 subfamily	6.11	10.04	-0.7962
TR37032 c0_g1	LjSGA_021748.1.1	Glyma.07G023700.1	cinnamoyl-CoA reductase (<i>CCR</i>)	7.23	1.2	2.4086
TR4013 c1_g1	LjSGA_073761.1	Glyma.12G019700.1	Cinnamyl-alcohol dehydrogenase (<i>CAD</i>)	24.28	21.27	0.0141
TR39650 c0_g1	chr2.CM0249.890.r2.m	Glyma.11G072000.1	Peroxidase (<i>POD</i>)	5.78	1.47	1.7657
TR39650 c0_g2	chr2.CM0249.890.r2.m	Glyma.11G072000.1	Peroxidase (<i>POD</i>)	0.52	0.01	5.4510
TR10667 c0_g1	chr4.CM0131.580.r2.m	Glyma.02G052700.1	Peroxidase (<i>POD</i>)	7.12	0	4.4836
TR47026 c0_g1	chr1.CM1911.100.r2.m	Glyma.03G122000.1	coumaroylquininate(coumaroylshikimate) 3'-monooxygenase (<i>C3'H</i>)	38.65	16.98	0.9787
TR47026 c0_g2	chr1.CM1911.100.r2.m	Glyma.03G122000.1	coumaroylquininate(coumaroylshikimate) 3'-monooxygenase (<i>C3'H</i>)	182.35	58.95	1.4244
TR31562 c0_g1	chr1.CM0017.470.r2.a	Glyma.08G362900.1	Transaldolase.	3.1	0	4.4836
	chr1.CM0017.470.r2.a	Glyma.08G362900.1	Transaldolase.	318.62	69.82	1.9860
TR123 c0_g1	chr2.CM0191.680.r2.m	Glyma.04G227700.1	Quercetin 3-O-methyltransferase (<i>3-OMT</i>)	0.27	1.84	-2.7781

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR39944 c0_g1	chr3.CM0574.80.r2.d	Glyma.19G182300.1	Phenylalanine ammonia-lyase (<i>PAL</i>)	1.33	2.63	-1.0588
TR49219 c1_g1	chr3.CM0574.80.r2.d	Glyma.03G181600.1	Phenylalanine ammonia-lyase (<i>PAL</i>)	0	16.69	-7.1991
TR36182 c0_g1	chr2.CM0177.130.r2.d	Glyma.01G020900.1	L-THREONINE 3-DEHYDROGENASE (<i>TDH</i>)	27.5	43.99	-0.8770
TR50128 c0_g2	chr6.CM0055.420.r2.m	Glyma.15G059500.1	L-THREONINE 3-DEHYDROGENASE (<i>TDH</i>)	1.62	3.31	-1.2317
TR50128 c0_g1	LjSGA_070619.1	Glyma.15G059500.1	L-THREONINE 3-DEHYDROGENASE (<i>TDH</i>)	15.98	4.57	1.6040
TR34200 c0_g1	chr6.CM0055.420.r2.m	Glyma.15G059500.1	L-THREONINE 3-DEHYDROGENASE (<i>TDH</i>)	5.23	3.64	0.3340
TR63600 c0_g2	chr6.CM0139.260.r2.m	Glyma.15G106000.1	GLYCOSYL HYDROLASE	2.21	2.27	-0.2523
TR63600 c0_g1	chr6.CM0139.260.r2.m	Glyma.15G106000.1	GLYCOSYL HYDROLASE	1.45	0.87	0.5441
TR53088 c0_g1	chr1.CM0104.480.r2.a	Glyma.16G050200.1	V-type H ⁺ -transporting ATPase subunit C	108.06	64.97	0.5319
TR20842 c1_g1	LjSGA_014979.1	Glyma.18G104100.1	Transferase family	1731.14	1737.77	1.4098
TR40987 c0_g1	chr6.LjT34E09.150.r2.m	Glyma.13G073200.1	elongation factor EF-1 beta subunit	114.07	34.18	1.5444
TR12259 c0_g1	LjSGA_075859.1	Glyma.20G175500.1	No apical meristem (NAM) protein	2.65	6.65	-1.5147

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR21604 c0_g2	chr2.CM0249.600.r2.d	Glyma.01G169200.1	ferulate-5-hydroxylase	268.84	239.65	-0.0414
TR60250 c0_g1	chr6.LjT35D18.30.r2.m	NA		5.93	0.89	2.5390
TR40104 c1_g1	LjSGA_035655.2	NA		7.01	2.59	1.3203
TR13073 c0_g1	chr3.CM0792.180.r2.d	NA		3.01	0	5.6687
TR32183 c1_g1	chr3.CM0792.180.r2.d	NA		3.99	1.27	1.5431
TR54702 c0_g1	chr3.CM0792.180.r2.d	NA		6.21	0	4.8824
TR42774 c0_g1	chr2.CM0249.600.r2.d	NA		5.72	3.81	0.6162
TR13250 c0_g1	chr6.CM0139.240.r2.m	NA		2.3	0	5.6687
TR26810 c0_g1	chr6.CM0139.240.r2.m	NA		9.59	0	5.1945
TR26810 c1_g1	chr6.CM0139.240.r2.m	NA		4.9	0	3.9305
TR35810 c0_g1	chr6.CM0139.240.r2.m	NA		3.96	0.52	2.6856
TR8127 c0_g2	chr6.CM0139.240.r2.m	NA		10.45	0	5.4510
TR34627 c0_g1	chr1.CM0023.30.r2.d	NA		1.62	0.95	0.6162
TR61337 c0_g2	chr5.CM0200.3250.r2.m	Glyma.20G128600.1		11.69	1.97	2.3536
TR61337 c0_g4	chr5.CM0200.3250.r2.m	Glyma.20G128600.1		5.23	4.36	0.0739
TR60726 c0_g1	chr4.CM0042.2100.r2.m	NA		5.45	0	5.19459
TR34461 c3_g1	chr1.CM0147.370.r2.m	NA		6.89	0	4.4836

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR10653 c0_g1	chr3.LjT45M09.50.r2.d	NA		2.36	1.12	0.9283
TR19271 c0_g1	chr3.LjT45M09.50.r2.d	NA		5.23	0	4.8824
TR65913 c1_g1	chr2.CM0191.680.r2.m	NA		1.2	9.4	-2.7781
TR36166 c0_g1	LjSGA_035751.3	NA		0	4.07	-4.2662
TR5553 c0_g1	chr3.CM0452.240.r2.d	NA		2.23	3.14	-0.3171
TR37416 c0_g1	chr6.CM0084.30.r2.m	NA		0	6.27	-5.2282
TR3728 c0_g1	chr1.CM0476.410.r2.a	NA		2.45	3.5	-0.3171
TR45016 c0_g1	LjSGA_073761.1	NA		1.36	3.22	-1.3166

Table S19. Tophat Results: Significantly enriched GO terms corresponding to the red stage of a *L. sessilifolius* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0034654nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process	21	5.00E-22
GO:0010035response to inorganic substance	36	1.80E-14
GO:0055114oxidation reduction	41	1.10E-12
GO:0034641cellular nitrogen compound metabolic process	52	2.20E-12
GO:0046686response to cadmium ion	20	4.00E-12
GO:0010038response to metal ion	29	4.60E-12
GO:0055085transmembrane transport	42	5.40E-12
GO:0018130heterocycle biosynthetic process	25	5.30E-11
GO:0019438aromatic compound biosynthetic process	36	1.90E-09
GO:0006725cellular aromatic compound metabolic process	47	3.30E-09
GO:0009611response to wounding	28	4.80E-08
GO:0009718anthocyanin biosynthetic process	12	5.60E-08
GO:0044271cellular nitrogen compound biosynthetic process	36	6.50E-08
GO:0031540regulation of anthocyanin biosynthetic process	8	4.50E-07
GO:0046283anthocyanin metabolic process	12	7.70E-07
GO:0031537regulation of anthocyanin metabolic process	8	6.70E-06
GO:0034655nucleobase, nucleoside, nucleotide and nucleic acid catabolic process	6	9.10E-06
GO:0034656nucleobase, nucleoside and nucleotide catabolic process	6	9.10E-06
GO:0042221response to chemical stimulus	104	1.10E-05
GO:0046483heterocycle metabolic process	40	1.40E-05
GO:0009962regulation of flavonoid biosynthetic process	9	2.00E-05
GO:0048610reproductive cellular process	9	5.90E-05
GO:0048868pollen tube development	12	6.50E-05
GO:0009605response to external stimulus	42	6.70E-05
GO:0009753response to jasmonic acid stimulus	23	0.00012
GO:0009651response to salt stress	27	0.00015
GO:0006970response to osmotic stress	31	0.00019
GO:0046148pigment biosynthetic process	16	0.00059
GO:0042440pigment metabolic process	17	0.00085
GO:0010039response to iron ion	6	0.001
GO:0031668cellular response to extracellular stimulus	13	0.0016
GO:0005984disaccharide metabolic process	9	0.0023
GO:0016036cellular response to phosphate starvation	8	0.0025
GO:0006811ion transport	35	0.0033

Enriched GO Term: Description	Number of associated genes	FDR
GO:0006814sodium ion transport	6	0.0033
GO:0035295tube development	12	0.0035
GO:0016137glycoside metabolic process	11	0.0035
GO:0042398cellular amino acid derivative biosynthetic process	23	0.0037
GO:0009813flavonoid biosynthetic process	15	0.0041
GO:0043455regulation of secondary metabolic process	9	0.0042
GO:0031669cellular response to nutrient levels	11	0.0043
GO:0009860pollen tube growth	9	0.0043
GO:0009719response to endogenous stimulus	57	0.0051
GO:0006820anion transport	13	0.0058
GO:0060918auxin transport	10	0.0058
GO:0009812flavonoid metabolic process	15	0.0066
GO:0046700heterocycle catabolic process	8	0.0085
GO:0015698inorganic anion transport	9	0.0087
GO:0010224response to UV-B	9	0.01
GO:0010033response to organic substance	64	0.011
GO:0006595polyamine metabolic process	5	0.012
GO:0042742defense response to bacterium	17	0.012
GO:0009311oligosaccharide metabolic process	9	0.013
GO:0009310amine catabolic process	9	0.013
GO:0006979response to oxidative stress	25	0.013
GO:0009737response to abscisic acid stimulus	26	0.013
GO:0009914hormone transport	10	0.013
GO:0009932cell tip growth	10	0.018
GO:0009267cellular response to starvation	9	0.019
GO:0009926auxin polar transport	9	0.019
GO:0019439aromatic compound catabolic process	5	0.023
GO:0009856pollination	12	0.027
GO:0005991trehalose metabolic process	5	0.028
GO:0009699phenylpropanoid biosynthetic process	17	0.032
GO:0006575cellular amino acid derivative metabolic process	25	0.035
GO:0009991response to extracellular stimulus	14	0.035
GO:0009116nucleoside metabolic process	8	0.035
GO:0009725response to hormone stimulus	50	0.035
GO:0048588developmental cell growth	10	0.038
GO:0009063cellular amino acid catabolic process	8	0.039

Enriched GO Term: Description	Number of associated genes	FDR
GO:0032549ribonucleoside binding	32	2.40E-34
GO:0016210naringenin-chalcone synthase activity	8	6.70E-06
GO:0016791phosphatase activity	20	0.0073
GO:0022891substrate-specific transmembrane transporter activity	43	0.021
GO:0042578phosphoric ester hydrolase activity	21	0.021
GO:0015103inorganic anion transmembrane transporter activity	9	0.021
GO:0008509anion transmembrane transporter activity	11	0.049
GO:0005773vacuole	55	5.40E-23
GO:0005911cell-cell junction	24	6.70E-23
GO:0030054cell junction	24	4.50E-22
GO:0005774vacuolar membrane	28	1.40E-09
GO:0044437vacuolar part	28	2.10E-09
GO:0009941chloroplast envelope	17	3.40E-06
GO:0005886plasma membrane	68	4.50E-06
GO:0009705plant-type vacuole membrane	12	1.00E-05
GO:0000325plant-type vacuole	15	1.20E-05
GO:0048046apoplast	18	7.40E-05
GO:0009526plastid envelope	17	9.00E-05
GO:0016020membrane	158	0.00033
GO:0044459plasma membrane part	27	0.0004
GO:0009570chloroplast stroma	13	0.015
GO:0005777peroxisome	10	0.016
GO:0042579microbody	10	0.016
GO:0005802trans-Golgi network	7	0.017
GO:0005618cell wall	25	0.048

Table S20. Tophat Results: Significantly enriched GO terms corresponding to the yellow stage of a *L. sessilifolius* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0055114oxidation reduction	43	2.10E-20
GO:0034654nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process	15	1.80E-16
GO:0034641cellular nitrogen compound metabolic process	44	4.80E-14
GO:0018130heterocycle biosynthetic process	24	8.80E-14
GO:0044271cellular nitrogen compound biosynthetic process	34	3.10E-11
GO:0046686response to cadmium ion	15	8.40E-10
GO:0016053organic acid biosynthetic process	34	4.00E-09
GO:0046394carboxylic acid biosynthetic process	34	4.00E-09
GO:0019438aromatic compound biosynthetic process	28	8.00E-09
GO:0006725cellular aromatic compound metabolic process	35	4.70E-08
GO:0043648dicarboxylic acid metabolic process	14	2.80E-07
GO:0043436oxoacid metabolic process	46	3.70E-07
GO:0006082organic acid metabolic process	46	3.70E-07
GO:0019752carboxylic acid metabolic process	46	3.70E-07
GO:0010035response to inorganic substance	20	4.20E-07
GO:0042180cellular ketone metabolic process	46	5.10E-07
GO:0010038response to metal ion	17	1.90E-06
GO:0046483heterocycle metabolic process	32	3.20E-06
GO:0009073aromatic amino acid family biosynthetic process	10	3.20E-06
GO:0046417chorismate metabolic process	10	3.20E-06
GO:0006633fatty acid biosynthetic process	17	3.30E-06
GO:0009072aromatic amino acid family metabolic process	11	8.40E-06
GO:0006631fatty acid metabolic process	20	1.30E-05
GO:0008610lipid biosynthetic process	28	1.60E-05
GO:0032787monocarboxylic acid metabolic process	28	2.20E-05
GO:0030244cellulose biosynthetic process	8	0.00015
GO:0010025wax biosynthetic process	7	0.00018
GO:0010102lateral root morphogenesis	9	0.00026

Enriched GO Term: Description	Number of associated genes	FDR
GO:0010101post-embryonic root morphogenesis	9	0.00026
GO:0009886post-embryonic morphogenesis	11	0.00042
GO:0010311lateral root formation	6	0.00042
GO:0010166wax metabolic process	7	0.00042
GO:0019439aromatic compound catabolic process	6	0.00046
GO:0006073cellular glucan metabolic process	14	0.00059
GO:0030243cellulose metabolic process	8	0.00062
GO:0006629lipid metabolic process	36	0.00071
GO:0044255cellular lipid metabolic process	31	0.00079
GO:0044042glucan metabolic process	14	0.0011
GO:0009309amine biosynthetic process	15	0.0015
GO:0008652cellular amino acid biosynthetic process	14	0.0017
GO:0044262cellular carbohydrate metabolic process	26	0.0026
GO:0009250glucan biosynthetic process	10	0.0042
GO:0008299isoprenoid biosynthetic process	10	0.0042
GO:0000162tryptophan biosynthetic process	5	0.0046
GO:0009664plant-type cell wall organization	8	0.0046
GO:0046219indolalkylamine biosynthetic process	5	0.0046
GO:0009416response to light stimulus	27	0.0046
GO:0044264cellular polysaccharide metabolic process	15	0.0057
GO:0009314response to radiation	27	0.007
GO:0016109tetraterpenoid biosynthetic process	5	0.0095
GO:0016117carotenoid biosynthetic process	5	0.0095
GO:0042401cellular biogenic amine biosynthetic process	6	0.012
GO:0006586indolalkylamine metabolic process	5	0.014
GO:0016114terpenoid biosynthetic process	8	0.014
GO:0006568tryptophan metabolic process	5	0.014
GO:0010015root morphogenesis	12	0.014
GO:0009555pollen development	10	0.014
GO:0042221response to chemical stimulus	63	0.014

Enriched GO Term: Description	Number of associated genes	FDR
GO:0051273beta-glucan metabolic process	5	0.014
GO:0051274beta-glucan biosynthetic process	5	0.014
GO:0044106cellular amine metabolic process	20	0.015
GO:0010218response to far red light	6	0.016
GO:0048527lateral root development	9	0.019
GO:0009628response to abiotic stimulus	48	0.02
GO:0005976polysaccharide metabolic process	16	0.02
GO:0006720isoprenoid metabolic process	10	0.021
GO:0033692cellular polysaccharide biosynthetic process	11	0.023
GO:0048528post-embryonic root development	9	0.024
GO:0042435indole derivative biosynthetic process	5	0.029
GO:0034285response to disaccharide stimulus	5	0.029
GO:0006520cellular amino acid metabolic process	19	0.032
GO:0000271polysaccharide biosynthetic process	11	0.036
GO:0016108tetraterpenoid metabolic process	5	0.037
GO:0016116carotenoid metabolic process	5	0.037
GO:0042434indole derivative metabolic process	5	0.039
GO:0042430indole and derivative metabolic process	5	0.039
GO:0030258lipid modification	6	0.039
GO:0019321pentose metabolic process	5	0.048
GO:0048229gametophyte development	13	0.048
GO:0034440lipid oxidation	5	0.049
GO:0032549ribonucleoside binding	27	3.90E-32
GO:0051213dioxygenase activity	10	5.30E-05
GO:0016835carbon-oxygen lyase activity	11	0.0075
GO:0005507copper ion binding	11	0.0077
GO:0016705oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	13	0.025
GO:0043167ion binding	81	0.025
GO:0016760cellulose synthase (UDP-forming) activity	5	0.025

Enriched GO Term: Description	Number of associated genes	FDR
GO:0044434chloroplast part	69	1.80E-35
GO:0009941chloroplast envelope	41	5.80E-35
GO:0044435plastid part	70	4.20E-33
GO:0009526plastid envelope	43	5.30E-33
GO:0009570chloroplast stroma	43	1.10E-32
GO:0005911cell-cell junction	27	4.30E-32
GO:0030054cell junction	27	4.90E-31
GO:0009532plastid stroma	43	9.40E-30
GO:0005886plasma membrane	70	3.20E-15
GO:0031967organelle envelope	44	2.00E-14
GO:0031975envelope	44	8.00E-14
GO:0009507chloroplast	81	2.40E-12
GO:0009536plastid	83	4.50E-11
GO:0009534chloroplast thylakoid	26	8.10E-11
GO:0031976plastid thylakoid	26	3.80E-10
GO:0031984organelle subcompartment	26	5.50E-10
GO:0044459plasma membrane part	30	2.00E-09
GO:0048046apoplast	20	3.60E-09
GO:0009579thylakoid	28	5.00E-07
GO:0016020membrane	121	9.10E-07
GO:0009535chloroplast thylakoid membrane	18	4.00E-06
GO:0055035plastid thylakoid membrane	18	1.10E-05
GO:0044446intracellular organelle part	95	1.20E-05
GO:0044422organelle part	95	1.20E-05
GO:0005576extracellular region	28	2.00E-05
GO:0042651thylakoid membrane	18	4.00E-05
GO:0044436thylakoid part	19	4.70E-05
GO:0005618cell wall	23	0.00072
GO:0005802trans-Golgi network	7	0.0012
GO:0005773vacuole	17	0.0019

Enriched GO Term: Description	Number of associated genes	FDR
GO:0030312external encapsulating structure	23	0.002
GO:0034357photosynthetic membrane	18	0.0025
GO:0044444cytoplasmic part	135	0.0026
GO:0010287plastoglobule	6	0.0081
GO:0022626cytosolic ribosome	9	0.02

Table S21. Red Stage TopHat Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR = 0.0066) . The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0021.2820.r2.m	Glyma.11G027700.1	Ribonuclease T(2) / Ribonuclease T2	11	303	-4.87472
chr2.CM0608.560.r2.m	Glyma.01G048200.1	NARINGENIN,2-OXOGLUTARATE 3-DIOXYGENASE	4	53	-3.79248
chr2.CM0018.1190.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	16	-7.06546
chr2.CM0018.760.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	1	103	-6.61962
chr1.LjT39K18.30.r2.m	Glyma.04G147700.1	Naringenin-chalcone synthase / Flavonone synthase	78	686	-3.24183
chr2.CM0018.730.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	83	-9.43178
chr2.CM0018.1300.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	43	-8.48495
chr5.CM0077.110.r2.m	Glyma.02G158700.1	Naringenin-chalcone synthase / Flavonone synthase	6	108	-4.248
chr1.CM0284.240.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	7	316	-5.57781
chr1.CM0104.1140.r2.m	Glyma.08G109300.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	27	108	-2.10201
chr2.CM0018.1150.r2.m	Glyma.01G228700.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	0	20	-7.38524
chr4.CM0119.240.r2.m	Glyma.01G166200.1	Thromboxane-A synthase / Thromboxane synthetase	11	82	-2.99064
chr1.CM0133.560.r2.m	Glyma.04G052100.1	Heme oxygenase (biliverdin-producing) / Heme oxygenase (decyclizing)	0	31	-8.01444
chr5.CM0077.210.r2.m	Glyma.02G158700.1	Naringenin-chalcone synthase / Flavonone synthase	1	27	-4.69274
chr2.CM0018.1200.r2.m	Glyma.01G228700.1	LEUCOANTHOCYANIDIN DIOXYGENASE-RELATED	1	83	-6.30857

***Flavonoid biosynthetic process was not found to be enriched at the yellow stage of floral colour change in *L. sessilifolius*, given the TopHat results.**

***Lignin biosynthetic process was not found to be expressed at the yellow or red stages of floral colour change in *L. sessilifolius*, given the TopHat results.**

Table S22. STAR Results: Significantly enriched GO terms corresponding to the yellow stage of a *L. sessilifolius* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0044281small molecule metabolic process	105	4.90E-11
GO:0008610lipid biosynthetic process	40	1.10E-08
GO:0044283small molecule biosynthetic process	57	1.30E-08
GO:0006629lipid metabolic process	57	1.30E-08
GO:0046394carboxylic acid biosynthetic process	41	4.80E-08
GO:0043436oxoacid metabolic process	58	4.80E-08
GO:0016053organic acid biosynthetic process	41	4.80E-08
GO:0006082organic acid metabolic process	58	4.80E-08
GO:0019752carboxylic acid metabolic process	58	4.80E-08
GO:0042180cellular ketone metabolic process	59	5.60E-08
GO:0044255cellular lipid metabolic process	41	1.50E-06
GO:0006633fatty acid biosynthetic process	22	9.80E-06
GO:0032787monocarboxylic acid metabolic process	34	1.30E-05
GO:0006631fatty acid metabolic process	25	2.00E-05
GO:0050896response to stimulus	144	3.20E-05
GO:0055114oxidation reduction	63	3.20E-05
GO:0046417chorismate metabolic process	11	3.90E-05
GO:0009073aromatic amino acid family biosynthetic process	11	3.90E-05
GO:0043648dicarboxylic acid metabolic process	15	4.30E-05
GO:0009628response to abiotic stimulus	70	6.20E-05
GO:0042221response to chemical stimulus	93	7.70E-05
GO:0008152metabolic process	288	0.00014
GO:0009072aromatic amino acid family metabolic process	12	0.00037
GO:0008299isoprenoid biosynthetic process	14	0.00046
GO:0010025wax biosynthetic process	8	0.00047
GO:0010166wax metabolic process	8	0.00047
GO:0009058biosynthetic process	135	0.00065

Enriched GO Term: Description	Number of associated genes	FDR
GO:0044262cellular carbohydrate metabolic process	38	0.0008
GO:0006720isoprenoid metabolic process	14	0.0021
GO:0006950response to stress	90	0.0021
GO:0044264cellular polysaccharide metabolic process	17	0.0025
GO:0044249cellular biosynthetic process	128	0.003
GO:0005976polysaccharide metabolic process	18	0.0041
GO:0009416response to light stimulus	33	0.006
GO:0030244cellulose biosynthetic process	7	0.006
GO:0016114terpenoid biosynthetic process	10	0.0063
GO:0006519cellular amino acid and derivative metabolic process	36	0.0063
GO:0016126sterol biosynthetic process	5	0.0063
GO:0033692cellular polysaccharide biosynthetic process	12	0.0063
GO:0005975carbohydrate metabolic process	47	0.0063
GO:0006073cellular glucan metabolic process	14	0.0063
GO:0009314response to radiation	33	0.0068
GO:0044042glucan metabolic process	14	0.0071
GO:0000271polysaccharide biosynthetic process	12	0.0071
GO:0030243cellulose metabolic process	7	0.0071
GO:0016109tetraterpenoid biosynthetic process	6	0.012
GO:0009607response to biotic stimulus	40	0.012
GO:0016117carotenoid biosynthetic process	6	0.012
GO:0046148pigment biosynthetic process	13	0.012
GO:0010038response to metal ion	28	0.013
GO:0009657plastid organization	13	0.013
GO:0006520cellular amino acid metabolic process	24	0.013
GO:0010218response to far red light	7	0.013
GO:0010102lateral root morphogenesis	9	0.014
GO:0010101post-embryonic root morphogenesis	9	0.014
GO:0044237cellular metabolic process	220	0.014
GO:0071669plant-type cell wall organization or biogenesis	12	0.014
GO:0051707response to other organism	38	0.014
GO:0042214terpene metabolic process	7	0.014
GO:0008652cellular amino acid biosynthetic process	15	0.014
GO:0006066alcohol metabolic process	23	0.015
GO:0019748secondary metabolic process	24	0.015
GO:0034641cellular nitrogen compound metabolic process	58	0.015

Enriched GO Term: Description	Number of associated genes	FDR
GO:0009309amine biosynthetic process	16	0.015
GO:0007275multicellular organismal development	74	0.015
GO:0009250glucan biosynthetic process	9	0.015
GO:0009719response to endogenous stimulus	46	0.015
GO:0044106cellular amine metabolic process	25	0.015
GO:0010015root morphogenesis	15	0.015
GO:0016108tetraterpenoid metabolic process	6	0.015
GO:0016116carotenoid metabolic process	6	0.015
GO:0032502developmental process	78	0.015
GO:0010035response to inorganic substance	32	0.016
GO:0044238primary metabolic process	227	0.016
GO:0000162tryptophan biosynthetic process	5	0.016
GO:0010214seed coat development	5	0.016
GO:0046219indolalkylamine biosynthetic process	5	0.016
GO:0007017microtubule-based process	12	0.016
GO:0032501multicellular organismal process	76	0.016
GO:0006721terpenoid metabolic process	10	0.017
GO:0046246terpene biosynthetic process	6	0.017
GO:0042335cuticle development	5	0.019
GO:0051704multi-organism process	44	0.02
GO:0019438aromatic compound biosynthetic process	38	0.021
GO:0044271cellular nitrogen compound biosynthetic process	43	0.022
GO:0009834secondary cell wall biogenesis	5	0.024
GO:0046686response to cadmium ion	22	0.027
GO:0010033response to organic substance	53	0.029
GO:0016125sterol metabolic process	5	0.029
GO:0048528post-embryonic root development	10	0.029
GO:0009110vitamin biosynthetic process	7	0.031
GO:0009308amine metabolic process	26	0.032
GO:0006586indolalkylamine metabolic process	5	0.032
GO:0051273beta-glucan metabolic process	5	0.032
GO:0051274beta-glucan biosynthetic process	5	0.032
GO:0042440pigment metabolic process	13	0.032
GO:0006568tryptophan metabolic process	5	0.032
GO:0071704organic substance metabolic process	67	0.033
GO:0006766vitamin metabolic process	7	0.033

Enriched GO Term: Description	Number of associated genes	FDR
GO:0071554cell wall organization or biogenesis	18	0.034
GO:0009886post-embryonic morphogenesis	12	0.039
GO:0016051carbohydrate biosynthetic process	16	0.041
GO:0009611response to wounding	15	0.047
GO:0003824catalytic activity	281	1.20E-11
GO:0016491oxidoreductase activity	70	6.70E-05
GO:0005506iron ion binding	27	0.019
GO:0004506squalene monooxygenase activity	5	0.019
GO:0004497monooxygenase activity	11	0.025
GO:0016740transferase activity	100	0.033
GO:0046527glucosyltransferase activity	9	0.042
GO:0016760cellulose synthase (UDP-forming) activity	5	0.047
GO:0016759cellulose synthase activity	5	0.047
GO:0044435plastid part	91	1.20E-19
GO:0044434chloroplast part	90	1.20E-19
GO:0009507chloroplast	113	6.70E-17
GO:0009536plastid	115	8.30E-16
GO:0009526plastid envelope	56	8.90E-15
GO:0009570chloroplast stroma	57	8.90E-15
GO:0009532plastid stroma	57	2.30E-14
GO:0009941chloroplast envelope	53	1.20E-13
GO:0031967organelle envelope	57	1.90E-10
GO:0031975envelope	57	2.00E-10
GO:0044464cell part	311	2.00E-09
GO:0005623cell	311	2.00E-09
GO:0044444cytoplasmic part	197	3.00E-09
GO:0043229intracellular organelle	223	3.00E-08
GO:0043226organelle	223	3.20E-08
GO:0005737cytoplasm	204	3.20E-08
GO:0043231intracellular membrane-bounded organelle	212	3.60E-08
GO:0043227membrane-bounded organelle	212	4.50E-08
GO:0044446intracellular organelle part	126	7.30E-08
GO:0044422organelle part	126	9.90E-08
GO:0009579thylakoid	36	7.30E-07

Enriched GO Term: Description	Number of associated genes	FDR
GO:0044424intracellular part	242	8.90E-07
GO:0005622intracellular	243	4.10E-06
GO:0005618cell wall	39	6.30E-06
GO:0030312external encapsulating structure	39	7.20E-06
GO:0031976plastid thylakoid	29	9.80E-06
GO:0009534chloroplast thylakoid	29	9.80E-06
GO:0031984organelle subcompartment	29	1.00E-05
GO:0005576extracellular region	36	6.00E-05
GO:0048046apoplast	26	0.0003
GO:0044436thylakoid part	23	0.00082
GO:0016020membrane	154	0.0015
GO:0005886plasma membrane	83	0.0059
GO:0031977thylakoid lumen	8	0.0061
GO:0009535chloroplast thylakoid membrane	19	0.0061
GO:0055035plastid thylakoid membrane	19	0.0062
GO:0010287plastoglobule	7	0.01
GO:0042651thylakoid membrane	19	0.011
GO:0034357photosynthetic membrane	19	0.02

Table S23. STAR Results: Significantly enriched GO terms corresponding to the red stage of a *L. sessilifolius* flower.

Enriched GO Term: Description	Number of associated genes	FDR
GO:0009605response to external stimulus	60	1.60E-09
GO:0042221response to chemical stimulus	143	7.20E-08
GO:0009611response to wounding	34	4.90E-07
GO:0050896response to stimulus	209	9.70E-07
GO:0006810transport	123	9.70E-07
GO:0051234establishment of localization	123	1.00E-06
GO:0051179localization	126	1.60E-06
GO:0009753response to jasmonic acid stimulus	30	1.60E-06
GO:0010033response to organic substance	94	2.20E-06
GO:0009714chalcone metabolic process	11	2.70E-06
GO:0042181ketone biosynthetic process	11	2.70E-06
GO:0009715chalcone biosynthetic process	11	2.70E-06
GO:0009719response to endogenous stimulus	79	3.40E-06
GO:0031540regulation of anthocyanin biosynthetic process	11	2.50E-05
GO:0009725response to hormone stimulus	71	2.90E-05
GO:0009718anthocyanin biosynthetic process	14	3.30E-05
GO:0009962regulation of flavonoid biosynthetic process	12	4.00E-05
GO:0055085transmembrane transport	55	4.40E-05
GO:0006950response to stress	133	5.00E-05
GO:0009813flavonoid biosynthetic process	18	5.10E-05
GO:0043455regulation of secondary metabolic process	14	7.30E-05
GO:0009991response to extracellular stimulus	23	0.00011
GO:0031537regulation of anthocyanin metabolic process	11	0.00025
GO:0046283anthocyanin metabolic process	14	0.00047
GO:0006979response to oxidative stress	35	0.00054
GO:0009699phenylpropanoid biosynthetic process	20	0.00059
GO:0031668cellular response to extracellular stimulus	21	0.00059
GO:0009812flavonoid metabolic process	18	0.00059
GO:0042398cellular amino acid derivative biosynthetic process	27	0.00059

Enriched GO Term: Description	Number of associated genes	FDR
GO:0071496cellular response to external stimulus	21	0.00059
GO:0016137glycoside metabolic process	16	0.00088
GO:0005975carbohydrate metabolic process	67	0.0012
GO:0060918auxin transport	13	0.0012
GO:0009914hormone transport	13	0.0014
GO:0009056catabolic process	57	0.0014
GO:0009926auxin polar transport	12	0.0015
GO:0009628response to abiotic stimulus	87	0.0018
GO:0009629response to gravity	13	0.0018
GO:0031667response to nutrient levels	18	0.0018
GO:0044248cellular catabolic process	52	0.0019
GO:0046148pigment biosynthetic process	18	0.0028
GO:0016139glycoside catabolic process	7	0.0033
GO:0016036cellular response to phosphate starvation	12	0.0033
GO:0010224response to UV-B	14	0.0034
GO:0006811ion transport	41	0.0037
GO:0044283small molecule biosynthetic process	54	0.0038
GO:0019748secondary metabolic process	34	0.0038
GO:0006527arginine catabolic process	5	0.0051
GO:0051704multi-organism process	63	0.0074
GO:0042440pigment metabolic process	19	0.0082
GO:0042594response to starvation	15	0.0083
GO:0031669cellular response to nutrient levels	16	0.0093
GO:0006814sodium ion transport	7	0.0098
GO:0009737response to abscisic acid stimulus	33	0.01
GO:0009311oligosaccharide metabolic process	11	0.01
GO:0009733response to auxin stimulus	25	0.012
GO:0005984disaccharide metabolic process	10	0.012
GO:0009698phenylpropanoid metabolic process	21	0.013
GO:0009607response to biotic stimulus	53	0.016

Enriched GO Term: Description	Number of associated genes	FDR
GO:0006575cellular amino acid derivative metabolic process	29	0.016
GO:0046395carboxylic acid catabolic process	13	0.016
GO:0016054organic acid catabolic process	13	0.016
GO:0044282small molecule catabolic process	28	0.017
GO:0009267cellular response to starvation	14	0.018
GO:0009411response to UV	15	0.018
GO:0016052carbohydrate catabolic process	17	0.019
GO:0008216spermidine metabolic process	5	0.024
GO:0008295spermidine biosynthetic process	5	0.024
GO:0006595polyamine metabolic process	6	0.024
GO:0051707response to other organism	50	0.024
GO:0030001metal ion transport	23	0.024
GO:0005991trehalose metabolic process	6	0.024
GO:0006970response to osmotic stress	35	0.026
GO:0010817regulation of hormone levels	19	0.026
GO:0009065glutamine family amino acid catabolic process	5	0.031
GO:0009310amine catabolic process	10	0.032
GO:0044275cellular carbohydrate catabolic process	15	0.039
GO:0006820anion transport	15	0.039
GO:0015698inorganic anion transport	10	0.041
GO:0009414response to water deprivation	22	0.045
GO:0010035response to inorganic substance	41	0.047
GO:0030036actin cytoskeleton organization	11	0.047
GO:0005215transporter activity	79	2.20E-07
GO:0022857transmembrane transporter activity	69	2.20E-07
GO:0022891substrate-specific transmembrane transporter activity	59	1.40E-06
GO:0022892substrate-specific transporter activity	62	1.40E-06
GO:0016210naringenin-chalcone synthase activity	11	1.90E-06
GO:0016798hydrolase activity, acting on glycosyl bonds	40	2.10E-05
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	38	2.70E-05

Enriched GO Term: Description	Number of associated genes	FDR
GO:0015291secondary active transmembrane transporter activity	26	3.20E-05
GO:0016791phosphatase activity	24	0.00012
GO:0042578phosphoric ester hydrolase activity	27	0.00015
GO:0016787hydrolase activity	152	0.00017
GO:0022804active transmembrane transporter activity	36	0.00021
GO:0003779actin binding	12	0.00066
GO:0015144carbohydrate transmembrane transporter activity	11	0.00073
GO:0003824catalytic activity	346	0.00073
GO:0015075ion transmembrane transporter activity	37	0.0041
GO:0015103inorganic anion transmembrane transporter activity	11	0.018
GO:0004650polygalacturonase activity	10	0.024
GO:0016746transferase activity, transferring acyl groups	26	0.024
GO:0008509anion transmembrane transporter activity	14	0.029
GO:0015294solute:cation symporter activity	7	0.031
GO:0015295solute:hydrogen symporter activity	6	0.031
GO:0008092cytoskeletal protein binding	15	0.031
GO:0008234cysteine-type peptidase activity	12	0.031
GO:0008324cation transmembrane transporter activity	26	0.031
GO:0051119sugar transmembrane transporter activity	7	0.031
GO:0046943carboxylic acid transmembrane transporter activity	10	0.031
GO:0015297antiporter activity	13	0.035
GO:0005342organic acid transmembrane transporter activity	10	0.041
GO:0015293symporter activity	7	0.041
GO:0005773vacuole	77	6.50E-07
GO:0000325plant-type vacuole	20	0.00014
GO:0016020membrane	229	0.00017
GO:0009705plant-type vacuole membrane	17	0.00075
GO:0044437vacuolar part	41	0.022
GO:0005774vacuolar membrane	41	0.022
GO:0044464cell part	397	0.023

Table S24. Yellow Stage STAR Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=1). The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr5.CM1077.590.r2.m	Glyma.03G187000.1	Trans-zeatin O-beta-D-glucosyltransferase / Zeatin O-glucosyltransferase	11.82	0	6.526304
chr3.LjT10E18.60.r2.m	Glyma.10G210600.1	B3 DNA binding domain (B3) // Auxin response factor (Auxin_resp)	43	6	2.664685
chr1.CM0410.460.r2.m	Glyma.18G268100.1	Isoflavone-7-O-beta-glucoside 6"-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	24.83	0	7.577035
chr3.CM0106.330.r2.m	Glyma.13G344700.1	B-BOX TYPE ZINC FINGER-CONTAINING PROTEIN-RELATED	157	18	2.963609
chr1.CM0476.410.r2.a	Glyma.04G227700.1	FLAVONE 3'-O-METHYLTRANSFERASE 1	495	98	2.182379

Table S25. Red Stage STAR Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR = 0.00059) . The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr2.CM0608.560.r2.m	Glyma.01G048200.1	Ribonuclease T(2) / Ribonuclease T2	11	128	-3.67758
chr4.CM0119.240.r2.m	Glyma.01G166200.1	NARINGENIN,2-OXOGLUTARATE 3-DIOXYGENASE	155.05	921.11	-2.72273
chr3.LjB14006.120.r2.a	Glyma.01G166200.1	NARINGENIN,2-OXOGLUTARATE 3-DIOXYGENASE	2.95	51.89	-4.20965
chr3.CM0590.770.r2.d	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	39.5	-8.40463
chr1.CM0284.250.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	48.1	-8.66696
chr3.CM0590.840.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	39.5	-8.40463
chr2.CM0018.1190.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	8.81	182.32	-4.4707
chr2.CM0018.760.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	2.85	307.57	-6.77327
chr2.CM0018.730.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	8.15	367.98	-5.65332
chr2.CM0018.1300.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	0	72.45	-9.25074
chr1.CM0284.240.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	11.42	767.23	-6.25955
chr2.CM0018.1150.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	5.77	256.13	-5.53724
chr2.CM0018.1200.r2.m	Glyma.01G228700.1	Naringenin-chalcone synthase / Flavonone synthase	9.38	468.71	-5.83578
chr5.CM0077.110.r2.m	Glyma.02G158700.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	20	321	-4.14841
chr1.CM0133.560.r2.m	Glyma.04G052100.1	Thromboxane-A synthase / Thromboxane synthetase	0	41	-8.44015
chr1.LjT39K18.30.r2.m	Glyma.04G147700.1	Heme oxygenase (biliverdin-producing) / Heme oxygenase (decyclizing)	115	1030	-3.31429

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr1.CM0104.1140.r2.m	Glyma.08G109300.1	Naringenin-chalcone synthase / Flavonone synthase	70	354	-2.48893
chr2.CM0021.2820.r2.m	Glyma.11G027700.1	LEUCOANTHOCYANIDIN DIOXYGENASE-RELATED	14	833	-6.03436

Table S26. Yellow Stage STAR Results: Genes for enzymes annotated to GO:0009808 (Lignin metabolic process, FDR=0.92). The first column lists differentially expressed (FDR<0.005) assembled RNA-seq reads that have mapped to the *Lotus japonicus* genome. The second column is their soybean (*Glycine max*) BLAST homolog, followed by expression at two stages of flower development.

<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	<i>G. max</i> description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
chr6.CM0385.240.r2.d	Glyma.09G038500.1	CHITINASE-LIKE PROTEIN 1	159	16	3.15060
chr6.CM0084.30.r2.m	Glyma.02G224800.1	LACCASE-16	51	3	3.88250
chr4.CM0387.960.r2.m	Glyma.07G258700.1	BETA-GLUCOSIDASE 45-RELATED	16	0	6.93742
chr2.CM0249.1520.r2.m	Glyma.01G173500.1	LACCASE-11	46	1	5.21342
chr1.CM0476.410.r2.a	Glyma.04G227700.1	FLAVONE 3'-O-METHYLTRANSFERASE 1	495	98	2.18237

*Lignin biosynthetic process was not found to be expressed at the red stage of floral colour change in *L. sessilifolius*, given the STAR results.

Table S27. Trinity Results: Significantly enriched GO terms corresponding to the yellow stage of a *L. sessilifolius* flower.

Enriched GO category	Number of associated genes	FDR
GO:0034641cellular nitrogen compound metabolic process	80	4.90E-07
GO:0009791post-embryonic development	138	4.90E-07
GO:0009628response to abiotic stimulus	183	4.90E-07
GO:0043436oxoacid metabolic process	119	2.30E-06
GO:0006082organic acid metabolic process	119	2.30E-06
GO:0019752carboxylic acid metabolic process	119	2.30E-06
GO:0042180cellular ketone metabolic process	120	2.40E-06
GO:0006629lipid metabolic process	111	1.70E-05
GO:0009416response to light stimulus	87	2.80E-05
GO:0048610reproductive cellular process	15	2.80E-05
GO:0032787monocarboxylic acid metabolic process	71	3.50E-05
GO:0009606tropism	29	4.00E-05
GO:0009314response to radiation	88	4.10E-05
GO:0018130heterocycle biosynthetic process	28	0.0002
GO:0009734auxin mediated signaling pathway	28	0.0002
GO:0044262cellular carbohydrate metabolic process	77	0.0002
GO:0046394carboxylic acid biosynthetic process	64	0.0002
GO:0016053organic acid biosynthetic process	64	0.0002
GO:0046417chorismate metabolic process	16	0.0002
GO:0009073aromatic amino acid family biosynthetic process	16	0.0002
GO:0016114terpenoid biosynthetic process	23	0.00024
GO:0009072aromatic amino acid family metabolic process	19	0.00029
GO:0016108tetraterpenoid metabolic process	15	0.00039
GO:0016116carotenoid metabolic process	15	0.00039
GO:0009886post-embryonic morphogenesis	24	0.00043
GO:0046148pigment biosynthetic process	29	0.0006
GO:0016109tetraterpenoid biosynthetic process	12	0.00071
GO:0016117carotenoid biosynthetic process	12	0.00071
GO:0008299isoprenoid biosynthetic process	25	0.00073
GO:0010382cellular cell wall macromolecule metabolic process	8	0.00077
GO:0006631fatty acid metabolic process	41	0.0009
GO:0042440pigment metabolic process	31	0.00098
GO:0009629response to gravity	22	0.001
GO:0043648dicarboxylic acid metabolic process	20	0.0012
GO:0010876lipid localization	8	0.002
GO:0006721terpenoid metabolic process	25	0.0021
GO:0006720isoprenoid metabolic process	28	0.0021

Enriched GO category	Number of associated genes	FDR
GO:0046246terpene biosynthetic process	11	0.0031
GO:0044264cellular polysaccharide metabolic process	39	0.0037
GO:0008610lipid biosynthetic process	61	0.0038
GO:0019748secondary metabolic process	71	0.004
GO:0044271cellular nitrogen compound biosynthetic process	50	0.0046
GO:0006073cellular glucan metabolic process	30	0.0046
GO:0009605response to external stimulus	75	0.0063
GO:0019438aromatic compound biosynthetic process	44	0.0078
GO:0016051carbohydrate biosynthetic process	48	0.0079
GO:0009630gravitropism	19	0.0082
GO:0010166wax metabolic process	11	0.011
GO:0044042glucan metabolic process	30	0.012
GO:0010025wax biosynthetic process	10	0.013
GO:0009733response to auxin stimulus	50	0.014
GO:0009637response to blue light	15	0.014
GO:0050896response to stimulus	362	0.018
GO:0044255cellular lipid metabolic process	78	0.023
GO:0005976polysaccharide metabolic process	43	0.028
GO:0042214terpene metabolic process	12	0.028
GO:0006633fatty acid biosynthetic process	26	0.034
GO:0009266response to temperature stimulus	66	0.035
GO:0030244cellulose biosynthetic process	11	0.035
GO:0009638phototropism	11	0.038
GO:0005975carbohydrate metabolic process	106	0.038
GO:0009834secondary cell wall biogenesis	12	0.04
GO:0042546cell wall biogenesis	24	0.045
GO:0033692cellular polysaccharide biosynthetic process	27	0.05
GO:0016829lyase activity	55	0.017
GO:0016491oxidoreductase activity	152	0.017
GO:0035251UDP-glucosyltransferase activity	26	0.017
GO:0009055electron carrier activity	69	0.02
GO:0016860intramolecular oxidoreductase activity	12	0.02
GO:0008194UDP-glycosyltransferase activity	35	0.02
GO:0046527glucosyltransferase activity	28	0.02
GO:0016835carbon-oxygen lyase activity	23	0.028
GO:0016760cellulose synthase (UDP-forming) activity	10	0.029
GO:0008757S-adenosylmethionine-dependent methyltransferase activity	21	0.029

Enriched GO category	Number of associated genes	FDR
GO:0004497monooxygenase activity	37	0.038
GO:0016758transferase activity, transferring hexosyl groups	46	0.038
GO:0009507chloroplast	222	2.30E-12
GO:0009536plastid	234	1.50E-11
GO:0044434chloroplast part	80	2.50E-08
GO:0044435plastid part	85	2.30E-07
GO:0005618cell wall	74	5.00E-07
GO:0009505plant-type cell wall	65	1.00E-06
GO:0030312external encapsulating structure	76	1.70E-06
GO:0009535chloroplast thylakoid membrane	42	1.90E-06
GO:0055035plastid thylakoid membrane	42	1.10E-05
GO:0044444cytoplasmic part	485	1.50E-05
GO:0009534chloroplast thylakoid	42	3.40E-05
GO:0005737cytoplasm	542	9.40E-05
GO:0010287plastoglobule	17	9.40E-05
GO:0042651thylakoid membrane	42	9.80E-05
GO:0031976plastid thylakoid	42	0.00016
GO:0031984organelle subcompartment	42	0.00023
GO:0044436thylakoid part	42	0.001
GO:0044464cell part	901	0.001
GO:0005623cell	901	0.001
GO:0009941chloroplast envelope	21	0.0057
GO:0031225anchored to membrane	27	0.0064
GO:0009514glyoxysome	5	0.033
GO:0034357photosynthetic membrane	43	0.04
GO:0043227membrane-bounded organelle	561	0.049

Table S28. Trinity Results: Significantly enriched GO terms corresponding to the red stage of a *L. sessilifolius* flower.

Enriched GO category: Description	Number of associated genes	FDR
GO:0050896response to stimulus	503	3.40E-08
GO:0010035response to inorganic substance	54	5.50E-08
GO:0048610reproductive cellular process	20	4.00E-07
GO:0042221response to chemical stimulus	284	6.20E-07
GO:0009414response to water deprivation	57	4.40E-06
GO:0009415response to water	58	6.00E-06
GO:0010033response to organic substance	197	6.00E-06
GO:0006950response to stress	325	1.40E-05
GO:0009719response to endogenous stimulus	168	1.50E-05
GO:0034641cellular nitrogen compound metabolic process	85	6.20E-05
GO:0009737response to abscisic acid stimulus	75	1.10E-04
GO:0010876lipid localization	11	1.10E-04
GO:0009624response to nematode	48	1.10E-04
GO:0006970response to osmotic stress	73	0.00015
GO:0009628response to abiotic stimulus	202	0.00019
GO:0009651response to salt stress	61	0.00019
GO:0009725response to hormone stimulus	151	0.00039
GO:0009607response to biotic stimulus	137	0.0013
GO:0005975carbohydrate metabolic process	140	0.0016
GO:0030001metal ion transport	52	0.0022
GO:0044262cellular carbohydrate metabolic process	87	0.0024
GO:0051707response to other organism	126	0.0024
GO:0042398cellular amino acid derivative biosynthetic process	54	0.0056
GO:0010038response to metal ion	32	0.01
GO:0043436oxoacid metabolic process	121	0.01
GO:0046174polyol catabolic process	8	0.01
GO:0006082organic acid metabolic process	121	0.01
GO:0009611response to wounding	41	0.01
GO:0019752carboxylic acid metabolic process	121	0.01
GO:0051704multi-organism process	152	0.012
GO:0006813potassium ion transport	19	0.017
GO:0042180cellular ketone metabolic process	121	0.017
GO:0006811ion transport	83	0.02
GO:0009605response to external stimulus	88	0.021
GO:0006575cellular amino acid derivative metabolic process	65	0.026
GO:0006519cellular amino acid and derivative metabolic process	107	0.026
GO:0009751response to salicylic acid stimulus	39	0.026

Enriched GO category	Number of associated genes	FDR
GO:0010107potassium ion import	6	0.034
GO:0044106cellular amine metabolic process	66	0.034
GO:0009991response to extracellular stimulus	33	0.049
GO:0016791phosphatase activity	50	0.00063
GO:0042578phosphoric ester hydrolase activity	56	0.0011
GO:0022891substrate-specific transmembrane transporter activity	116	0.0063
GO:0022857transmembrane transporter activity	140	0.008
GO:0015291secondary active transmembrane transporter activity	54	0.012
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	62	0.012
GO:0022892substrate-specific transporter activity	132	0.013
GO:0009055electron carrier activity	83	0.013
GO:0016798hydrolase activity, acting on glycosyl bonds	65	0.015
GO:0005506iron ion binding	85	0.017
GO:0022804active transmembrane transporter activity	98	0.017
GO:0005215transporter activity	173	0.023
GO:0015075ion transmembrane transporter activity	86	0.026
GO:0008374O-acyltransferase activity	18	0.031
GO:0020037heme binding	51	0.043
GO:0016021integral to membrane	199	0.0099
GO:0031224intrinsic to membrane	224	0.03
GO:0050896response to stimulus	503	3.40E-08
GO:0010035response to inorganic substance	54	5.50E-08
GO:0048610reproductive cellular process	20	4.00E-07
GO:0042221response to chemical stimulus	284	6.20E-07
GO:0009414response to water deprivation	57	4.40E-06
GO:0009415response to water	58	6.00E-06
GO:0010033response to organic substance	197	6.00E-06
GO:0006950response to stress	325	1.40E-05
GO:0009719response to endogenous stimulus	168	1.50E-05
GO:0034641cellular nitrogen compound metabolic process	85	6.20E-05
GO:0009737response to abscisic acid stimulus	75	1.10E-04
GO:0010876lipid localization	11	1.10E-04
GO:0009624response to nematode	48	1.10E-04
GO:0006970response to osmotic stress	73	0.00015
GO:0009628response to abiotic stimulus	202	0.00019
GO:0009651response to salt stress	61	0.00019
GO:0009725response to hormone stimulus	151	0.00039
GO:0009607response to biotic stimulus	137	0.0013

Enriched GO category	Number of associated genes	FDR
GO:0005975carbohydrate metabolic process	140	0.0016
GO:0030001metal ion transport	52	0.0022
GO:0044262cellular carbohydrate metabolic process	87	0.0024
GO:0051707response to other organism	126	0.0024
GO:0042398cellular amino acid derivative biosynthetic process	54	0.0056
GO:0010038response to metal ion	32	0.01
GO:0043436oxoacid metabolic process	121	0.01
GO:0046174polyol catabolic process	8	0.01
GO:0006082organic acid metabolic process	121	0.01
GO:0009611response to wounding	41	0.01
GO:0019752carboxylic acid metabolic process	121	0.01
GO:0051704multi-organism process	152	0.012
GO:0006813potassium ion transport	19	0.017
GO:0042180cellular ketone metabolic process	121	0.017
GO:0006811ion transport	83	0.02
GO:0009605response to external stimulus	88	0.021
GO:0006575cellular amino acid derivative metabolic process	65	0.026
GO:0006519cellular amino acid and derivative metabolic process	107	0.026
GO:0009751response to salicylic acid stimulus	39	0.026
GO:0010107potassium ion import	6	0.034
GO:0044106cellular amine metabolic process	66	0.034
GO:0009991response to extracellular stimulus	33	0.049
GO:0016791phosphatase activity	50	0.00063
GO:0042578phosphoric ester hydrolase activity	56	0.0011
GO:0022891substrate-specific transmembrane transporter activity	116	0.0063
GO:0022857transmembrane transporter activity	140	0.008
GO:0015291secondary active transmembrane transporter activity	54	0.012
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	62	0.012
GO:0022892substrate-specific transporter activity	132	0.013
GO:0009055electron carrier activity	83	0.013
GO:0016798hydrolase activity, acting on glycosyl bonds	65	0.015
GO:0005506iron ion binding	85	0.017
GO:0022804active transmembrane transporter activity	98	0.017
GO:0005215transporter activity	173	0.023
GO:0015075ion transmembrane transporter activity	86	0.026
GO:0008374O-acyltransferase activity	18	0.031
GO:0020037heme binding	51	0.043
GO:0016021integral to membrane	199	0.0099

Enriched GO category	Number of associated genes	FDR
GO:0031224intrinsic to membrane	224	0.03
GO:0022857transmembrane transporter activity	140	0.008
GO:0015291secondary active transmembrane transporter activity	54	0.012
GO:0004553hydrolase activity, hydrolyzing O-glycosyl compounds	62	0.012
GO:0022892substrate-specific transporter activity	132	0.013
GO:0009055electron carrier activity	83	0.013
GO:0016798hydrolase activity, acting on glycosyl bonds	65	0.015
GO:0005506iron ion binding	85	0.017
GO:0022804active transmembrane transporter activity	98	0.017
GO:0005215transporter activity	173	0.023
GO:0015075ion transmembrane transporter activity	86	0.026
GO:0008374O-acyltransferase activity	18	0.031
GO:0020037heme binding	51	0.043
GO:0016021integral to membrane	199	0.0099
GO:0031224intrinsic to membrane	224	0.03

Table S29. Red Stage Trinity Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=0.11). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN51676_chr2.CM0641.560.nc c1_g2		Glyma.01G048200.1	Ribonuclease T(2) / Ribonuclease T2	11	353	-5.0898
TRINITY_DN51676_chr2.CM0641.550.nc c1_g1		Glyma.01G048400.1	Ribonuclease T(2) / Ribonuclease T2	19	1992	-6.8045
TRINITY_DN50283_chr2.CM0124.40.nc c0_g1		Glyma.01G067100.1	F11M15.8 PROTEIN- RELATED	17	300	-4.2331
TRINITY_DN52390_chr4.CM0119.240.nc c0_g1		Glyma.02G048400.1	NARINGENIN,2- OXOGLUTARATE 3- DIOXYGENASE (F3H)	276	1893	-2.8794
TRINITY_DN52643_chr2.CM0124.280.nd c1_g1		Glyma.02G124700.1	Codeine 3-O- demethylase / Codeine O-demethylase	180	884	-2.3973
TRINITY_DN44592_chr1.CM0133.170.nc c1_g1		Glyma.04G052100.1	Thromboxane-A synthase / Thromboxane synthetase	0	38	-8.3044
TRINITY_DN52292_LjT09A12.100.nd c0_g1		Glyma.06G103200.1	CRYPTOCHROME-1	365	4597	-3.7563
TRINITY_DN52780_chr4.CM0429.360.nc c0_g1		Glyma.06G202300.1	FLAVONOID 3'- MONOOXYGENASE	7	2073	-8.2859
TRINITY_DN36546_chr3.CM0208.350.nc c0_g1		Glyma.07G048600.1	F3H7.17 PROTEIN	93	520	-2.5836
TRINITY_DN45435_chr4.CM0387.330.nc c0_g1		Glyma.08G092800.1	2-OXOGLUTARATE (2OG) AND FE(II)- DEPENDENT OXYGENASE SUPERFAMILY PROTEIN-RELATED	0	2374	-14.265
TRINITY_DN46871_LjSGA_066111.1 c0_g2		Glyma.08G110000.1	Quinate O- hydroxycinnamoyltrans ferase / Hydroxycinnamoyl coenzyme A-quinat e transferase	5	35	-2.8775

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN45808_LjSGA_081858.1 c0_g1		Glyma.10G138800.1	TRANSCRIPTION FACTOR PIF3	0	47	-8.6102
TRINITY_DN54168_chr1.CM0284.240.nc c0_g1		Glyma.11G011500.1	Naringenin-chalcone synthase / Flavonone synthase	17	1134	-6.1510
TRINITY_DN54168_chr1.CM0591.360.nd c0_g2		Glyma.11G011500.1	Naringenin-chalcone synthase / Flavonone synthase	65	4382	-6.1743
TRINITY_DN53094_chr2.CM0304.350.nc c0_g1		Glyma.11G027700.1	LEUCOANTHOCYANIDIN DIOXYGENASE-RELATED (ANS)	30	3761	-7.0659
TRINITY_DN47758_LjT35E03.200.nd c0_g1		Glyma.12G194000.1	MULTIDRUG RESISTANCE PROTEIN	50	664	-3.8298
TRINITY_DN53202_chr3.CM0711.140.nd c0_g1		Glyma.12G199800.1	Transferase family (Transferase)	12	456	-5.3349
TRINITY_DN53800_chr6.CM0055.290.nc c0_g1		Glyma.13G255800.1	Anthocyanidin 3-O-glucosyltransferase / Uridine diphosphoglucose-anthocyanidin 3-O-glucosyltransferase // Flavonol 3-O-glucosyltransferase / UDP-glucose flavonol 3-O-glucosyltransferase	262	3390	-3.7951
TRINITY_DN50775_chr3.CM0711.100.nd c0_g1		Glyma.13G302300.1	COUMAROYL-COA:ANTHOCYANIDIN 3-O-GLUCOSIDE-6"-O-COUMAROYLTRANSFERASE 1-RELATED	0	6587.81	-15.737
TRINITY_DN49460_LjSGA_019127.1 c0_g1		Glyma.17G061000.1	SPERMIDINE HYDROXYCINNAMOYL TRANSFERASE	1	786	-9.5451
TRINITY_DN52789_chr5.CM0077.790.nc c0_g1		Glyma.17G252200.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	19	434	-4.6063

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN52652_chr1.CM0593.380.nc c1_g1		Glyma.19G105100.1	Naringenin-chalcone synthase / Flavonone synthase	145	660	-2.2875
TRINITY_DN49414_chr1.CM0012.680.nd c0_g1		Glyma.19G187000.1	UDP-GLYCOSYLTRANSFERASE 73C7	136	1143	-3.1720

Table S30. Yellow Stage Trinity Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=1). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST equivalent, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN45233_chr4.CM0227.650.nc c0_g1		Glyma.02G104600.1	GLUCOSYL/GLUCURON OSYL TRANSFERASES	65	9	2.7339
TRINITY_DN45432_chr2.CM0124.350.nd c0_g1		Glyma.02G125100.1	2-OXOGLUTARATE (2OG) AND FE(II)- DEPENDENT OXYGENASE SUPERFAMILY PROTEIN-RELATED	319	45	2.7201
TRINITY_DN51740_LjSGA_017175.2 c2_g2		Glyma.04G227700.1	FLAVONE 3'-O- METHYLTRANSFERASE 1	1247	281	2.0472
TRINITY_DN45545_chr3.LjT10E18.60.nc c0_g1		Glyma.12G174100.1	AUXIN RESPONSE FACTOR 10-RELATED	23	2.4	3.3450
TRINITY_DN52633_LjSGA_025000.1 c4_g1		Glyma.12G235100.1	OXIDOREDUCTASE, 2OG-FE II OXYGENASE FAMILY PROTEIN	962	155	2.5307
TRINITY_DN39800_chr1.LjB17A22.120.nc c0_g1		Glyma.13G082300.1	FLAVONOL SYNTHASE 3-RELATED	1601	112	3.7338
TRINITY_DN51578_LjSGA_065630.1 c0_g1		Glyma.18G208600.1	GLUCOSYL/GLUCURON OSYL TRANSFERASES	1413	162	3.0216

Table S31. Similarly Expressed Trinity Results: Genes for enzymes annotated to GO:0009813 (Flavonoid biosynthetic process, FDR=). The first two columns list Trinity contigs expresand gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log2FC
TRINITY_DN41192_chr5.CM0077.690.nc c0_g1		Glyma.02G158700.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	2	0	4.0402
TRINITY_DN25538_chr1.CM0122.460.nc c0_g1		Glyma.03G222100.2	NITROGEN REGULATORY PROTEIN P-II	38	14	1.3309
TRINITY_DN48324_chr1.CM0476.340.nd c0_g1		Glyma.04G228000.1	PROTEIN TRANSPARENT TESTA GLABRA 1	493	570	-0.3114
TRINITY_DN49922_chr4.CM0007.970.nc c0_g1		Glyma.05G127300.2	26S PROTEASOME NON-ATPASE REGULATORY SUBUNIT 8	214	247	-0.3088
TRINITY_DN41679_chr4.CM0042.2280.nc c0_g1		Glyma.05G245800.1	PATHOGENESIS-RELATED THAUMATIN-LIKE PROTEIN-RELATED	3	0	4.5957
TRINITY_DN36305_chr1.CM0033.750.nc c0_g1		Glyma.10G053500.1	B3 DNA binding domain (B3) // Auxin response factor (Auxin_resp)	6	0	5.5655
TRINITY_DN38832_LjSGA_024944.1 c0_g1		Glyma.10G210600.1	B3 DNA binding domain (B3) // Auxin response factor (Auxin_resp)	28	4	2.6690
TRINITY_DN31599_chr3.LjT10E18.60.nc c0_g1		Glyma.11G145500.1	AUXIN RESPONSE FACTOR 10-RELATED	8	0	5.9729
TRINITY_DN18931_LjSGA_078422.1 c0_g1		Glyma.11G145500.1	AUXIN RESPONSE FACTOR 10-RELATED	1	2	-1.0110
TRINITY_DN45545_chr3.LjT10E18.60.nc c0_g1		Glyma.12G174100.1	AUXIN RESPONSE FACTOR 10-RELATED	23	2.4	3.3450
TRINITY_DN56369_chr3.CM0590.460.nd c0_g1		Glyma.13G279600.1	COP1-INTERACTING PROTEIN 7 (CIP7)	4	0	4.9957

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN50775_chr3.CM0711.100.nd c0_g1		Glyma.13G302300.1	COUMAROYL-COA:ANTHOCYANIDIN 3-O-GLUCOSIDE-6"-O-COUMAROYLTRANSFERASE 1-RELATED	0	6587.81	-15.7376
TRINITY_DN45545_chr3.LjT10E18.60.nc c2_g1		Glyma.13G325200.1	AUXIN RESPONSE FACTOR 10-RELATED	4	0	4.9957
TRINITY_DN52789_chr5.CM0077.690.nc c0_g2		Glyma.14G072700.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	18	3	2.4363
TRINITY_DN55140_chr5.CM0077.690.nc c0_g1		Glyma.14G072700.1	Dihydrokaempferol 4-reductase / NADPH-dihydromyricetin reductase	0	2	-4.1364
TRINITY_DN53766_chr1.CM0410.100.nc c0_g2		Glyma.18G268100.1	Isoflavone-7-O-beta-glucoside 6"-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	187	206	-0.2415
TRINITY_DN33124_chr1.CM0410.130.nc c0_g1		Glyma.18G268100.1	Isoflavone-7-O-beta-glucoside 6"-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	10	6	0.6246
TRINITY_DN57182_chr1.CM0410.130.nc c0_g1		Glyma.18G268100.1	Isoflavone-7-O-beta-glucoside 6"-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	12	7	0.6663

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN6325_c0_g1	chr1.CM0410.130.nc	Glyma.18G268100.1	Isoflavone-7-O-beta-glucoside 6''-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	5	1	2.0924
TRINITY_DN53766_c0_g2	chr1.CM0410.80.nc	Glyma.18G268100.1	Isoflavone-7-O-beta-glucoside 6''-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	187	206	-0.2415
TRINITY_DN54421_c2_g4	chr1.CM0410.40.nc	Glyma.18G268600.1	Isoflavone-7-O-beta-glucoside 6''-O-malonyltransferase / Flavone/flavonol 7-O-beta-D-glucoside malonyltransferase	407	211.72	0.8384
TRINITY_DN53649_c3_g1	chr1.CM0222.40.nc	Glyma.19G106900.1	WD REPEAT CONTAINING PROTEIN	211	148	0.4092
TRINITY_DN41800_c0_g1	LjSGA_045968.2	Glyma.20G180000.1	Glyma.20G180000.1.p - (M=20) PF02362//PF06507 -	4	0	4.9957
TRINITY_DN45569_c1_g1	LjSGA_045968.2	Glyma.20G180000.1	Glyma.20G180000.1.p - (M=20) PF02362//PF06507 -	1	00	

Table S32. Yellow Stage Trinity Results: Genes for enzymes annotated to the over-represented GO:0009809 (Lignin biosynthetic process, FDR=1). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN45532_chr2.CM0249.1210.nc c2_g1		Glyma.01G173600.1	LACCASE-11	23	0	7.4814
TRINITY_DN45532_chr2.CM0249.1210.nc c4_g1		Glyma.01G173600.1	LACCASE-11	17	0	7.0482
TRINITY_DN46194_chr4.CM0227.500.nc c0_g1		Glyma.02G103500.1	O-METHYLTRANSFERASE-RELATED	8146	826	3.1995
TRINITY_DN51740_LjSGA_017175.2 c2_g2		Glyma.04G227700.1	FLAVONE 3'-O-METHYLTRANSFERASE 1	1247	281	2.0472
TRINITY_DN44781_LjSGA_035751.3 c0_g1		Glyma.07G089700.1	Abieta-7,13-dien-18-ol hydroxylase / CYP720B1	8110	225	5.0688
TRINITY_DN43710_LjSGA_035751.3 c0_g1		Glyma.07G089700.1	Abieta-7,13-dien-18-ol hydroxylase / CYP720B1	91.32	12	2.8083
TRINITY_DN44818_chr4.CM0042.560.nd c0_g1		Glyma.07G089800.1	Cytochrome P450 (p450)	40	2	4.1399
TRINITY_DN45532_chr2.CM0249.1210.nc c0_g1		Glyma.11G069500.1	LACCASE-11	25	0	7.6011
TRINITY_DN52386_LjSGA_011490.1 c0_g1		Glyma.12G019700.1	ALCOHOL DEHYDROGENASE-RELATED	76	11	2.6730
TRINITY_DN34978_chr1.LjB20I01.110.nd c0_g2		Glyma.14G062300.1	MULTI-COPPER OXIDASE	11	0	6.4261
TRINITY_DN45845_CM0385.150.nd c0_g1		Glyma.15G143600.1	CHITINASE-LIKE PROTEIN 1	225	15	3.7940

Table S33. Red Stage Trinity Results: Genes for enzymes annotated to the over-represented GO:0009809 (Lignin biosynthetic process, FDR=). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN49460_LjSGA_019127.1 c0_g1		Glyma.17G061000.1	SPERMIDINE HYDROXYCINNAMOYL TRANSFERASE	1	786	-9.5451
TRINITY_DN45071_LjSGA_021886.2 c0_g1		Glyma.17G171100.1	O-METHYLTRANSFERASE- RELATED	2	72	-5.1840
TRINITY_DN49973_LjSGA_030253.1 c0_g1		Glyma.13G141600.1	S-ADENOSYLMETHIONINE SYNTHASE 3	206	707	-1.8805

Table S34. Similarly Expressed Trinity Results: Genes for enzymes annotated to GO:0009813 (Lignin process, FDR=1). The first two columns list Trinity contigs expresand gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN53597_chr2.CM0249.290.nd c0_g1		Glyma.01G169200.1	Cytochrome P450 CYP2 subfamily	170	92	0.7828
TRINITY_DN41837_LjSGA_021886.2 c0_g1		Glyma.01G187700.1	O-METHYLTRANSFERASE-RELATED	1831	863	0.9829
TRINITY_DN52969_LjSGA_096258.1 c0_g1		Glyma.03G122000.1	CYTOCHROME P450 98A3-RELATED	1429	1471	-0.1439
TRINITY_DN49973_LjSGA_030253.1 c2_g1		Glyma.03G184000.1	S-ADENOSYLMETHIONINE SYNTHASE 3	4222	10944	-1.4762
TRINITY_DN16999_chr1.CM0063.200.nc c0_g1		Glyma.05G203000.1	OMEGA-HYDROXYPALMITATE O-FERULOYL TRANSFERASE	0	10	-6.3911
TRINITY_DN49267_LjT16L08.100.nd c1_g2		Glyma.05G233200.1	Cytochrome P450 CYP2 subfamily	144	441	-1.7159
TRINITY_DN17837_chr3.LjT40P18.60.nc c0_g1		Glyma.07G021600.1	SHIKIMATE O-HYDROXYCINNAMOYLTRANSFERASE	5	6	-0.3570
TRINITY_DN53428_LjSGA_040039.1 c1_g1		Glyma.07G026300.1	CINNAMOYL-COA:NADP OXIDOREDUCTASE-LIKE 1-RELATED	156	46	1.6571
TRINITY_DN33065_LjSGA_027787.1 c0_g1		Glyma.07G050600.1	NAC DOMAIN-CONTAINING PROTEIN 43-RELATED	3	0	4.59573
TRINITY_DN33065_LjSGA_027787.1 c0_g2		Glyma.07G050600.1	NAC DOMAIN-CONTAINING PROTEIN 43-RELATED	0	00	
TRINITY_DN66560_LjSGA_027787.1 c0_g1		Glyma.07G050600.1	NAC DOMAIN-CONTAINING PROTEIN 43-RELATED	1	00	
TRINITY_DN17072_LjSGA_035751.3 c0_g1		Glyma.07G089700.1	Abieta-7,13-dien-18-ol hydroxylase / CYP720B1	102	31	1.61231
TRINITY_DN49973_LjSGA_140570.1 c3_g1		Glyma.07G233800.1	S-ADENOSYLMETHIONINE SYNTHASE 4	13	15	-0.3058
TRINITY_DN51745_chr3.CM0452.240.nd c0_g1		Glyma.08G220200.1	SHIKIMATE O-HYDROXYCINNAMOYLTRANSFERASE	568	908	-0.7788

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN42496_chr1.CM0410.190.nc c1_g1		Glyma.08G246100.1	LRR RECEPTOR-LIKE SERINE/THREONINE- PROTEIN KINASE RPK2	10	7	0.406346
TRINITY_DN47225_chr1.CM0410.190.nc c0_g1		Glyma.08G246100.1	LRR RECEPTOR-LIKE SERINE/THREONINE- PROTEIN KINASE RPK2	24	5	2.13427
TRINITY_DN37023_LjSGA_030670.1 c0_g2		Glyma.08G311800.1	SPERMIDINE HYDROXYCINNAMOYL TRANSFERASE	4	1	1.77955
TRINITY_DN37010_chr1.CM0147.770.nc c0_g1		Glyma.08G359100.1	LACCASE-17	1	1	-0.0908
TRINITY_DN32484_LjSGA_011490.1 c0_g1		Glyma.12G019700.1	ALCOHOL DEHYDROGENASE- RELATED	5	4	0.21381
TRINITY_DN41325_LjSGA_011490.1 c0_g1		Glyma.12G019700.1	ALCOHOL DEHYDROGENASE- RELATED	5	1	2.09241
TRINITY_DN41325_LjSGA_011490.1 c0_g2		Glyma.12G019700.1	ALCOHOL DEHYDROGENASE- RELATED	4	6	-0.6698
TRINITY_DN49071_LjSGA_015763.2 c2_g3		Glyma.12G019700.1	ALCOHOL DEHYDROGENASE- RELATED	2	2	-0.0961
TRINITY_DN54121_LjSGA_013806.1 c0_g1		Glyma.12G109800.1	O-METHYLTRANSFERASE FAMILY PROTEIN	124	228	-0.9800
TRINITY_DN49611_LjT45M09.100.nd c0_g1		Glyma.13G369800.1	NAD DEPENDENT EPIMERASE/DEHYDRATA SE	1434	375	1.83260
TRINITY_DN35253_chr1.LjB20I01.110.nd c1_g1		Glyma.14G062300.1	MULTI-COPPER OXIDASE	3	0	4.59573
TRINITY_DN50309_LjSGA_070619.1 c0_g1		Glyma.15G059500.1	ALCOHOL DEHYDROGENASE RELATED	48	24	0.89451
TRINITY_DN51988_LjT35D18.30.nc c1_g1		Glyma.15G190500.1	S-ADENOSYLMETHIONINE SYNTHASE 4	8441	2869	1.45468

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN2300_c0_g1	chr1.CM0593.100.nd	Glyma.16G047300.1	Protein kinase domain (Pkinase) // Leucine rich repeat N-terminal domain (LRRNT_2) // Leucine rich repeat (LRR_8)	6	4	0.47076
TRINITY_DN38374_c0_g1	chr1.CM0593.100.nd	Glyma.16G047300.1	Protein kinase domain (Pkinase) // Leucine rich repeat N-terminal domain (LRRNT_2) // Leucine rich repeat (LRR_8)	0	10	-6.3911
TRINITY_DN38272_c0_g1	chr1.CM0593.40.nd	Glyma.16G047300.1	Protein kinase domain (Pkinase) // Leucine rich repeat N-terminal domain (LRRNT_2) // Leucine rich repeat (LRR_8)	10	0	6.29031
TRINITY_DN49071_c1_g1	LjT48C16.160.nd	Glyma.18G091500.1	ALCOHOL DEHYDROGENASE-RELATED	93	70	0.30726
TRINITY_DN37010_c1_g1	chr1.CM0147.770.nc	Glyma.18G177200.1	MULTI-COPPER OXIDASE	4	0	4.99577
TRINITY_DN47225_c1_g1	chr1.CM0410.190.nc	Glyma.18G267000.1	LRR RECEPTOR-LIKE SERINE/THREONINE-PROTEIN KINASE RPK2	61.37	47	0.27337
TRINITY_DN47225_c1_g2	chr1.CM0410.190.nc	Glyma.18G267000.1	LRR RECEPTOR-LIKE SERINE/THREONINE-PROTEIN KINASE RPK2	1.63	0	4.04029
TRINITY_DN50519_c0_g1	chr5.CM0200.1180.nc	Glyma.20G128600.1	CINNAMYL ALCOHOL DEHYDROGENASE 4-RELATED	423	183	1.1061
TRINITY_DN58801_c0_g1	LjSGA_075859.1	Glyma.20G175500.1	No apical meristem (NAM) protein (NAM)	4	1	1.7795

Table S35. *L. filicaulis* Yellow Stage Trinity Results: Genes for enzymes annotated to GO:00016117 (Carotenoid biosynthetic process, FDR=0.11). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR50146 c0_g1	chr3.CM0996.270.r2.a	Glyma.01G028900.1	ATP-CITRATE SYNTHASE	34.96	2	3.7015
TR45451 c0_g1	chr4.CM0570.280.r2.a	Glyma.01G064200.1	AGAMOUS-LIKE MADS-BOX PROTEIN AGL8-RELATED	1414	151	2.8697
TR18321 c0_g1	LjSGA_059500.1	Glyma.01G186200.1	ZEAXANTHIN EPOXIDASE, CHLOROPLASTIC	310	63	1.9404
TR32448 c0_g2	chr5.CM1125.160.r2.m	Glyma.02G149400.1	Not Available	10.54	0.64	3.1901
TR52234 c0_g2	LjSGA_013562.2	Glyma.02G188200.1	Prolycopene isomerase / CRTISO	6.01	18.45	2.2958
TR33198 c0_g1	chr2.LjT08I01.60.r2.a	Glyma.02G240200.1	PHYTOENE SYNTHASE, CHLOROPLASTIC	4765	1049	1.8268
TR64960 c0_g1	chr3.CM0282.920.r2.m	Glyma.06G238100.1	SQUAMOSA PROMOTER-BINDING-LIKE PROTEIN 4-RELATED	117	1	6.3654
TR22838 c0_g1	chr5.CM0456.560.r2.m	Glyma.09G132200.1	Beta-carotene 3-hydroxylase / Beta-carotene 3,3'-monooxygenase	4	00	
TR60004 c0_g2	chr5.CM1077.670.r2.m	Glyma.10G062900.1	Carotene epsilon-monooxygenase / LUT1	33.2	8	7.9261
TR3908 c0_g1	chr5.CM1323.150.r2.m	Glyma.10G160500.1	9,9'-di-cis-zeta-carotene desaturase / Zeta-carotene desaturase	21.7	5.36	1.9696
TR53492 c0_g1	chr4.CM0044.50.r2.m	Glyma.11G252500.1	Zeta-carotene isomerase / 15-cis-zeta-carotene isomerase	446	18	4.2661
TR15422 c0_g1	chr1.CM0104.3370.r2.m	Glyma.16G036100.1	POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE 1, CHLOROPLASTIC	27.03	0	7.5950
TR18321 c1_g1	LjSGA_078276.1	NA	zeaxanthin epoxidase	4274.46	305.41	3.4517
TR18321 c2_g1	LjSGA_024159.1	NA	zeaxanthin epoxidase	366	77	1.8908

Table S36. *L. filicaulis* Similarly Expressed Trinity Results: Genes for enzymes annotated to GO:00016117 (Carotenoid biosynthetic process, FDR=1). The first two columns list Trinity contigs expresand gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR52234 c0_g1	LjSGA_013562.2	Glyma.02G188200.1	ATP-CITRATE SYNTHASE	6.01	18.45	-1.9168
TR15191 c0_g1	chr6.CM0013.1810.r2.d	Glyma.09G066200.1	Not Available	143	46	1.27775
TR7103 c0_g1	LjSGA_075584.1	Glyma.07G273600.1	Prolycopene isomerase / CRTISO	129.46	16	3.93057
TR32448 c0_g1	chr5.CM1125.160.r2.m	Glyma.02G149400.1	Glutathione transferase / S-(hydroxyalkyl)glutathione lyase	10.54	0.64	2.97110
TR15422 c0_g1	chr1.CM0104.3370.r2.m	Glyma.16G036100.1	LYCOPENE BETA CYCLASE, CHLOROPLASTIC	27.03	0	1.93647
TR51506 c0_g1	chr6.CM0367.660.r2.d	Glyma.09G008500.1	Prolycopene isomerase / CRTISO	2279	594	1.58315
TR42882 c0_g1	LjSGA_050034.1	Glyma.09G000600.1	MONOOXYGENASE	4	0	4.88247
TR64376 c0_g1	LjSGA_046377.1	Glyma.02G283400.1	UBIQUINOL OXIDASE 4, CHLOROPLASTIC/CHROMOPLASTIC	222	571	-1.7188
TR61986 c0_g1	LjSGA_045520.1	Glyma.03G128900.1	LYCOPENE EPSILON CYCLASE, CHLOROPLASTIC	635	164	1.59588
TR60004 c0_g1	chr5.CM1077.670.r2.m	Glyma.10G062900.1	Beta-carotene 3-hydroxylase / Beta-carotene 3,3'-monooxygenase	33.2	8	1.67434
TR13886 c0_g1	chr5.CM1323.150.r2.m	Glyma.10G160500.1	PROTEIN LUTEIN DEFICIENT 5, CHLOROPLASTIC	21.7	5.36	1.75885
TR33253 c0_g1	chr1.CM1413.250.r2.a	Glyma.09G252800.1	Carotene epsilon-monooxygenase / LUT1	963	463	0.69985
TR61205 c0_g1	chr6.LjT13N05.10.r2.m	Glyma.11G253000.1	9,9'-di-cis-zeta-carotene desaturase / Zeta-carotene desaturase	1488	691	0.7499

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TR6924 c0_g1	chr2.CM0545.170.r2.m	Glyma.14G046600.1	15-cis-phytoene desaturase / Plant-type phytoene desaturase	81	23	1.45539
TR24925 c0_g1	LjSGA_031176.1	Glyma.17G174500.1	LYCOPENE BETA/EPSILON CYCLASE PROTEIN	241	58	1.6964
TR22838 c0_g1	chr5.CM0456.560.r2.m	Glyma.09G132200.1	POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSF ERASE 1, CHLOROPLASTIC	4	0	4.88247
TR50146 c2_g1	chr2.CM0008.330.r2.m	Glyma.01G028900.1	ZEAXANTHIN EPOXIDASE, CHLOROPLASTIC	840	625	0.06998

Table S37. *L. sessilifolius* Yellow Stage Trinity Results: Genes for enzymes annotated to GO:00016117 (Carotenoid biosynthetic process, FDR=0.00027). The first two columns list differentially expressed (FDR<0.05) Trinity contigs and gene names as annotated by BLASTing the *Lotus japonicus* genome. The third column is their soybean BLAST homolog, followed by expression at two stages of flower development.

Trinity contig	<i>L. japonicus</i> ID	<i>G. max</i> BLAST ID	Gene description	Counts per gene (Yellow)	Counts per gene (Red)	Log ₂ FC
TRINITY_DN53748_LjSGA_045520.1 c0_g1		Glyma.03G128900.1	LYCOPENE BETA CYCLASE, CHLOROPLASTIC	38.81	0	8.240016
TRINITY_DN50287_LjSGA_046377.1 c0_g1		Glyma.08G306200.1	Glutathione transferase / S-(hydroxyalkyl)glutathione lyase // 15-cis-phytoene synthase / PSase	102	10	3.232895
TRINITY_DN49175_LjT17B03.90.nd c0_g1		Glyma.09G008500.1	UBIQUINOL OXIDASE 4, CHLOROPLASTIC/CHROMOPLASTIC	1268	232	2.347616
TRINITY_DN50928_LjSGA_055879.1 c1_g1		Glyma.09G233300.1	Protein disulfide-isomerase / S-S rearrangase	187	42	2.0493
TRINITY_DN48832_LjSGA_013562.2 c0_g1		Glyma.10G086700.1	Prolycopene isomerase / CRTISO	249	43	2.428307
TRINITY_DN50878_chr2.CM0018.70.nc c0_g1		Glyma.11G016200.1	CYTOCHROME P450 97B3, CHLOROPLASTIC	170	13	3.59459
TRINITY_DN50044_chr6.LjT13N05.100.nc c0_g1		Glyma06g11820.1	15-cis-phytoene desaturase / Plant-type phytoene desaturase	2187	425	2.260955
TRINITY_DN53304_LjT01G11.200.nd c1_g1		Glyma.17G174500.1	ZEAXANTHIN EPOXIDASE, CHLOROPLASTIC	73	14	2.270513
TRINITY_DN38311_LjT40E15.150.nc c0_g2		Glyma.20G227900.1	9,9'-di-cis-zeta-carotene desaturase / Zeta-carotene desaturase	8.33	7.97	-0.10057
TRINITY_DN53720_LjSGA_021331.1 c2_g1		Glyma10g07210.1	carotene epsilon-monooxygenase	204	21	3.170623
TRINITY_DN53304_LjSGA_078276.1 c3_g1		Glyma11g05960.1	zeaxanthin epoxidase	185	12	3.830864
TRINITY_DN53304_LjSGA_024159.1 c5_g1		Glyma17g20020.1	zeaxanthin epoxidase	2027	321	2.556112

Table S38 Red stage *L. filicaulis*: 19 contigs with greatest magnitude of log fold-change recovered from entire transcriptome.

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (UniProt and AnnoMine)
TR7064 c2_g1	-16.5797	NA	LjT31N05.100.r2.d	SJCHGC01014 protein (Fragment) OS=Schistosoma japonicum
TR34298 c0_g1	-13.9764	NA	chr4.CM0026.290.r2.m	Putative senescence-associated protein (Fragment) OS=Pisum sativum
TR50987 c0_g1	-13.3631	LjSGA_061080.1.1	chr4.CM0026.290.r2.m	Putative senescence-associated protein (Fragment) OS=Pisum sativum
TR3216 c0_g3	-13.3391	LjSGA_089366.0.1	LjSGA_089366.0.1	NA
TR29498 c0_g1	-13.3057	chr4.CM0042.1410.nc	chr4.CM0042.1350.r2.m	Dynamin OS=Cucumis melo subsp. melo
TR48381 c3_g1	-13.1261	chr5.CM1439.180.nd	chr5.CM1439.390.r2.m	RING/U-box superfamily protein
TR42412 c0_g1	-13.1231	LjT03L03.160.nd	chr5.CM0696.1060.r2.a	NA
TR53102 c0_g1	-13.1113	chr4.CM0229.70.nc	chr4.CM0229.70.r2.m	Probable polygalacturonase At2g43860
TR40587 c0_g1	-13.0933	chr3.CM0129.380.nd	chr3.CM0129.50.r2.d	NA
TR2085 c0_g1	-13.0644	LjSGA_078330.2	chr4.CM0026.290.r2.m	NA
TR1538 c0_g2	-12.7012	LjSGA_125503.1	LjT31N05.100.r2.d	SJCHGC01014 protein (Fragment) OS=Schistosoma japonicum
TR3124 c0_g3	-12.6853	chr5.CM0040.150.nd	chr5.CM0040.150.r2.d	Protein of unknown function DUF1677, plant
TR4260 c0_g1	-12.5775	chr4.CM0288.1020.nd	chr4.CM0288.1020.r2.d	1-aminocyclopropane-1-carboxylate oxidase OS=Trifolium repens
TR14717 c0_g3	-12.558	LjSGA_070603.1	chr3.CM0213.540.r2.m	ATP-binding cassette transporter, putative OS=Ricinus communis
TR20264 c0_g1	-12.4303	LjSGA_025923.1	LjSGA_025923.1	NA

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (UniProt and AnnoMine)
TR48491 c0_g1	-12.3818	chr1.LjT16003.100.nd	chr1.LjT16003.100.r2.d	Legumain-like protease (Precursor) OS=Ixodes ricinus
TR46724 c0_g3	-12.3162	chr5.CM0200.550.nc	chr5.CM0200.2620.r2.m	Zinc finger protein OS=Cicer arietinum
TR49591 c0_g1	-12.3111	LjT07J01.20.nd	chr5.CM0052.370.r2.a	GmCK2p OS=Glycine max
TR23809 c0_g6	-12.1554	LjSGA_022872.2	LjSGA_022872.2	Elongation factor 1- alpha OS=Ricinus communis

Table S39 Yellow stage *L. filicaulis*: 19 contigs with greatest magnitude of log fold-change recovered from entire transcriptome.

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (Annomine and UniProt)
TR24089 c0_g1	14.89616	chr1.CM0088.370.r2.d	NA	NA
TR15262 c0_g2	14.61414	chr5.CM0072.330.r2.d	NA	Flowering- promoting factor 1- like protein 1
TR43132 c0_g1	14.19693	LjSGA_013861.2	LjSGA_013861.2	Putative auxin efflux carrier protein 6 OS=Medicago truncatula
TR55706 c0_g1	14.16986	chr6.CM0679.380.r2.m	NA	member of Alpha- Expansin Gene Family
TR47113 c0_g4	13.92623	LjSGA_045866.1	LjSGA_045866.1	Early-responsive to dehydration OS=Medicago truncatula
TR33662 c0_g2	13.54451	LjSGA_020750.1	LjSGA_020750.1	ribonuclease T2
TR16126 c0_g1	13.45947	LjSGA_032186.1	LjSGA_032186.1	NA
TR33992 c0_g1	13.24861	chr1.LjT44L17.240.r2.d	NA	NA
TR50837 c0_g1	13.23161	LjSGA_071404.1	LjSGA_071404.1	Lachrymatory-factor synthase, putative OS=Ricinus communis
TR56525 c0_g2	13.20029	chr3.CM0152.240.r2.m	NA	Beta-amylase
TR17308 c2_g2	13.02148	chr6.CM0037.300.r2.m	NA	Cellulose synthase 3 OS=Eucalyptus grandis
TR63082 c0_g1	12.93094	LjSGA_009749.1	LjSGA_009749.1	Putative ripening related protein OS=Cicer arietinum
TR50996 c0_g1	12.92962	chr3.CM0129.210.r2.d	NA	TPR domain protein OS=Arthrospira platensis NIES-39
TR48869 c0_g3	12.91239	LjSGA_012731.1	LjSGA_012731.1	DEAD (Asp-Glu-Ala- Asp) box polypeptide 5 OS=Mus musculus

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID		Annotation (UniProt and AnnoMine)
TR3230 c0_g2		12.8773	chr5.CM0052.240.r2.d	NA	RANBP2-like and GRIP domain-containing protein 5/6
TR16751 c0_g1		12.83714	chr6.CM0066.630.r2.m	NA	NA
TR53791 c0_g1		12.81876	LjSGA_020451.1	LjSGA_020451.1	WD repeat-containing protein OS=Medicago truncatula
TR8063 c0_g1		12.76956	chr4.CM0126.800.r2.d	NA	Glucose-methanol-choline oxidoreductase OS=Halococcus salifodinae DSM 8989
TR49365 c1_g1		12.73229	chr3.CM0136.330.r2.m	NA	Fasciclin-like AGP 11 OS=Populus canescens

Table S40 Red stage *L. sessilifolius* 19 contigs with greatest magnitude of log fold-change recovered from entire transcriptome.

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (Annomine and Uniprot)
TRINITY_DN50775_c0_g1	-15.7376	chr3.CM0711.100.nd	chr3.CM0711.100.r2.d	Malonyl- coenzyme:anthocyanin 5-O-glucoside-6'''-O- malonyltransferase
TRINITY_DN45435_c0_g1	-14.2652	chr4.CM0387.330.nc	chr4.CM0387.1100.r2.m	Thebaine 6-O- demethylase
TRINITY_DN11784_c0_g2	-13.6968	LjT33L17.60.nc	chr2.CM0660.220.r2.m	Bark storage protein A
TRINITY_DN54303_c0_g1	-12.8579	LjSGA_138785.1	LjSGA_138785.1	Raffionse synthase 2 OS=Glycine max
TRINITY_DN49759_c0_g1	-12.802	LjSGA_010954.1	LjSGA_010954.1	Catalytic, putative OS=Ricinus communis
TRINITY_DN50868_c0_g1	-12.5777	chr2.CM1150.550.nd	chr2.CM0028.320.r2.d	Metalloendoproteinase 1
TRINITY_DN52783_c0_g1	-12.3214	chr4.CM0307.20.nc	chr4.CM0307.20.r2.m	DNA polymerase III PolC-type
TRINITY_DN40888_c0_g2	-11.9258	chr1.CM0141.350.nd	chr1.CM0141.360.r2.a	R2R3 MYB related transcription factor OS=Ipomoea batatas
TRINITY_DN16795_c0_g1	-11.8435	LjT03L03.160.nd	chr5.CM0696.1060.r2.a	Peroxidase OS=Populus trichocarpa
TRINITY_DN47218_c0_g1	-11.5483	chr1.LjT43005.150.nd	chr1.LjT43005.70.r2.a	pathogenesis-related family protein
TRINITY_DN42182_c0_g1	-11.5201	LjSGA_034431.0.1	LjSGA_034431.0.1	TIFY Transcription factor
TRINITY_DN50876_c1_g1	-11.4787	LjT03E06.40.nd	chr1.CM0012.70.r2.d	Protein of unknown function DUF3741
TRINITY_DN48639_c0_g2	-11.4787	chr3.CM0164.80.nd	chr3.CM0164.80.r2.d	dichloro-2,5- cyclohexadiene-1,4- diol dehydrogenase
TRINITY_DN49642_c0_g1	-11.4448	LjSGA_111612.1.2	chr4.CM0007.970.r2.a	Putative senescence- associated protein (Fragment) OS=Ipomoea nil

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (UniProt and AnnoMine)
TRINITY_DN34011_c0_g1	-11.4362	LjSGA_053993.1	LjSGA_053993.1	Beta-fructofuranosidase OS=Citrus sinensis
TRINITY_DN48126_c0_g1	-11.3968	LjT45L18.200.nd	LjT45L18.200.r2.d	Proteinase inhibitor I3, Kunitz legume
TRINITY_DN50442_c0_g1	-11.3789	LjSGA_147445.1.1	chr4.LjT13G24.170.r2.a	SJCHGC01393 protein OS=Schistosoma japonicum
TRINITY_DN49028_c1_g1	-11.3333	chr1.CM0295.60.nd	chr1.CM0295.140.r2.m	Phosphate transporter OS=Lotus japonicus
TRINITY_DN53079_c2_g1	-11.3287	LjSGA_008068.1	LjSGA_008068.1	D-galactose transporter GalP OS=Edwardsiella tarda

Table S41 Yellow stage *L. sessilifolius*: 19 contigs with greatest magnitude of log fold-change recovered from entire transcriptome.

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (Annomine and UniProt)
TRINITY_DN11722_c0_g1	11.10024	chr1.CM0141.260.nc	chr1.CM0141.260.r2.m	Expansin-A15 Ribose-phosphate pyrophosphokinase 3 OS=Capsaspora owczarzaki (strain ATCC 30864)
TRINITY_DN51824_c2_g1	10.13103	LjSGA_043454.1	LjSGA_043454.1	Homeodomain- related (Fragment) OS=Medicago
TRINITY_DN48232_c0_g1	10.02803	chr1.CM0088.1030.nc	chr1.CM0088.1030.r2.a	truncatula Calmodulin-binding protein OS=Medicago
TRINITY_DN43262_c0_g2	9.784361	chr6.LjT10104.70.nd	chr6.CM0778.70.r2.d	truncatula Gibberellin induced protein OS=Medicago
TRINITY_DN51657_c0_g1	9.652068	chr1.CM0398.130.nd	chr1.CM0398.250.r2.d	truncatula Glucan 1,3-beta- glucosidase OS=Medicago
TRINITY_DN43721_c0_g2	9.624089	LjSGA_011684.1	LjSGA_011684.1	truncatula Disease resistance response protein 1
TRINITY_DN47886_c0_g1	9.595556	LjSGA_025651.1.1	chr3.CM2163.270.r2.m	OS=Glycine max
TRINITY_DN52257_c0_g4	8.907039	chr6.CM1829.200.nc	chr6.CM0139.1580.r2.m	NA energized vacuolar membrane proton pump
TRINITY_DN46161_c0_g1	8.907039	chr4.CM0219.190.nc	chr4.CM0219.190.r2.m	CER1-like 2
TRINITY_DN40465_c0_g3	8.811031	chr1.CM0009.350.nc	chr1.CM0009.350.r2.m	NA
TRINITY_DN34483_c0_g2	8.760519	LjT26P12.80.nd	LjT26P12.100.r2.a	Heparanase-like protein OS=Medicago
TRINITY_DN47039_c0_g2	8.708175	chr4.CM0219.380.nd	chr4.CM0219.380.r2.d	truncatula
TRINITY_DN47671_c0_g1	8.708175	LjSGA_043568.1	LjSGA_043568.1	NA

Trinity_ID	LogFC	Lotus1.0_ID	Lotus2.5_ID	Annotation (UniProt and AnnoMine)
				alpha/beta-Hydrolases
TRINITY_DN40925_c0_g3	8.681273	LjSGA_031888.1.1	chr1.CM0104.3590.r2.m	superfamily protein PMR5 N-terminal
TRINITY_DN42547_c0_g1	8.681273	LjSGA_027078.2	LjSGA_027078.2	domain
				Major facilitator superfamily domain, general substrate
TRINITY_DN36333_c0_g1	8.65386	LjSGA_092888.1	LjSGA_092888.1	transporter
				xylosidase/alpha-L-arabinofuranosidase
TRINITY_DN25191_c0_g1	8.625916	chr5.CM0239.470.nd	chr5.CM0239.440.r2.d	2
TRINITY_DN44884_c1_g1	8.625916	CM0466.10.nd	chr5.CM0466.460.r2.d	NA
TRINITY_DN46831_c0_g2	8.568349	LjSGA_031037.1	chr4.CM0387.700.r2.a	NA

Table S42 Results of hand pollination experiment. Hand pollinated treatments are denoted by P, controls are denoted by N.

Treatment	Plant number	Days to PACC from anthesis (mean of 3 reps)
P	1	3.33333
P	2	3.33333
P	3	3
P	4	3
P	5	2.66667
P	6	3.33333
P	7	2.33333
P	8	2.66667
P	9	2.33333
P	10	2.33333
P	11	2.33333
P	12	2.66667
P	13	2
P	14	2
P	15	2
P	16	2.66667
P	17	2.33333
P	18	2
P	19	2
P	20	2
P	21	2
N	1	5
N	2	5.33333
N	3	4.6667
N	4	4.66667
N	5	4.66667
N	6	5

Treatment	Plant number	Days to PACC from anthesis (mean of 3 reps)
N	7	4.66667
N	8	4.33333
N	9	4.33333
N	10	4
N	11	4.33333
N	12	4.66667
N	13	4
N	14	4
N	15	4.66667
N	16	4
N	17	5
N	18	5
N	19	4.33333
N	20	5
N	21	4.66667

Mean number of days to observe colour change in hand pollinated treatment = **2.5 ± 0.2 days**

Mean number of days to observe colour change in control = **4.6 ± 0.2 days**

$p = 2.2 \times 10^{-16}$

$n = 21$

Appendix 4

Supplementary Figures S1 – S10 Amino acid alignments for non-differentially expressed carotenoid genes in *Lotus filicaulis*

Figure S1. Amino acid alignment for β -CHY

Lfil	1	-----KPYPLTLTVKSSINKNSTKPSSWVSPDWLTSLSRSLTAGKNDDSGIPI
Ljap	1	SLSLPLSLFPPKPYPLTLTVKSSINKNSTKPSSWVSPDWLTSLSRSLTAGKNDDSGIPI
Gmax	1	SLSLPLSLVFPPTPSPKRLSVKSSINKPTTKPSSWVSPDWLTSLSRSLTAG-NDVSGIPV
Mtru	1	SLPLSHLSLSSFKRYPLRLTLTKSSINKNITTKSTSWVSPDWLTSLSKSLTTSKNDDSNIPPI
Atha	1	SSSYSSLFTAKPSPKPKTFSIRSSIEKPSSKSQSWVSPDWLTTLTRLTLSSGKNDESGIPI
Lfil	49	ASAKLDDVSDLLGGALFLPLFKWMKEYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYG
Ljap	61	ASAKLDDVSDLLGGALFLPLFKWMKEYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYG
Gmax	60	ASAKLDDVSDLLGGALFLPLFKWMQDYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYA
Mtru	61	ASAKLDDVSDLLGGALFLPLFKWMNEYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYG
Atha	61	ANAKLDDVADLLGGALFLPLYKWMNEYGPIYRLAAGPRNFVTVSDPAIAKHVLKNYPKYA
Lfil	109	KGLVAEVSEFLFGSGFAIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRCAERLVEKLQ
Ljap	121	KGLVAEVSEFLFGSGFAIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRCAERLVEKLQ
Gmax	120	KGLVAEVSEFLFGSGFAIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRCAERLVEKLQ
Mtru	121	KGLVAEVSEFLFGDGFIAEGPLWTARRRAVPSLHKRYLSIMVDRVFCRCAERLVEKLQ
Atha	121	KGLVAEVSEFLFGSGFAIAEGPLWTARRRAVPSLHRRYLSVIVERVFCRCAERLVEKLQ
Lfil	169	PDALNGTAVNMEDKFSQTLTDVIGLSVFNYDFDSLNDSPVIGAVYTALKEAEARSTDLL
Ljap	181	PDALNGTAVNMEDKFSQTLTDVIGLSVFNYDFDSLNDSPVIGAVYTALKEAEARSTDLL
Gmax	180	PDALNGTAVNMEA KFSQTLTDVIGLSVFNYNFDLSLNTDSPVIEAVYTALKEAEARSTDLL
Mtru	181	ADAVNGTAVNMEDKFSQTLTDVIGLSVFNYNFDALNSDSPVIEAVYTALKEAEARSTDLL
Atha	181	PYAEDGSAVNMEA KFSQMTLDVIGLSLFNYNFDSLTTDSPVIEAVYTALKEAELRSTDLL
Lfil	229	PYWQVEFLRKIIPRQIKA EKAVSIIRKTVENLIEKCKEIVESEGERIDGDEYVNDSDPSI
Ljap	241	PYWQVEFLRKIIPRQIKAENAVTIIRKTVEDLIEKCKEIVESEGERIDGDEYVNDSDPSI
Gmax	240	PYWKFKFLCKIIPRQIKAEEAVSVIRKTVEDLIEKCKEIVESEGERIDVEEYVNDSDPSI
Mtru	241	PYWKIDFLCKIIPRQIKAENAVTVIRKTVEDLIEQCKEIVESEGERIDADEYVNDADPSI
Atha	241	PYWKIDALCKIVPRQVKA EKAVTLIRETVEDLIAKCKEIVEREGEGERINDEEYVNDADPSI
Lfil	289	LRFLLASREEVSSDQLRDDLLSMLVAGHETGTVLTWTLYLLSKDSSSLAKVQEEVDRVL
Ljap	301	LRFLLASREEVSSDQLRDDLLSMLVAGHETGTVLTWTLYLLSKDSSSLAKVQEEVDRVL
Gmax	300	LRFLLASREEVSSVQLRDDLLSLVAGHETGTVLTWTLYLLSKDSSSLAKAQEEVDRVL
Mtru	301	LRFLLASREEVSSVQLRDDLLSMLVAGHETGTVLTWTLYLLSKDSSSLAKAQEEVDRVL
Atha	301	LRFLLASREEVSSVQLRDDLLSMLVAGHETGTVLTWTLYLLSKNSSALRKAQEEVDRVL
Lfil	349	QGRRPTEFDMKSLKFLTRCITESLRLYPHPPVLIRRAQVPDELPGAYKVNAGQDIMISVY
Ljap	361	QGRRPTEFDMKNLKFLTRCITESLRLYPHPPVLIRRAQVPDELPGAYKVNAGQDIMISVY

Gmax 360 QGRRPTYEDIKDLKFLTRCIIESLRLYPHPPVLIRRAQVPDELPGGYKLDAGQDIMISVY
 Mtru 361 QGRRPTYEDMKDLKFLNRCIIESLRLYPHPPVLIRRSQIPDELPGDYKIDAGQDIMISVY
 Atha 361 EGRNPAPAFEDIKELKYITRCINESMRLYPHPPVLIRRAQVPDIILPGNYKVNTGQDIMISVY

Lfil 409 NIIHSSEVWDRAEEFMPERFDLDGMPNETNTDFRFIPFSGGPRKCVGDQFALLEATVSL
 Ljap 421 NIIHSSEVWDRAEEFLPERFDLDGMPNETNTDFRFIPFSGGPRKCVGDQFALLEATVSL
 Gmax 420 NIHRSSSEVWDRAEEFVPERFDLDGVPNETNTDFRFIPFSGGPRKCVGDQFALMEATVAL
 Mtru 421 NIIHSSKVVWDRAEEFLPERFDLDGVPNETNTDFRFIPFSGGPRKCVGDQFALLEATVAL
 Atha 421 NIHRSSSEVWEKAEEFLPERFDIDGAIIPNETNTDFKFIIPFSGGPRKCVGDQFALMEATVAL

Lfil 469 AIFLQHMFELVPDQNISMTTGATIHTTNGLYMKLSQRVK
 Ljap 481 AIFLQHMFELVPDQNISMTTGATIHTTNGLYMKLSQRVK
 Gmax 480 AIFLQHMFELVPDQNISMTTGATIHTTNGLYMKLSRRLK
 Mtru 481 AVFLQHMFELVPDQNIGMTTGATIHTTNGLYMKLSQRLK
 Atha 481 AVFLQRLNVELVPDQTIISMTTGATIHTTNGLYMKVSQR--

Figure S2. Amino acid alignment for CDS

Lfil	1	MAGCGILSAVNLNYHGGARNTIKLCSSSSSTDSSISLSFGGSESIRLG--PSTRARLRNH
Ljap	1	MAGCGILSAVNLNYHAGARNTIKLGSSSSSTDSSISLSFGGSESIRLG--PSTRARLRNH
Gmax	1	MAACGYISAANFNLYLVGARNISKFAS----SDATISFSFGGSDSMGLTLRA----PKRNH
Mtru	1	MALYGSISSPNLNWQIGPKTISKSTSSSMNSYTTISLSFSSSVSMGLNLRPSTRARLRNH
Atha	1	MVVFQNVSAANLPYQNGFLE-----ALSSGGCELMGHSFRPTSQARRRST

Lfil	59	GSPLRVVCIDYPRPQLENTVNFLEAAYLSSTFRASPRPTKPLKVVIAGAGLAGLSTAKYL
Ljap	59	GSPFRVVCIDYPRPQLENTVNFLEAAYLSSTFRASPRPAKPLKVVIAGAGLAGLSTAKYL
Gmax	53	FSPLRVVCVDYPRPELENTVNFVEAAYLSSTFRASPRPLKPLNIVIAAGAGLAGLSTAKYL
Mtru	61	GSPLKVVCIDYPRPELDDTVNFIEASYLSSTFRASPRPTKPLKVVIAGAGLAGLSTAKYL
Atha	46	AGPLQVVCVDIPRPELENTVNFLEAASLSASFSAAPRPAKPLKVVIAGAGLAGLSTAKYL

Lfil	119	ADAGHKPILLEARDVLGGKVAAWKDEGDWYETGLHIFFGAYPNVQNLFGELGINDRLQW
Ljap	119	ADAGHKPILLEARDVLGGKVAAWKDEGDWYETGLHIFFGAYPNVQNLFGELGINDRLQW
Gmax	113	ADAGHKPILLEARDVLGGKVAAWKDKGDWYETGLHIFFGAYPNVQNLFGELGINDRLQW
Mtru	121	ADAGHKPILLEARDVLGGKVAAWKDEGDWYETGLHIFFGAYPNVQNLFGELGINDRLQW
Atha	106	ADAGHKPILLEARDVLGGKIAAWKDEGDWYETGLHIFFGAYPNVQNLFGELGINDRLQW

Lfil	179	KEHSMIFAMPSKPGEFSTRFDFAEVLPA
Ljap	179	KEHSMIFAMPSKPGEFSTRFDFAEVLPA
Gmax	173	KEHSMIFAMPNKPGEFSTRFDPEVLPS
Mtru	181	KEHSMIFAMPSKPGEFSTRFDPEVLPS
Atha	166	KEHSMIFAMPSKPGEFSTRFDPEVLPA

Figure S3. Amino acid alignment for CRTISO-2

```

Lfil 1 EADV VVIGSG I GGLCCAALLARYEQ NVLVLESHDL PGGAHSFDIKGYKFDSGPSLFSGL
Ljap 1 EADV VVIGSG I GGLCCAALLARYEQ NVLVLESHDL PG-----
Gmax 1 EADV VVIGSG I GGLCCGALLARYEE DVLVLESHDQ PGGAHSFDIKGYNFDSGPSLFSGL
Mtru 1 ETDV I VIGSG I GGLS CAALLARYEQ DVVVFESH DHAGGAHSFDV KGYKFDSGPSLFSGL
Atha 1 EADV VVIGSG I GGLCCGALLARYDQ DVIVLESHDHPGGAHSF EIKGYKFDSGPSLFSGL

Lfil 61 QSRGPQANPLAQVLDALGESLP CATYDSWKVYLPEGEFLSRIGPTEFSKDLEKYGGPNAV
Ljap 38 -----SWMRWVSHSHALLM-----
Gmax 61 QSRGPQANPLAQVLDALGESLP CATYDSWMVHIPEGQFLSRIGPTEFFKDLQQYAGPNAV
Mtru 61 QSRGPQANPLAQVLDALGESVPCATYDSWMVHVPEADFLSRIGPTEFLKDLHNYAGPEAV
Atha 61 QSRGPQANPLAQVLDALGESFPCCKKYDSWMVYLPEGDFLSRIGPTDFFKDLKYAGPSAV

Lfil 121 QEWRKLLDAVLPLSTAAMALPPLSIRGDLGVLYTAAARYAPSL LKSFLQMGPQGALGATK
Ljap 52 -----
Gmax 121 LEWRKLLDAVLPLSTAAMALPPLSIRGDLGVLS TAAARYAPSLFKSFLQMGPRA SLRATK
Mtru 121 QEWQKLLDAVLPLSTAAMALPPLSVRGDFGVLYTAAARYAPSLFNTFFQMGPQAALRSTQ
Atha 121 QEWEKLLGATLPLSSAAMALPPLSIRGDLGVLS TAAARYAPSL LKSFIKMGPKGALGATK

Lfil 181 LLRPFKEILDGLELKDPPFIRNWIDLLSFLLAG-----
Ljap 52 -----
Gmax 181 LLRPFSEILD DLQLNDPFI RNWIDLLSFLLAGVKANAVLSAEMIYMF AEWKPGCCLEYP
Mtru 181 LLS PFSQILDSLQLKNPFI RNWIDLLSFLLAGVKS DSI LSAEMVYMF AEWKPGGCLEYP
Atha 181 LLRPFSEI VDSL ELKDPPFIRNWIDLLAFLLAGVKS DGI LSAEMIYMF AEWKPGCTLEYP

Lfil -----
Ljap 52 -----IHGRFIY LKVNIV
Gmax 241 LHGSAAIIDALVRGLEKFGGRISLQSHVENI VVENDRAIGVKLGSGQFIRAKKAVV
Mtru 241 LDGAAGIVDALIRGLEKFGGRISLQSHVEKI VVENDRAIGVKLRSGQFIRAKKAVV
Atha 241 IDGTGAVVEALVRGLEKFGGRISLQSHVENI VIENGKAVGVKLRNGQFVRARKAVV

```

Figure S4. Amino acid alignment for CRTISO-3

```

Lfil 1 SSSIFSSLNKFNP LSLRSP LRI R AT VVK DTEG KVVVEKTFPSKIPLQTGA EEDSDSSPPT
Ljap 1 SSSIFSSLNKFNP LSLRSP LRV R AT VVK DTEG KVVVEKTFPSKIPLQTGA EEDSDSSPPT
Gmax 1 SSSIFSP LPK ---HTAPQLRARSLLQDNEKKVIVHDSFPSKTPLHTADK-GGNSINTS
Mtru 1 CDS SSSLFTPLNKF RPLSFRPPLRIRSSLLQDKEDKVITQNTFPSKTPLDSVTENDDDTS
Atha 1 SSGTLRPLVTHLPLSSRRLLRVQATILQDDEEKVVVEESFKAETSTGTEPLENMSSSSTS

```

```

Lfil 61 AFELGV I KVEQS I NIFLTDSVIKILDALYRDRNYARFFVLETIARVPYFAFMSVLHMYES
Ljap 61 AFELGV I KVEQS I NIFLTDSVIKILDALYRDRNYARFFVLETIARVPYFAFMSVLHMYES
Gmax 56 AFEKRI I KVEQSVNIFLTDSVIKILDTLYHDRHYARFFVLETIARVPYFAFMSVLHMYES
Mtru 61 STSAWEKGV I KVEQSVNIFLTDSVIKILDALYRDRNYARFFVLETIARVPYFAFMSILHM
Atha 61 AFETWI I KLEQGVNVFLTDSVIKILDTLYRDRTYARFFVLETIARVPYFAFMSVLHMYET

```

```

Lfil 121 FGWRRADY LKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYFMTALMYTIS
Ljap 121 FGWRRADY LKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYFMTALMYTIS
Gmax 116 FGWRRADY LKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAVS
Mtru 121 YESFGWRRADY LKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYFMTALMY
Atha 121 FGWRRADY LKVHFAESWNEMHLLIMEELGNSWWFDRFLAQHIATFYFMTVFLYILS

```

```

Lfil 181 PRMAYHFSECVENHAFETYDKFIKDQ---
Ljap 181 PRMAYHFSECVENHAFETYDKFIKDQ---
Gmax 176 PRMAYHFSECVESHA FETYDKFIKVQ---
Mtru 181 LISPRMAYHFSECVESHA FETYDKFIKEQ
Atha 181 PRMAYHFSECVESHA YETYDKFLKAS---

```

Figure S5. Amino acid alignment for CYB-1

Lfil	1	-----
Ljap	1	ILVETDHIIGRQASGSVTVTDGETIVYTTVCLDDVPSEPSDFFPLSVNYQERFSAAGRTSG
Gmax	1	ILVETAHIIGRQASGSVTVTDGETIVYTTVCLDDVPCEPSDFFPLSVIYQERFSAAGRTSG
Mtru	1	ILVETGHMGRQASGSVTVTDGETIVYTTVCLDDIPSEPSDFFPLSVNYQERFSAAGRTSG
Atha	1	ILVETGLMGRQASSAVTVTDGETIVYTSVCLADVPSEPSDFLLPLYVHYQERFSAVGRRTSG
Lfil	1	-----
Ljap	61	GFFKREGKTKDHEVLICRLIDRPLRPSMPKGFYHETQILSWVLSYDGLHSPDSLAVTAAG
Gmax	61	GFFKREGKTKDHEVLICRLIDRPLRPTMPKGFYHETQILSWVLSYDGLHSPDSLAVTAAG
Mtru	61	GFFKREGKTKDHEVLICRLIDRPLRPTMPKGFYHETQILSWVLSYDGLHIPDSLAVTAAG
Atha	61	GFFKREGRTKDHEVLICRLIDRPLRPTMPKGFYNETQILSWVLSYDGLHAPDALAVTSAG
Lfil	1	-----
Ljap	121	IAValsevPMSKTVAGVRIGLVGDKFLVNPTTEEMENSELDLILLAGTDSAILMIEGYSNF
Gmax	121	IAValsevPMSKAVAGVRVGLVGDKYIVNPTTEEMEHSELDDLILLAGTDNAILMIEGYSNF
Mtru	121	IAIALSEVPTSKAVAGVRVGLVGDKYIVNPTSEEMENSELDLMLAGTDSAILMIEGYGNF
Atha	121	IAValsevPNAKATAGVRVGLIGGEFIVNPTVKEMEESQLDLFLAGTDTAILTIEGYSNF
Lfil	1	-----
Ljap	181	LPEEKLLAEVEVGQDAVRAICKEVAALVKKCGKPKMFDAIKLPPPELYKHVEEIIAGDELV
Gmax	181	LPEEKLLKAVEVGQNAVRAICNEVEALVKKCGKPKMLDAIKLPPPELYEHVEAIIASDELV
Mtru	181	LPEEKLLKAVEVGQDAVRAICNEVEALVKKCGKPKMVDIAIKLPPPELYKHVEEIIAGDELV
Atha	181	LPEEMLLQAVKVGQDAVQATCIAIEVLAKKYGKPKMLDAIRLPPPELYKHVKELAGEELT
Lfil	1	-----
Ljap	241	KVLQIRNKIPRRKALSSLEEKVLIKILTENGVTNDATQRSNTETIAEILEDEDEDEEVIV
Gmax	241	KVLQIRNKIPRRKALSSLEEKVLKILTENGFVSNSTLRNSNETIAEILEVEDEDEEVIV
Mtru	241	KVLQIRNKIPRRKALSSLEEKVLKILSENGFVTNDTAPRNNAETIAEILEDEDEDEEVIV
Atha	241	KALQIKSKISRKAISLEEKVLTILTEKGYVIDEVA-FGTIEAQPDLLEDEDEDEEVIVP
Lfil	1	-----VTSKFLRKRIVEGGKRS DGRKPN EIRPIN
Ljap	301	DGEVDEGDVHIKPTARKPLFSEVDVKLVFKEVTSKFLRKRIVEGGKRS DGRKPN EIRPIN
Gmax	301	DGEVDEGDVHIKPTPRKPLFSEVDVKLVFKEVTSKYWRKRIVEGGKRS DGRTPDGIRPIY
Mtru	301	DGEVDEGDVHIKPTPRKPLFSEVDVKLVFKDVTSKFLRKRIVEGGKRS DGRTPNEIRPIN
Atha	300	EGEVDQGDVHIRPIPRKPLFSEVDVKLVFKEVSSKLLRRRIVEGGKRS DGRTLDEIRPIN
Lfil	30	SRCGLLPRAHGSTLFTRGETQSLAVVTLGDKQMAQRIDNLMGVDEFKKFYLYQYSFPPSCV
Ljap	361	SRCGLLPRAHGSTLFTRGETQSLAVVTLGDKQMAQRIDNLMGVDEFKKFYLYQYSFPPSCV
Gmax	361	SRCGLLPRAHGSTLFTRGETQALAVVTLGDKQMAQRIDSIDGVDEFKRFYLYQYSFPPSCV
Mtru	361	SSCGLLPRAHGSALFTRGETQSLAVATLGDNKMAQRIDNLMVDVDYKRFYLYQYSFPPSCV
Atha	360	SRCGLLPRAHGSTLFTRGETQALAVVTLGDKQMAQRIDNLEGSDEYKRFYLYQYTFPPSSV
Lfil	90	GEVGRIGAPSRREIGHGMLAERSLEPILPSDKDFPYTIRVESTITESNGSSSMASVCGGC
Ljap	421	GEVGRIGAPSRREIGHGMLAERSLEPILPSDKDFPYTIRVESTITESNGSSSMASVCGGC
Gmax	421	GEVGRIGAPSRREIGHGMLAERSLEPILPSDDFPYTIRVESTITESNGSSSMASVCGGC
Mtru	421	GEVGRIGAPSRREIGHGMLAERSLETILPSDKDFPYTIRVESTITESNGSSSMASVCGGC
Atha	420	GEVGRIGAPSRREIGHGTLAERALETILPSDDFPYTIRVESTIVIESNGSSSMASVCGGC

```

Lfil 150 LALQDAGVPIKSAIAGIAMGMVLDTKEFGGDGTPILISDITGSEDASGDMDFKVAGNEHG
Ljap 481 LALQDAGVPIKSAIAGIAMGMVLDTKEFGGDGTPILISDITGSEDASGDMDFKVAGNEHG
Gmax 481 LALQDAGVPIKSGIAGIAMGMVLDTKEFGGDGTPILISDITGSEDASGDMDFKVAGNEDG
Mtru 481 LALQDAGVPIKSSIAGIAMGIIVLDTKEFGGDGTPILISDITGSEDASGDMDFKVAGNEDG
Atha 480 LALQDAGVPVKCSVAGIAMGMVWDTTEFGGDGSPILISDITGAEDASGDMDFKVAGNEDG

```

```

Lfil 210 ITAFQMDIKVGGITLPVMREALLQARDGRKHILGEMMNCSPPPAKIILSKYAPLIHIMKVR
Ljap 541 ITAFQMDIKVGGITLPVMREALLQARDGRKHILGEMMNCSPPPAKIILSKYAPLIHIMKVR
Gmax 541 ITAFQMDIKVGGITLPIMRKALLQARDGRKHILGAMNNSPPPAKRLSKYAPLIHVMKVR
Mtru 541 ITAFQMDIKVGGITLAIMREALLQAKDGRKHILGEMMKCLPPPAKRLSKYAPLIHVMKVR
Atha 540 VTAFQMDIKVGGITLIEIMEKALIQAKAGRRHILAEMAKCSPPPTLSLSKYAPLILIMKVH
Lfil 270 PDKVNLIIGSGGKKVKSIIIEESGIEAIDTDDNGTVKIFARDLSSSLERSKAMISNLTMVPA
Ljap 601 PDKVNLIIGSGGKKVKSIIIEESGIEAIDTDDNGTVKIFARDLSSSLERSKAMISNLTMVPA
Gmax 601 PDRINLIIGSGGKKVKSIIIEESGVDITDTEDDGTVKIFAKDLPSLEKSKAIINSNLTMVPT
Mtru 601 PDKINLIIGSGGKKVKSIIIEQSGIEAIDTEDDNGTVKIFARDLASLEMSKAIINSNLTMVPT
Atha 600 PSKVYSLIGSGGKKVKSIIIEESGVEAIDMQDDGTVKIMAITDVASLERAKAIISGLTMVPS

```

```

Lfil 330 IGDYRNCEIKSIAPYGVFVEIAPGREGLCHISELSSGWLPKAEDVFKVGDRIDVKLIEI
Ljap 661 IGDYRNCEIKSIAPYGVFVEIAPGREGLCHISELSSGWLPKAEDVFKVGDRIDVKLIEI
Gmax 661 IGDYRSCEIKSIASISYGVFVEIAPGREGLCHISELSSGWLPKAEDVFKVGDRIDVKLIEI
Mtru 661 IGDYRNCEIKSIVPYGAFVEIAPGREGLCHISELSSAWLAKAEDAFKVGDRIDVKLIEI
Atha 660 VGDYRNCEIKSMAPYGAFAVEIAPGREGLCHISELSAEWLAKPEDAYKVGDRIDVKLIEV

```

```

Lfil 390 NDKGQLRLSHRALLPDTDPDGLIAEEKIEQPKDKINTPKGDSSKRSSRDDSDALPSKKFIRR
Ljap 721 NDKGQLRLSHRALLPDTDPDGLIAEEKIEQPKDKINTPKGDSSKRSSRDDSDALPSKKFIRR
Gmax 721 NEKGQLRLSHRALLPDTEASGLVEEKEEQSKDKSSTPNVTSSKRSEDDSVLPSKKFVRR
Mtru 721 NDKGQLRLSHRALLPDADPDGLVEEKEEQN-AKVNTRKGRSSKRSEDDNSALPSKKFIRR
Atha 720 NEKGQLRLSVRALLPESETDDSTKDKSSQR-KYVNTSSKDAKSVSSGDELVL-KKKDVRR

```

```

Lfil 450 MVSPSQDKPVTNKDKIKRSSNKASDKDESSLVGEAA
Ljap 781 MVSPSQDKPVTNKDKIKRSSNKASDKDESSLVSEEA
Gmax 781 LVSSSLDKPFETNKDKTKKSSNKASGKDESSLVSGEA
Mtru 780 S---SEDKPVTNKDKIKKSDN----KDESSLVSEEA
Atha 778 ATGGSSDKTMNS-----SSTNEESLVNGEA

```

Figure S6. Amino acid alignment for CYB-2

Lfil	1	LRASSALLELVPEFKKENLDFELPLYDSTKGAVVNLAVVGGGPAGLAVAQQVSEAGLSV
Ljap	1	LRASSALLELVPEVKKENLDFELPLYDSMKGAVVDLAVVGGGPAGLAVAQQVSEAGLSV
Gmax	1	LRASSALLELVPEFKKENLDFELPLYDSSKGAMVDLAVVGGGPAGLAVAQQVSEAGLSV
Mtru	1	LRASSALLELVPEFKKENLDFELPLYDSTKGTIVVDLVVGGGPAGLAVAQQVSEAGLSV
Atha	1	VVSGSAALLDLVPETKKKENLDFELPLYDTSKSQLVVDLAVVGGGPAGLAVAQQVSEAGLSV
Lfil	61	CAIDPNPRLIWPNNYGVWVDEFEAMDLLDCLDTTWSGAVVYIDDKTKKYLDLPYGRVNRK
Ljap	61	CAIDPNPRLIWPNNYGVWVDEFEAMDLLDCLDTTWSGAVVYIDDKTKKYLDLPYGRVNRK
Gmax	61	CAIDPNPRLIWPNNYGVWVDEFEAMDLLDCLDTTWSGAVVYIDDKTKKDLDRPYGRVNRK
Mtru	61	CAIDPNPRLIWPNNYGVWVDEFEAMDLLDCLDKTWSGAVVYIDDKTKKDLDRPYGRVNRK
Atha	61	CSIDPSPKLIWPNNYGVWVDEFEAMDLLDCLDTTWSGAVVYVDEGVKKDLSPRYGRVNRK
Lfil	121	LLKSKMLQKCISNGVKFHQAQVIKVIHEESKSMLICNDGVTIQATVVLDATGFSRSLVQY
Ljap	121	LLKSKMLQKCISNGVKFHQAQVIKVIHEESKSMLICNDGVTIQATVVLDATGFSRSLVQY
Gmax	121	LLKSKMLQKCISNGVKFHQAQVIKVIHEEAKSLICNDGVTVHATVVLDATGFSRCLVQY
Mtru	121	LLKSKMLQKCIANGVKFHQAQVIKVIHEESKSMLICNDGVTIQATVVLDATGFSRCLVQY
Atha	121	QLKSKMLQKCITNGVKFHQS KVTNVVHEEANSTVVCSDGVKIQASVVLDATGFSRCLVQY
Lfil	181	DKPYNPGYQVAYGILAEVEEHFPDVKMLFMDWRDShLDNNMELKEKNSRIPTFLYAMPF
Ljap	181	DKPYNPGYQVAYGILAEVEEHFPDVKMLFMDWRDShLDNNMELKEKNSRIPTFLYAMPF
Gmax	181	DKPYNPGYQVAYGILAEVDEHPFDVKMLFMDWRDShLDNDMELKQNSRIPTFLYAMPF
Mtru	181	DKPYNPGYQVAYGILAEVEEHFPDVKMLFMDWRDShLDNDLKLKERNKSKIPTFLYAMPF
Atha	181	DKPYNPGYQVAYGIVAEVDGHPFDVKMVFMDWRDKHLDSPYELKERNKSKIPTFLYAMPF
Lfil	241	SSTKIFLEETSLVARPGLRMDDIQDRMVARLKHGLGKVKSEIEDEHCVIPMGGPLPVLPO
Ljap	241	SSTKIFLEETSLVARPGLRMDDIQDRMVARLKHGLGKVKSEIEDEHCVIPMGGPLPVLPO
Gmax	241	SSTKIFLEETSLVARPGLRMDDIQERMAARLKHGLGIRVKSEIEDEHCVIPMGGPLPVLPO
Mtru	241	SSTKIFLEETSLVARPGLRMDDIQDRMVARLKHGLGINVKSEIEDEQCVIPMGGPLPVLPO
Atha	241	SSNRIFLEETSLVARPGLRMDIQERMAARLKHGLGINVKRIEEDERCVIPMGGPLPVLPO
Lfil	301	RVIGIGGTAGMVHPSTGYMVARTLAAPIVANAIVQYLGSDRGISGDGLSSLVWKDLWPI
Ljap	301	RVIGIGGTAGMVHPSTGYMVARTLAAPIVANAIVQYLGSDRGISGDGLSSLVWKDLWPI
Gmax	301	RVVGIGGTAGMVHPSTGYMVARTLAAPIVANSIVQCLGSDRGFSGDEISAQVWKDLWPI
Mtru	301	RVVGIGGTAGMVHPSTGYMVARTLAAPIVANAIVQYLGSDRGILGDEV SARVWKDLWPI
Atha	301	RVVGIGGTAGMVHPSTGYMVARTLAAPIVANAIVRYLGSPNSLRGDQLSAEVWRDLWPI
Lfil	361	ERRRQREFFFCFGMDILLKLDLPGTRRFFDAFFDLEPHYWHGFLSSRLFLPELLVFGLSLF
Ljap	361	ERRRQREFFFCFGMDILLKLDLPGTRRFFDAFFDLEPHYWHGFLSSRLFLPELLVFGLSLF
Gmax	361	QRRRQREFFFCFGMDILLKLDLPGTKRFFDAFFDLEPHYWHGFLSSRLFLHELLVFGLSLF
Mtru	361	ERRRQREFFFCFGMDVLLKLDLPGTRRFFDAFFNLEPHYWHGFLSSRLYLPELFTFGLSLF
Atha	361	ERRRQREFFFCFGMDILLKLDLPATRRFFDAFFDLQPHYWHGFLSSRLFLPELLVFGLSLF
Lfil	421	SYASNTSRIEIMAKGTVP
Ljap	421	SYASNTSRIEIMAKGTVP
Gmax	421	SYASNTSRVEIMAKGTLP
Mtru	421	SYASNTSRLEIMAKGTLP
Atha	421	SHASNTSRLEIMTKGTVP

Figure S7. Amino acid alignment for CYE

Lfil	1	MLLLQLQPFNGSHLTVGTPFPFLKKCSRRTLLLRPQAVPSRTQRIMESVSVSGEVGGAGGA
Ljap	1	MLLLQLQPFNGSHLTVGTPFPFQKKCSRRTLLLRPQAVPSRTQRIMESVSVSGEVGGAGGA
Gmax	1	MLLLRLQPIINGSQLSVGTSFPHKKCSRRLPLSRPQAVPSRTQRIMESVSVSGEVGGAGGA
Mtru	1	MSLLQLQPIINGSQLLV-----QKKCSRKAFFRAQAVPSKTQRIMESVSVSGEVGGAGGA
Atha	1	MVVLQIHQFNGSSVSRAQTNQGRRRASTTVCVQTQLAPSLTQKIMESISVGGGEAGGAGGA
Lfil	61	YSYEALKRLDQLWSSICSTQEVVQEPQQVVSIPSLFRSSDLANKAEGSYDVVVCGGTILG
Ljap	61	YSYEALKRLDQLWSSICSAQEVVQEPQQVVSIPSLFRSSDLANKAEGSYDVVVCGGTILG
Gmax	61	YSYEALKRLDQQWSSICSSQEVVQEPQQVVSIPSLFGSSDLADKAEGSYDVIVCGGTILG
Mtru	56	YSYEALKRLDQLWSNICSPQEVVQEPQQIVSTIPSLFTSSDLADKIEGSDVIVCGGTILG
Atha	61	YSYNALKRLDNIWSNICIQPTGPGQETQQIVSRVSGFSQDYSMGNNLVGTFDIIVCGGTILG
Lfil	121	IFIATALCARGLKVAIVERNVLKGREQDWNISRKELLEVEVGVLEEDDIERATSINFNPN
Ljap	121	IFIATALCARGLKVAIVERNVLKGREQDWNISRKELLEVEVGVLEEDDIERATSINFNPN
Gmax	121	IFIATALCARGLRVAIVERNVLKGRDQEWNIISTKELLEVEVGVLEEDDIERATAVKFNPN
Mtru	116	IFIATALCARGLRVAIVERNVLKGREQEWNIISRKELLEVEVGILEEDDIERVTTAKFNPN
Atha	121	IFIATALCAKGLRVAIVERNATKGRDQEWNIISRKEMKELTEVRVLTEDEIEEVIAAKFNPN
Lfil	181	NRCGFESKGDIVVNNILHLGVSPVKLLEIVKERFISLGGVIFEGCSVSCINIYEDA AVLK
Ljap	181	NRCGFESKGDIVVNNILHLGVSPVKLLEIVKERFISLGGVIFEGCSVSCINIYEDA AVLK
Gmax	181	NRCGFERKGDIVVNDILNLGVSPAKLIEIVKTRFISLGGVIFEGCVSCINVYEDA AVLK
Mtru	176	NRCGFESKGDIVVNNILNLGVSPVRLIEIVKKRFISLGGVFEGYSVSFINVYEDA AVLK
Atha	181	NRCGFENLGDIVVEDILNLGVSPAKLVETVKQRFISLGGVILEDSSISSIVILYNDLAVMQ
Lfil	241	LSGDKILSSRLIIDAMGNFSPIVKQIRGRKPDGVCLVVGTCARGFESNSTSDVIFSSSS
Ljap	241	LSGDKILSSRLIIDAMGNFSPIVKQIRGRKPDGVCLVVGTCARGFESNSTSDVIFSSSS
Gmax	241	LSGDKILSSRLIIDAMGNFSPVVKQIRGGRKPDGVCLVVGTCARGFETNSTSDVIFSSSS
Mtru	236	LSGGKVLSSRLIIDAMGNFSPVVKQIRGRKPDGVCLVVGTCARGFENNNSNDVIFSSSS
Atha	241	LSKGDTLSSRLVIDAMGNFSPILKQIKRGRKPDGMCLVVGSCAHGFKENSSSSDVIYSSSS
Lfil	301	VKKVGDSQAQYFWEAFPAASGPLDRTTYMFTYVEAQPGSPKLEELLEAYWDLMPYQGVS
Ljap	301	VKKVGDSQAQYFWEAFPAASGPLDRTTYMFTYVEAQPGSPKLEELLEAYWDLMPYQGVS
Gmax	301	IKKVGDSKAQLFWEAFPAAGSGPLDRTTYLFTYVEQPASPKLEELLEAYWDLMPYQGVS
Mtru	296	VKRVGDSKAQYFWEAFPAAGSGPLDRTTYMFTYVEQPQPGSPKLEELLEAYWDLMPYQGV
Atha	301	VTRVADSNVQLFWEAFPAAGSGPLDRTTYMFTYTEPQSTSPSLEDLLEEYWKLMPKYQGV
Lfil	361	LDNLEILRVIYGIFPTYRESPLPAAFSRVLQFGDASGIQSPVSFGGFGSLTRHLGRISAG
Ljap	361	LDNLEILRVIYGIFPTYRESPLPAAFCRVLQFGDASGIQSPVSFGGFGSLTRHLGRISAG
Gmax	361	LDNLEILRVIYGIFPTYRQSPLPAAFSRVLQFGDASGIQSPVSFGGFGSLTRHLGRISAG
Mtru	356	LDNLEILRVIYGIFPTYRESPLPAAFSRVLQFGDASGIQSPVSFGGFGSLTRHLGRISAG
Atha	361	LDELEILRVVYGIFPTYRNSPLPAAFDRVLQFGDASGIQSPVSFGGFGSLTRHLGRISNG
Lfil	421	IYEAINGDYLDSDYNLSLLNPYPNLSASWLFQRAMSAKKQSNVPPDFINELLYANFSCMQ
Ljap	421	IYEAINGDYLDSDYNLSLLNPYPNLSASWLFQRAMSAKKQSNVPPDFINELLYANFSCMQ
Gmax	421	IHEAINGDYLDSDYSLSLLNPYPNLSASWLFQRAMSAKKQSNVPADFINELLYTNFSCMQ
Mtru	416	IQEAIIDGYLDSDYNLSLLNPYPNLSASWLFQRAMSAKKQSDVPEDFINELLYANFSCMQ
Atha	421	IYDAIDGDLSDSLSKLNYPYPNLSASWLFQRAMSAKKQLDVSRGFTNELLHVNFSCMQ

Lfil	481	RLGDPVLRPFLQDVVQFGPLSKTLGLVMLTKPQILPSIFK
Ljap	481	RLGDPVLRPFLQDVVQFGPLSKTLGLVMLTKPQILPSIFK
Gmax	481	RLGDPVLRPFLQDVVQFGPLSKTLGLVMLTNPKILPSIFK
Mtru	476	RLGDPVLRPFLQDVVQFGALSKTLGLVMLTRPQILPSIFK
Atha	481	RLGDPVLRPFLQDIQFGPLAKTLGLVMLTKPQIIPSIIR

Figure S8. Amino acid alignment for E-CHY

Lfil	1	MASHAALLHVPPPLSIPTQRFYHKQICINRLKPSTTSSSSSWFPCSSSSVIVCSSSNGKS
Ljap	1	MASHAALLHVPPPLSIPTQRFYHKQICINRLKPSTTSSSSSWFPCSSSSVIVCSSSNGKS
Gmax	1	MASHVALLRVPPPLSISTQRFHAKQICINGLKLNTSSSSSCFPCSCSSVITCSSSNGRD
Mtru	1	MASHLTLHLHAPPPLSLQTKTFHISKYITIKPLKPTTTSSSCSLFPCSCSSFIACSSSNGRS
Atha	1	MAMAFPL-SYTPTI-----TVKP-----V-SNFVVFSSSSNGRD
Lfil	61	SDSGDDG-----IKSKQRAELSARIAKGEFTVNKQSGFPSDLKKSLSQVGVNPNEILDFL
Ljap	61	SDSGDDG-----IKSKQRAELSARIAKGEFTVNKQSGFPSDLKKSLSQVGVNPNEILDFL
Gmax	61	PNSVDIEEKQVERILEEKRRRAALSAKIASGEFTVKQKSGLLSI-MEGLAKVGVNPNEVLEFL
Mtru	61	PDSVDDGKSADQLLEEKRRRAELSAKIASGEFTVKQESGLPSILKKSLSNLGVSNNEILEFL
Atha	33	PNSVPNGKSLKLEKQEEKRRRAELSARIAAGFTVRKS-SFPSTVKNGLSKIGIPSNVLDLFM
Lfil	115	FGWIEGGGYYPKVPEAKGSINAVRNEAFFIPLYELYLTYYGGIFRLTFGPKSFLIVSDPSI
Ljap	115	FGWIEGGGYYPKVPEAKGSINAVRNEAFFIPLYELYLTYYGGIFRLTFGPKSFLIVSDPSI
Gmax	120	FGWIEGGGEHPKIPKAVRSVAFIPLYELYLTYYGGIFRLTFGPKSFLIVSDPSI
Mtru	121	FG-----LYPKIPEAKGSISATRSEAFFIPLYELYITYGGIFRLNFGPKSFLIVSDPAI
Atha	92	FDWTGSDQDYPKVPEAKGSIQAVRNEAFFIPLYELFLTYGGIFRLTFGPKSFLIVSDPSI
Lfil	175	AKHILKDNAKGYSKGILAEILDVFMGKGLIPADGEIWRVRRRAIVPALHQKYVAAMIGLF
Ljap	175	AKHILKDNAKGYSKGILAEILDVFMGKGLIPADGEIWRVRRRAIVPALHQKYVAAMIGLF
Gmax	180	AKHILRENSKAYSKGILAEILDVFMGKGLIPADGEIWRVRRRAIVPALHQKYVAAMIGLF
Mtru	175	AKHILKDNSKAYSKGILAEILDVFMGKGLIPADGEIWRVRRRTIVPALHLKFAAMIGLF
Atha	152	AKHILKDNAKAYSKGILAEILDVFMGKGLIPADGEIWRVRRRAIVPALHQKYVAAMISLF
Lfil	235	GQATDRLCKKLDTAASDGEDVEMESLFSRLTLDVIGKAVFNDFDLSLNDTGII EAVYNV
Ljap	235	GQATDRLCKKLDTAASDGEDVEMESLFSRLTLDVIGKAVFNDFDLSLNDTGII EAVYNV
Gmax	240	GQAADRLCQKLDAAASDGEDVEMESLFSRLTLDIIGKAVFNDFDLSLNDTGII EAVYTV
Mtru	235	GQATDRLCQKLDTAASDGEDVEMESLFSRLTLDVIGKAVFNDFDLSLNDTGII EAVYTV
Atha	212	GEASDRLCQKLDAAALKGEEVEMESLFSRLTLDIIGKAVFNDFDSLINDTGVI EAVYTV
Lfil	295	LREAEDRSVSPIPTWEIPIWKDISPRQKKVNAALKLVNDTLDLIAICKRMVDEEELQFH
Ljap	295	LREAEDRSVSPIPTWEIPIWKDISPRQKKVNAALKLVNDTLDLIAICKRMVDEEELQFH
Gmax	300	LREAEDRSVAPIPVWEIPIWKDVSPRLRKVNAALKLINDTLDLIAICKRMVDEEELQFH
Mtru	295	LREAEDRSISPIPVWDLPIWKDISPRQKVTAALKLVNDTLNNLIAICKRMVDEEELQFH
Atha	272	LREAEDRSVSPIPVWDIPIWKDISPRQKVATSLKLINDTLDLIAICKRMVEEELQFH
Lfil	355	EEYMNEQDPSILHFLLAAGDDVSSKQLRDDLMTILIAGHETSAAVLTWTFYLLSKEPSVM
Ljap	355	EEYMNEQDPSILHFLLAAGDDVSSKQLRDDLMTILIAGHETSAAVLTWTFYLLSKEPSVM
Gmax	360	EEYMNEQDPSILHFLLASGDDVSSKQLRDDLMTMLIAGHETSAAVLTWTFYLLSKEPRVM
Mtru	355	EEYMNEQDPSILHFLLASGDDVTISKQLRDDLMTMLIAGHETSAAVLTWTFYLLSKEPSVM
Atha	332	EEYMNERDPSILHFLLASGDDVSSKQLRDDLMTMLIAGHETSAAVLTWTFYLLTTEPSV
Lfil	415	SKLQEEVDSVLGDRFPTIEDMKKLYTTRVINESLRLYPQPPVLIRRSIDDDVLGEYPIK
Ljap	415	SKLQEEVDSVLGDRFPTIEDMKKLYTTRVINESLRLYPQPPVLIRRSIDDDVLGEYPIK
Gmax	420	SKLQEEVDSVLGDQYPTIEDMKKLYTTRVINESLRLYPQPPVLIRRSIEDDVLGEYPIK
Mtru	415	SKLQEEVDSVLGDRFPTIEDMKKLYTTRVINESLRLYPQPPVLIRRSIEDDVLGEYPIK
Atha	392	AKLQEEVDSVIGDRFPTIQDMKKLYTTRVMNESLRLYPQPPVLIRRSIDNDILGEYPIK

```

Lfil  475  KGEDIFISLWNLHRSPLWDDADKFEPERWPLDGNPNNETNQNFRLPFGGGQRKCI GDM
Ljap  475  KGEDIFISLWNLHRSPLWDDADKFEPERWPLDGNPNNETNQNFRLPFGGGQRKCI GDM
Gmax  480  RNEDIFISVWNLHRSPLWDDADKFEPERWALDGPSPNETNQNFKYLPFGGGPRKCVGDL
Mtru  475  RGEDIFISVWNLHRSPTLWNLADKFEPERWPLDGNPNNETNQGFKYLPFGGGPRKCI GDM
Atha  452  RGEDIFISVWNLHRSPLHWDDAEKFNPERWPLDGNPNNETNQNFSLPFGGGPRKCI GDM

```

```

Lfil  535  FATFETVVALAMLVRRFNFQMAIGAPPVVMTTGATIHTTQGLNMTVTRRIQPPIVPSLHM
Ljap  535  FATFETVVALAMLVRRFNFQMAIGAPPVVMTTGATIHTTQGLNMTVTRRIQPPIVPSLHM
Gmax  540  FASYETVVALAMLVRRFNFQIIVGAPPVEMTTGATIHTTQGLKMTVTHRIKPPIVPSLQM
Mtru  535  FASYEVVVALAMLVRRFNFQMAV GAPPVVMTTGATIHTTQGLNMTVTRRIKPPIVPSLQM
Atha  512  FASFENVVAIAMLIIRRFNFQIAPGAPPVKMTTGATIHTTEGLKLTVTKRTKPLDIPSVPI

```

```

Lfil  595  STVEVDQKDQVYQAQ
Ljap  595  STVEVDQKDQVYQAQ
Gmax  600  STLEVDQKGEVYQAQ
Mtru  595  STLEVDQKDQVYQAQ
Atha  572  LPMD--SRDEVSSAL

```

Figure S9. Amino acid alignment for ZEP-1

Lfil	1	LCYNSLNPSTAAFSRTNFSVPSNKHSL	ETSPFSYVHRLGRATKQGKKVTHVKA	AAVAERE
Ljap	1	LCYNSLNPSTSAFSRTNFSVPSNKHSL	ETSPFSYVHRLGRATKQGKKVTHVKA	AAVAERE
Gmax	1	LCYNSLNPSTTVFSRTHFSVPLNKELPL	DASPGYNCGVGRTRKQRKKVMHVKCA	VV--E
Mtru	1	LSHKCLSPSMTTFSRTHFSNP	-----FSYGN--RTTKQRRKLMQV	KATVM-HE
Atha	1	FCY-SINPSKLDFTTRTHVFS	SPVSKQFYLDLSSFGKPGGV-SGFR	SRALLGVKAALVEKE

Lfil	61	APKKQLRVLVAGGGIGGLVFALA	AKKKGFEVVVFEKDLSA	VRGEGQYRGPIQIQSNALAA
Ljap	61	APKKQLRVLVAGGGIGGLVFALA	AKKKGFEVVVFEKDLSA	VRGEGQYRGPIQIQSNALAA
Gmax	59	APKKQLRVLVAGGGIGGLVFALA	AKRKGFVVMFEKDLSA	IRGEGQYRGPIQIQSNALAA
Mtru	46	APKKQLKVLVAGGGIGGLVFALA	AKRKGFVVMFEKDLSA	IRGEGQYRGPIQIQSNALAA
Atha	59	EKKKKSRVLVAGGGIGGLVFALA	AKKKGFDVIVFEKDLSA	IRGEGQYRGPIQIQSNALAA

Lfil	121	LEAID	-----
Ljap	121	LEAIDP	QVADEVMLWCITGDRINGL
Gmax	119	LEAIDSE	VADEVMRVGCITGDRINGL
Mtru	106	LEAIDMN	VADEVMRVGCITGDRINGL
Atha	119	LEAIDIE	VAEQVMEAGCITGDRINGL

Figure S10. Amino acid alignment for ZEP-2

Lfil	1	----KIREYDSKRLVKEHFKRLSGKELPIKSAQVTGSTDFSELQDKEPWLSSSKLVVKPD
Ljap	1	MARKKIREYDSKRLVKEHFKRLSGKELPIKSAQVTGSTDFSELQDKEPWLSSSKLVVKPD
Gmax	1	MARKKIREYDSKRLVKEHFKRLSGQELPIKSAQVTESTNFSELAEKEPWLSSSKLVVKPD
Mtru	1	MARKKIREYDSKRLVKEHFKRLSGDLPIKSAQVTEATDFTELQDKEQWLSSSKLVVKPD
Atha	1	MARKKIREYDSKRLVKEHFKRLSGQELPIRSVQINQETDLNELVEREPWLSSEKLVVKPD
Lfil	57	MLFGKRGKSGLVALNLDFAQVASFVKERLGKEVEMGGCKGPITTFIVEPFIPHNEEFYLN
Ljap	61	MLFGKRGKSGLVALNLDFAQVASFVKERLGKEVEMGGCKGPITTFIVEPFIPHNEEFYLN
Gmax	61	MLFGKRGKSGLVALNLDLAEVDSFVKERLGKEVEMGGCKGPITTFIVEPFIPHNEEFYLN
Mtru	61	MLFGKRGKSGLVALNLDLAEQVASFVKERLGKEVEMGGCKGPITTFIVEPFIPHNEEFYLN
Atha	61	MLFGKRGKSGLVALNLDFAVATFVKERLGKEVEMSGCKGPITTFIVEPFIPHNEEFYLN
Lfil	117	IVSERLGNSISFSECGGIDIEENWDKVKTVFIPTGVSLTSEIVAPLVATLPLEIKGEIEE
Ljap	121	IVSERLGNSISFSECGGIDIEENWDKVKTVFIPTGVSLTSEIVAPLVATLPLEIKGEIEE
Gmax	121	IVSERLGNSISFSECGGIEEENWDKVKTVFMPTGVSLTSESIAPLVATLPLEIKGEIEE
Mtru	121	IVSDRLGNSISFSECGGIDIEENWDKVKTVFIPTGVSLTSEIIAPLVATLPLEIKGEIEE
Atha	121	IVSDRLGCSISFSECGGIDIEENWDKVKTTIPTGASLTFEICAPLVATLPLEIKGEIED
Lfil	177	FLKVIFTLFDLDFTFLEMNPFITLVDGKPYPLDMRGELDDTAAAFKNFKKWGNIEFPLPFG
Ljap	181	FLKVIFTLFDLDFTFLEMNPFITLVDGKPCPLDMRGELDDTAAAFKNFKKWGNIEFPLPFG
Gmax	181	FLKVIFTLFDLDFTFLEMNPFITLVNGKPYPLDMRGELDDTAAAFKNFKKWGNIEFPLPFG
Mtru	181	FLKVIFTLFDLDFTFLEMNPFITLVDGKPYPLDMRGELDDTAAAFKNFKKWGNIEFPLPFG
Atha	181	FIQVIFTLFEIDLDFTFLEMNPFITLVDGKPYPLDMRGELDDTAAAFKNFKKWGDIEFPMFPFG
Lfil	237	RVMSATESFVHELDEKTSASLKFTVLNPEGRIWTMVAGGGASVIYADTVGDLGYAAELGN
Ljap	241	RVMSATESFVHELDEKTSASLKFTVLNPKGRIWTMVAGGGASVIYADTVGDLGFASELGN
Gmax	241	RVMSATEAFIHGLDEKTSASLKFTVLNPMGRIWTMVAGGGASVIYADTVGDLGYAPELGN
Mtru	241	RVMSATESFIHGLDEKTSASLKFTVLNPKGRIWTMVAGGGASVIYADTVGDLGFANELGN
Atha	241	RVMSSTESFIHGLDEKTSASLKFTVLNPKGRIWTMVAGGGASVIYADTVGDLGYASELGN
Lfil	297	YAEYSGAPKEDEVLYARVVIDCATANPDGQKRALVIGGGIANFTDVAATFSGIIRALKE
Ljap	301	YAEYSGAPKEDEVLYARVVIDCATANPDGQKRALVIGGGIANFTDVAATFSGIIRALKE
Gmax	301	YAEYSGAPKEDEVLYARVVIDCATSNPDGQKRALVIGGGIANFTDVAATFSGIIRALKE
Mtru	301	YAEYSGAPNEFEVLQYARVVIDCATANPDGQKRALVIGGGIANFTDVAATFSGIIRALKE
Atha	301	YAEYSGAPKEDEVLYARVVIDCATANPDGKSRALVIGGGIANFTDVAATFNIGIIRALKE
Lfil	357	KEQKLKEAKMHLVYVRGGPNYQKGLAKMRALGEEIGIPIEVYGPEATMTGICKEAIQCIT
Ljap	361	KESKLKAARMHIYVRGGPNYQKGLALMRALGEEIGIPIEVYGPEATMTGICKEAIQCIT
Gmax	361	KEQKLKEAKMHLVYVRGGPNYQKGLAKMRALGEEIGIPIEVYGPEATMTGICKQAIQYIT
Mtru	361	KESKLKAARMHLVYVRGGPNYQKGLEKMRALGEEIGIPIEVYGPEATMTGICKQAIQCIT
Atha	361	KEAKLKAARMHIFVRRGGPNYQKGLAKMRSLGDEIGVPIEVYGPEATMTGICKEAIQYIT
Lfil	417	AAA
Ljap	421	AAA
Gmax	421	AAA
Mtru	421	ASA
Atha	421	AAA

Appendix 5

Supplementary Figures S11 – S17 Amino acid alignments for yellow-differentially expressed carotenoid genes in *Lotus filicaulis*

Figure S11. Amino acid alignment for CDS-1

Lfil	1	M--ASLIHCPAATSLSATR-ADSVGFLVLPSSSRFIKTHKCRVRCSLDSNVSDMSVNAPK
Ljap	1	M--ASLIHCPAATSLSATR-GDSVGFLV-PSSSRFIKTHKCRVRCSLDSNVSDMSVNAPK
Gmax	1	MAMASLIQCS-ATSLSAV-----PITTRFTRTHKSRLRCSLDANVSDMSVNAPK
Mtru	1	MM-TSLIHCP-ATSLSTTRPGDSVAFFV-PSRFTKTRSYKPRVRCSLDSNVSDMSTNAPK
Atha	1	M--ASSVVFAATGSLSVP-----P-----LKSRRFYVNSSLDSDVSDMSVNAPK
Lfil	58	GLFPPEPEHYRGPKLKVAIIAGLAGMSTAVELLDDQGHEVDIYESRPFIGGKVGVSFVDKR
Ljap	57	GLFPPEPEHYRGPKLKVAIIAGLAGMSTAVELLDDQGHEVDIYESRPFIGGKVGVSFVDKR
Gmax	49	GLFPPEPEHYRGPKLKVAIIAGLAGMSTAVELLDDQGHEVDIYESRPFIGGKVGVSFVDKG
Mtru	58	GLFPPEPEHYRGPKLKVAIIAGLAGMSTAVELLDDQGHEVDIYESRTFIGGKVGVSFVDKR
Atha	43	GLFPPEPVFYKGPGLKVAIIAGLAGMSTAVELLDDQGHEVDIYDSRTFIGGKVGVSFVDKR
Lfil	118	GNHIEMGLHVFFGTCYNNLFRLMKKVGAEENLLVKDHTHTFVNTGGQIGELDFRFPVIGAPL
Ljap	117	GNHIEMGLHVFFGTCYNNLFRLMKKVGAEENLLVKDHTHTFVNTGGQIGELDFRFPVIGAPL
Gmax	109	GNHIEMGLHVFFGTCYNNLFRLMKKVGAEENLLVKDHTHTFVNKGGQIGELDFRFPVIGAPL
Mtru	118	GNHIEMGLHVFFGTCYNNLFRLMKKVGAEENLLVKDHTHTFVNKGGQIGELDFRFPVIGAPL
Atha	103	GNHIEMGLHVFFGTCYNNLFRLMKKVGAEKNNLLVKDHTHTFTNKDGTIGELDFRFPVIGAPL
Lfil	178	HGITAFLTTNQLKTYDKARNALALALSPVVRALVDPDGAIRDINLDSVSFSDWFLSKGG
Ljap	177	HGITAFLTTNQLKTYDKARNALALALSPVVRALVDPDGAIRDINLDSVSFSDWFLSKGG
Gmax	169	HGIRAFLLTNQLNTYDKARNAVALALSPVVRALVDPDGAIRDINLDSVSFSDWFLSKGG
Mtru	178	HGINAFLTTNQLKTYDKARNAVALALSPVVRALVDPDGAIRDINLDSVSFSDWFLSKGG
Atha	163	HGIRAFLLVTNQLKPYDKLRNSLALALSPVVKALVDPDGAMRDINLDSVSFSDWFLSKGG
Lfil	238	TRTSIQKMWDVPVAYALGFIDCDNISARCMLTIFALFATKTEASLLRMLKGSPDVYLSGPI
Ljap	237	TRTSIQKMWDVPVAYALGFIDCDNISARCMLTIFALFATKTEASLLRMLKGSPDVYLSGPI
Gmax	229	TRMSITKMWDVPVAYALGFIDCDNISARCMLTIFALFATKTEASLLRMLKGSPDVYLSGPI
Mtru	238	TRTSIKKMWDVPVAYALGFIDCDNISARCMLTIFALFATKTEASLLRMLKGSPDVYLSGPI
Atha	223	TRASIQRMWDVPVAYALGFIDCDNMSARCMLTIFSLFATKTEASLLRMLKGSPDVYLSGPI
Lfil	298	RKYITDRGGRFHLRWGCREILYDESADGSTYVKGLSMSKATAKQIVKADAYVAACDVPGI
Ljap	297	RKYITDRGGRFHLRWGCREILYDESADGSTYVKGLSMSKATAKQIVKADAYVAACDVPGI
Gmax	289	RKYIMDRGGRFHLRWGCREILYDKSADGSTIYVRGLSMSKATAKIVKADAYVAACDVPGI
Mtru	298	RKYITDRGGRFHLRWGCREVLYDKSADGSTYVTGLSLSKATEKKIVEADAYVAACDVPGI
Atha	283	KQYITDRGGRIHLRWGCREILYDKSADGETYVTGLATSKATNKKIVKADVVAACDVPGI
Lfil	358	KRLIPSEWRENEFFNNIYELVGVPVTVQLRYNGWVTELQNLEKSRQLKKATGLDNLLYT
Ljap	357	KRLIPSEWRENEFFNNIYELVGVPVTVQLRYNGWVTELQNLEKSRQLKKATGLDNLLYT

```

Gmax 349 KRLLPSEWREQEFFNNIYELVGVPVTVQLRYNGWVTELDLEKSRRLGKAVGLDNLLYT
Mtru 358 KRLLPSEWREKEFFNNIYELVGVPVTVQLRYNGWVTELNLELSRQLKKATGLDNLLYT
Atha 343 KRLLPKEWRESRFFNDIYELEGVPVTVQLRYNGWVTELDIELARQLKRAVGLDNLLYT

```

```

Lfil 418 PDADFSCFADLALTSPEDYYIEQGGSLLQCVLTPGDPYMPLPNEEIIARVAKQ-----
Ljap 417 PDADFSCFADLALTSPEDYYIEQGGSLLQCVLTPGDPYMPLPNEEIIARVAKQ-----
Gmax 409 PDADFSCFADLALSSPEDYYIEQGGSLLQCVLTPGDPYMPLPNDEIIARVAKQVLALFPS
Mtru 418 PDADFSCFADLALASPEDYYIEQGGSLLQCVLTPGDPYMPLPNEEIIISRVAKQVISLFPS
Atha 403 PDADFSCFADLALASPADYYIEQGGLLQCVLTPGDPYMRMPNDKIIEKVAMQVTELFPS

```

```

Lfil -----
Ljap -----
Gmax 469 SQGLEVTWSSVVKIGQSLYREGPGKDPYRPDQKTPVRNFFLAGSYTKQDYIDSMEGATLS
Mtru 478 SQGLEVTWSSVVKIGQSLYREGPGKDPFRPDQKTPVKNNFFLSGSYTKQDYIDSMEGATLS
Atha 463 SRGLEVTWSSVVKIAQSLYREAPGKDPFRPDQKTPIKNFFLAGSYTKQDYIDSMEGATLS

```

```

Lfil -----
Ljap -----
Gmax 529 GRQASAYICDAGEELVALRKKLLDAEFKDDLKISNTKDELSLVX
Mtru 538 GRQTSAYICDAGEELVALRKELLAQSKDDIKFTNTKDELSLVX
Atha 523 GRQASSYICDAGEELAELENKKLSSS-----ATAVPDELSLVX

```

Figure S12. Amino acid alignment for CIS

Lfil	1	MASSPLVFSAPFFSHRTSSTKTHLSLPLSSSRSVFPSKPLLLSRKLLAPASVAEVGEDSA
Ljap	1	MASSPLVFSAPFFSHRTSSTKTHLSLPLSSSRSVFPSKPLLLSRKLLAPASVAEVGEDSA
Gmax	1	MAASVVV-STSLSVYGY-HRRPSISITKTNLKSNSNSVAPIYCRKFVARTSIGEVGEDSA
Mtru	1	MASTPLVFSTSISSHLL-HHHP--SLPSSSLCQYFSSNPLLLSHKLVTRVSSSEVGEDSA
Atha	1	MAVYHLILLSSPPSLLPPSPRRPNLRIPAH-PRSLSSSSPV-IRKILVRSTIREIGEDSA
Lfil	61	AFDLQSQKVSSWVYFTVILGVVLFALNVVWIDPSTGFGKDFVDAVSGISDSPEVVMILLI
Ljap	61	VFDLQSQKISSWVYFTVILGVVLFALNVVWIDPSTGFGKDFVDAVSGISDSPEVVMILLI
Gmax	59	VFELGKQKVSSWIYFTAILGVVLYVLNVAWIDNSTGYGKAFIDAVSTLSDSSEVVMILLI
Mtru	58	TFDIQNQKLSSWFYFTAVLGVVLSVLNVIWIDNSTGFGKAFVDSISGISDSHEVVMILLI
Atha	59	AFELGKQKLVSWVYFGVVLGVVLFILNVVWIDNSTGFGKSFIDAVSNISGSPEVAMILLI
Lfil	121	LIFAGVHSGLASFRDTGEKLIGERAFRVLFAGTSLPLALTIVYFINHRYDGLQLWQLQN
Ljap	121	LIFAGVHSGLASFRDTGEKLIGERAFRVLFAGTSLPLALTIVYFINHRYDGFQLWQLQN
Gmax	119	LIFAGVHSGLASFRNTGEKLIGERPFRVLFAGISLPLAVSTVYFINHRYDGLQLWQLQD
Mtru	118	LIFAGVHSGLASFRDTGEKLIGERAYRVLFAGTSLPLALTMIVYFINHRYDGLQLWQLQD
Atha	119	LIFAIHVHSGLASLRDICEKLIGERAFRVLFAGISLPLAMSTIVYFINHRYDGSQLWQLQG
Lfil	181	VPGLHQLLWLSNFI SFYFLYPSTFNLLEVAAVDKPKLHLWETGIMRITRHPQIVGQVIWC
Ljap	181	VPGLHQLLWLSNFI SFYFLYPSTFNLLEVAAVDKPKLHLWETGIMRITRHPQIVGQVIWC
Gmax	179	APGLHQLLWLSNFI SFYFLYPSTFNLLEVAAVDKPKLHLWETGIIRITRHPQMVGQVIWC
Mtru	178	VPGVHQLVWISNFI SFYFLYPATFNLLEVAAVDKPKLHLYETGIMRITRHPQMVGQVMWC
Atha	179	VPGVHEAIWVANFVSFFFLYPSTFNLLEVAAVDKPKMHLWETGIMRITRHPQVLNISLMC
Lfil	241	LAHTIWIGNSVAVAASVGLIGHHLFGAWNGDRRLAKRYGEDFDLVKERTSIVPFAAILDG
Ljap	241	LAHTIWIGNSVAVAASVGLIGHHLFGAWNGDRRLAKRYGEDFDLVKERTSIVPFAAILDG
Gmax	239	LAHTIWIGNSVAVAASIGLIAHHLFGVWNGDRRLAIRYGEDFELVKSRTSVVPFAAILDG
Mtru	238	LAHTVWIGNSVAVAASFALIAHHLFGVWNGDRKLAERYGEDFEIVKGRTSVVPFAAILDG
Atha	239	LNMKT-----I-----
Lfil	301	RQRLPKDFYKEFIRLPYLTITFLTGLGAYFAHPLMQSASFNLHW
Ljap	301	RQRLPKDFYKEFIRLPYLTITFLTGLGAYFAHPLMQSASFNLHW
Gmax	299	RQELPKDFYKEFIRLPYLTITAITLGLGAYFAHPLMQAASFNLHW
Mtru	298	RQKLPKDFYKEFIRLPYFTITAITLGLGAYFAHPLMRAASLKLHW
Atha	245	-----SFIFVQYRKIGFL-----

Figure S13. Amino acid alignment for CRTISO

Lfil	1	VKGARVLVLEKYVIPGGSSGFYQRDGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCKMQ
Ljap	1	VKGARVLVLEKYVIPGGSSGFYQRDGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCKMQ
Gmax	1	VKGARVLVLEKYVIPGGSSGFYQRDGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCRMQ
Mtru	1	VKGAKVLVLEKYVIPGGSSGFYHREGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCRMQ
Atha	1	VKEARVLVLEKYVIPGGSSGFYERDGYTFDVGSSVMFGFSDKGNLNLITQALKAVGRKME
Lfil	61	VVPDPTTVHFHLPNNLSVRVHREYDKFIEELTCYFPHEKEGILKFYGECWKIFNALNSLE
Ljap	61	VVPDPTTVHFHLPNNLSVRVHREYDKFIEELTCYFPHEKEGILKFYGECWKIFNALNLE
Gmax	61	VVPDPTTVHFHLPNNLSVRVHREYDKFIEELTSYFPHEKEGILKFYGECWKIFNALNSLE
Mtru	61	VIPDPTTVHFHLPNNLSVRVHREYDKFIEELTSYFPHEKDGILKFYGECWKIFNALNSLE
Atha	61	VIPDPTTVHFHLPNNLSVRTHREYDDFIAELTSKFPHEKEGILGFYGDCKWKIFNSLNSLE
Lfil	121	LKSLEEPLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Ljap	121	LKSLEEPLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Gmax	121	LKSLEEPLYLFGQFFRRPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Mtru	121	LKSLEEPLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Atha	121	LKSLEEPYLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIKDPQLLSFIDAECFIVSTVN
Lfil	181	ALQTPMINAAMVLCDRHFGGINYPGGVGGIKSLAKGLADQGSEILYKANVTSVIIIEQG
Ljap	181	ALQTPMINAAMVLCDRHFGGINYPGGVGGIKSLAKGLADQGSEILYKANVTSVIIIEQG
Gmax	181	ALQTPMINAAMVLCDRHFGGINYPGGVGGIKSLAKGLVDQGSEIVYKANVTSVIIIEQG
Mtru	181	ALQTPMINASMVLCDRHFGGINYPGGVGGIKSLAKGLVDQGSQILYKANVTSVIITEQG
Atha	181	ALQTPMINASMVLCDRHYGGINYPGGVGGIKSLAEGLVDQGSEIQYKANVKSIILDHG
Lfil	241	KAVGVRLSDGREFFAKTIISNATRWDT-----
Ljap	241	KAVGVRLSDGREFFAKTIISNATRWDTFG-----
Gmax	241	KAVGVRLSNGREFFAKTIISNATRWDTFGKLLKGVPLPKEEENFQKVYVKAPSFLSIHMG
Mtru	241	KAVGVRLSDGREFFAKTIISNATRWDTFGKLVKGESLPKEEENFQKVYVKAPSFLSIHMG
Atha	241	KAVGVRLADGREFFAKTIISNATRWDTFGKLLKGEKLPKEEENFQKVYVKAPSFLSIHMG
Lfil		-----
Ljap	270	-----
Gmax	301	VKAEVLPPDTDCHHFVLESNWWSKLEEPYGSIFLSIPTVLDSSLAPEGRHILHIFTTSSME
Mtru	301	VKAEVLPPDTDCHHFVLENNWWSLEEPYGSIFLSIPTVLDSSLAPEGRHILHIFTTSSMD
Atha	301	VKAEVLPPDTDCHHFVLEDDWKNLEEPYGSIFLSIPTVLDSSLAPDGRHILHIFTTSSIE
Lfil		-----
Ljap	270	-----
Gmax	361	DWEGLSRVEYEAKKQLVADEITSRLNKLFPGLRSSIDFMEVGTPKTHRRYLARDEGTYG
Mtru	361	DWQGLSRIDYEAKKQVVADEIISRLEKKLFPGLRSSIDFIEVGTPKTHRRYLARDEGTYG
Atha	361	DWEGLPPKEYEAKKEDVAARIIQRLKKLFPGLSSSITFKEVGTPRTHRRFLARDKGTYG
Lfil		-----
Ljap	270	-----
Gmax	421	PMPRRIPKGLLGMPFNNTTGIDGLYCVGDSCFPQGQVIAVAFSGVMCAHRVAADIGLEKKS
Mtru	421	PMPRSIPKGLLGMPFNNTTSIDGLYCVGDSCFPQGQVIAVAFSGVMCAHRVAADIGLEKKS

Atha 421 PMPRGTPKGLLGMPFNNTAIDGLYCVGDSCFPGQGVI AVAFSGVMCAHRVAADIGLEKKS

Lfil -----

Ljap 270 -----WER

Gmax 481 PVLD SMLLRLLGWLR

Mtru 481 PVLDGMLLGLLGWLR

Atha 481 RVLDVGLLGLLGWLR

Figure S14. Amino acid alignment for E-CHY

Lfil	1	-----TLWVLVTLPPLSMC-----SR
Ljap	1	-LSLSPLSLFPPSVKSSINKKPSSWVSPDWLTSLSRSLTAGKNDDSGIPIASAKLDDVSD
Gmax	1	SLSLTPLSVFPSPSVKSSINKKPSSWVSPDWLTSLSRSLTAG-NDVSGIPVASAKLDDVSD
Mtru	1	SLPLSHLSLSSFLTSSSTNKKSTSWVSPDWLTSLSKSLTTSKNDDSNIPIASAKLDDVSD
Atha	1	SSSYSSLFTAKPSIRSSIEKKSQSWVSPDWLTTLTRTLSSGKNDESGIPIANAKLDDVAD
Lfil	17	IMG-----NMARVLLRFRSFCGLGLVLQLLRV
Ljap	60	LLGGALFLPLFKWMKEYGPIYRLAAGPRNFVVSDPAIAKHVLRNYGKYGKGLVAEVSEF
Gmax	60	LLGGALFLPLFKWMQDYGPIYRLAAGPRNFVVSDPAIAKHVLRNYGKYAKGLVAEVSEF
Mtru	61	LLGGALFLPLFKWMNEYGPIYRLAAGPRNFVVSDPAIAKHVLRNYGKYGKGLVAEVSEF
Atha	61	LLGGALFLPLYKWMNEYGPIYRLAAGPRNFVIVSDPAIAKHVLRNYPKYAKGLVAEVSEF
Lfil	44	HSGGTCLL---XWICY-----FVDVLRDXWRSYNLMH-----LM
Ljap	120	LFGGFALIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRCAERLVEKLOPDALNGTAVNM
Gmax	120	LFGGFALIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRCAERLVEKLOPDALNGTAVNM
Mtru	121	LFGGFALIAEGPLWTARRRAVPSLHKRYLSIMVDRVFCRCAERLVEKLOQADAVNGTAVNM
Atha	121	LFGGFALIAEGPLWTARRRAVPSLHRRYLSVIVERVFCRCAERLVEKLOPYAEDGSAVNM
Lfil	75	ELQLTFHXSLLIMLLVYLYSTMTLTQTVLISVLFLLXKRQARSTDLLPYWQVEFLRKII
Ljap	180	EDKFSQLTLDVIGLSVFNYDFDSLNDSPVIGAVYTALKEAEARSTDLLPYWQVEFLRKII
Gmax	180	EAKFSQLTLDVIGLSVFNYNFDLNDSPVIEAVYTALKEAEARSTDLLPYWKFKFLCKII
Mtru	181	EDKFSQLTLDVIGLSVFNYNFDLNDSPVIEAVYTALKEAEARSTDLLPYWKIDFLCKII
Atha	181	EAKFSQMTLDVIGLSLFNYNFDSLTDSPVIEAVYTALKEAELRSTDLLPYWKIDALCKIV
Lfil	135	PRQIKAEKAVSIIRKTVENLIECKKEIVESEGERIDGDEYVNDSDPSILRFLLASREEVS
Ljap	240	PRQIKAENAVTIIRKTVEDLIECKKEIVESEGERIDGDEYVNDSDPSILRFLLASREEVS
Gmax	240	PRQIKAEKAVSVIRKTVEDLIECKREIVESEGERIDVEEYVNDSDPSILRFLLASREEVS
Mtru	241	PRQIKAENAVTVIRKTVEDLIEQCKEIVESEGERIDADEYVNDADPSILRFLLASREEVS
Atha	241	PRQVKAEKAVTLIRETVEDLIAKCKEIVEREGEINDEEYVNDADPSILRFLLASREEVS
Lfil	195	SDQLRDDLSSMLVAGHETTGSVLTWTLYLLSKDSSSLAKVQEEVDRVLOGRRPTFEDMKS
Ljap	300	SDQLRDDLSSMLVAGHETTGSVLTWTLYLLSKDSSSLAKVQEEVDRVLOGRRPTFEDMKN
Gmax	300	SVQLRDDLSSILVAGHETTGSVLTWTLYLLSKDSSSLAKAQEEVDRVLOGRRPTYEDIKD
Mtru	301	SVQLRDDLSSMLVAGHETTGSVLTWTLYLLSKDSSSLAKAQEEVDRVLOGRRPTYEDMKD
Atha	301	SVQLRDDLSSMLVAGHETTGSVLTWTLYLLSKNSSALRKAQEEVDRVLEGRNPAFEDIKE
Lfil	255	LKFLTRCITESLRLYPHPPVLIRRAQVPDELPGAYKVNAGQDIMISVYNIHHSSEVWDRA
Ljap	360	LKFLTRCITESLRLYPHPPVLIRRAQVPDELPGAYKVNAGQDIMISVYNIHHSSEVWDRA
Gmax	360	LKFLTRCIIIESLRLYPHPPVLIRRAQVPDELPGGYKLDAGQDIMISVYNIHRSSEVWDRA
Mtru	361	LKFLNRCIIIESLRLYPHPPVLIRRSQIPDELPGDYKIDAGQDIMISVYNIHSSKVWDRA
Atha	361	LKYITRCINESMRLYPHPPVLIRRAQVPDILPGNYKVNTGQDIMISVYNIHRSSEVWEKA
Lfil	315	EEFLPERFDLDGPMNETNTDFRFIPFSGGPRKCVGDQFALLEATVSLAIFLQHMNFELV
Ljap	420	EEFLPERFDLDGPMNETNTDFRFIPFSGGPRKCVGDQFALLEATVSLAIFLQHMNFELV
Gmax	420	EEFVPERFDLDGPPNETNTDFRFIPFSGGPRKCVGDQFALMEATVALAIFLQHMNFELV
Mtru	421	EEFLPERFDLDGPPNETNTDFRFIPFSGGPRKCVGDQFALLEATVALAVFLQHMNFELV

Atha	421	EEFLPERFDIDGAI	PNETNTDFK	FIPFSGGPRKCVGDQFALMEAI	IVALAVFLQRLNVELV
Lfil	375	PDQNISM	TTGATIHTTNGLYMKLSQRVK		
Ljap	480	PDQNISM	TTGATIHTTNGLYMKLSQRVK		
Gmax	480	PDQNISM	TTGATIHTTNGLYMKLSRRLK		
Mtru	481	PDQNI	GMTTGATIHTTNGLYMKLSQRLK		
Atha	481	PDQT	ISM	TTGATIHTTNGLYMKVSQR--	

Figure S15. Amino acid alignment for PSY

Lfil	1	VTLEFWASYPNLEVSNSTGSLHSTRDVRLFDSSKFMLPDFRSIRAKKDRKKRWRFCSLST
Ljap	1	VTLEFWASFPNLQVSNSTRLLDSIRDVRLFDSSKFMSRDFRSIRAKKDRKKRWRFCSLST
Gmax	1	VTFLWFVSSPGLEVSHSTGLLDSDVRHVKFLDSSKVMRSDFGSIRVEKDKEKGWRLCSLST
Mtru	1	VTLLRVASSPSLEVSCSNGLLDSFGVVKLVDSKFLSRDFVSIRAKKDKKKTRRFCSLST
Atha	1	VAVLWVAITSS-----LNPDPNNCGLVVRVLESSRLFSF-CQNQRLNKGKKKQI--PTWSS
Lfil	61	GVKYACIGQSGLESASSFPMISSVLANPAAAEVAVSSEQKVYDVVLKQASLIKRKLSAGD
Ljap	61	GVKYACVQGQSGLESASSFPMISNVLANPAAAEVAVSSEQKVYDVVLKQASLIKRKLSAGD
Gmax	61	DMKYACVGRSGLESASNFPLIANVLANPAAGEVAVSSEQKVYDVVLKQASLVKRKLGAG
Mtru	61	DIKYACVGGPGLSASNFPLISNVLANPTAGEVTVSSEQKVYDVVLKQASLVKRKLSSGE
Atha	53	SFVRNRSRR-----IGVVSSSLVASPSGEIALSSSEKVVYNVVLKQAALVNKQLRSYD
Lfil	121	LGVKQNVFLPGNLSLLSEAYDRCEICAEBYAKTFYLGTLMTERRRAIWAIVWCRRTD
Ljap	121	HGVKQNVFLPGNLSLLSEAYDRCEICAEBYAKTFYLGTLMTERRRAIWAIVWCRRTD
Gmax	121	LDAPKDIALPGNLSLLNEAYDRCEICAEBYAKTFYLGTLMTERRRAIWAIVWCRRTD
Mtru	121	LEVKPEIVLPGNLSLLSEAYDRCEICAEBYAKTFYLGTLMTERRRAIWAIVWCRRTD
Atha	105	LDVKQDVVLPGLSLLGEAYDRCEVCAEBYAKTFYLGTLMTERRRAIWAIVWCRRTD
Lfil	181	ELVDGPNASHITPTALDRWESRLEELFQGRPFDMLEAAVSDTVTKFPVDIQPFKDMIEGM
Ljap	181	ELVDGPNASHITPTALDRWESRLEELFQGRPFDMLEAAVSDTVTKFPVDIQPFKDMIEGM
Gmax	181	ELVDGPNASQITPTALDRWESRLEELFQGRPFDMLEAALADTVAKFPVDIQPFKDMIEGM
Mtru	181	ELVDGPNASHITATAMDREWSRLDELFGQGRPFDMLEAALSDTVNRFPVDIQPFKDMIEGM
Atha	165	ELVDGPNASHITPMALDRWEARLEDLFRGRPFDMLEAALADTVARYFPVDIQPFKDMIEGM
Lfil	241	RTDLRKSRYKNFDELYLYCYVAGTVGLMSVPIMGISPDSQTSSTESVYNAALALGIANQL
Ljap	241	RTDLRKSRYKNFDELYLYCYVAGTVGLMSVPIMGISPDSQTSSTESVYNAALALGIANQL
Gmax	241	RLDLKKPRYKNFDELYLYCYVAGTVGLMSVPIMGISPNSSEATTESVYNAALALGIANQL
Mtru	241	RMDLKKSRYSFDELYLYCYVAGTVGLMSVPMVGISTHSQATTESVYNAALALGIANQL
Atha	225	RMDLKKSRYSQNFDDLYLYCYVAGTVGLMSVPMVGIDPKSKATTESVYNAALALGIANQL
Lfil	301	TNILRDVGEDASRGVYLPQDELAQAGLSEDDIFAGKVTDKWRNFMKSQIKRARMFFDEA
Ljap	301	TNILRDVGEDASRGVYLPQDELAQAGLSEDDIFAGKVTDKWRNFMKSQIKRARMFFDEA
Gmax	301	TNILRDVGEDASRGVYLPQDELAQAGLSEDDIFAGKVTDKWRNFMKSQIKRARMFFDEA
Mtru	301	TNILRDVGEDARRGRVYLPQDELTLAGLSDDDDIFAGKVTDKWRNFMKSQIKRARTFFDEA
Atha	285	TNILRDVGEDARRGRVYLPQDELAQAGLSEDDIFAGKVTDKWRNFMKMQLKRARMFFDEA
Lfil	361	EKGVMELNKASRWPVLA SLLLYRQILDEIEANDYNNFTKRAYVSKTKKLLSLPAAYARTM
Ljap	361	EKGVMELNKASRWPVLA SLLLYRQILDEIEANDYNNFTKRAYVSKTKKLLSLPAAYARTM
Gmax	361	EKGVTELNEASRWPVWASLLLYRQILDEIEANDYNNFTRAYVSKAKKFLSLPAAYARS
Mtru	361	EKGVTELNEQSRWPVWASLLLYRQILDEIEANDYNNFTKRAYVSKTKKLLSLPLAYARSM
Atha	345	EKGVTELSAASRWPVWASLLLYRRILDEIEANDYNNFTKRAYVGKVKKIAALPLAYAKSV
Lfil	421	VAPPRKLSHV
Ljap	421	VAPPRKLSHV
Gmax	421	VPPSRKLSFV
Mtru	421	VPPSKKLSHV
Atha	405	LKTSSSRLSI

Figure S16. Amino acid alignment for ZEP-1

Lfil	1	-VXLCLVILSMPCSQIWAKED-----
Ljap	1	GVXLCLVILSMPCSQIWAKED-----
Gmax	1	GLPVTRVISRMVLQEILARAVGEDIIIMNASNVVNFVDDGNKVTVELENGQKYEGDVLVGA
Mtru	1	GLPVTRVISRMALQEILARAVGDDVIMNGSNVVDFFIDHETKVTVVLDNGQKYDGDLLVGA
Atha	1	GLPVTRVISRMTLQQILARAVGEDVIRNESNVVDFEDSGDKVTVVLENGQRYEGDLLVGA
Lfil	21	-----
Ljap	22	-----
Gmax	61	DGIWSKVRKQLFGLTEAVYSGYTCTGTGIADFVPADIEITVGYRVFLGHKQYFVSSDVGAGK
Mtru	61	DGIWSKVRTKLFGSTEATYSGYTCTGTGIADFVPPDIESVGYRVFLGHKQYFVSSDVGAGK
Atha	61	DGIWSKVRNNLFGRSEATYSGYTCTGTGIADFIPADIESVGYRVFLGHKQYFVSSDVGGGK
Lfil	21	-----AW
Ljap	22	-----AW
Gmax	121	MQWYAFHKEPPGGVDEPNGKKERLLRIFEGWCDNAVDLILATEEEAILRRDIYDRIPTLT
Mtru	121	MQWYAFHQEPAGGVDTPNGKKERLLKIFEGWCDNAIDLIVATEEEAILRRDIYDRTPTLT
Atha	121	MQWYAFHEEPAGGADAPNGMKKRLFEIFDGWCDNVLDLLHATEEEAILRRDIYDRSPGFT
Lfil	23	-----PLDSYQLASELDNAWEQSIKSGSPIEIDSSLRS
Ljap	24	-----PLDSYQLASELDNAWEQSIKSGSPIEIDSSLRS
Gmax	181	WGKGRVTLLGDSVHAMQPNMGQGGCMAIEDSYQLAWELNNAWEQSIKSGSPIDIDSSLRS
Mtru	181	WGKGRVTLLGDSVHAMQPNMGQGGCMAIEDGYQLAFELDNAWQSAKSGSTIDIASSLKRS
Atha	181	WGKGRVTLLGDSIHAMQPNMGQGGCMAIEDSFQLALELDEAWKQSVETTPVDVVSLLKR
Lfil	56	YERERRLRVAIIHGMAARMAALMASTYKAYLGVGLGPLEFLTFRIPHPGRVG
Ljap	57	YERERRLRVAIIHGMAARMAALMASTYKAYLGVGLGPLEFLTFRIPHPGRVG
Gmax	241	YERERRLRVAIIHGMAARMAALMASTYKAYLGVGLGPLEFLTFRIPHPGRVG
Mtru	241	YERERRLRVTFVHGMAARMAALMASTYKAYLGVGLGPFEEFLTFRIPHPGRVG
Atha	241	YEESSRLLRVAIIHAMARMAALMASTYKAYLGVGLGPLSFLTFRVPHHPGRVG

Figure S17. Amino acid alignment for ZEP-3

Lfil	1	KKNLRILVAGGGIGGLVFALAAKRKGFVVVF EKDM SAIRGEGKYRGPIQIQSNALAALE
Ljap	1	-----
Gmax	1	KKKLRVLVAGGGIGGLVFALAAKRKGFVVVF EKDM SAIRGEGQYRGPIQIQSNALAALE
Mtru	1	KKQLKVLVAGGGIGGLVFALAAKRKGFVVVF EKDL SAIRGEGQYRGPIQIQSNALAALE
Atha	1	KKKSRLVAGGGIGGLVFALAAKKGF DVLVF EKDL SAIRGEGKYRGPIQIQSNALAALE
Lfil	61	AIDPDVADEV MR LGCITGDRINGLVDGVS GSWYVKFDTFTPAVERGLPVTRVISRMVLQG
Ljap	1	-----
Gmax	61	AIDLEVAEEVLRVGCITGDRINGLVDGVS GSWYIKFDTFTPAERGLPVTRVISRMALQE
Mtru	61	AIDMNVAEV MR VGCITGDRINGLVDGVS GSWYIKFDTFTPAERGLPVTRVISRMALQE
Atha	61	AIDIEVAEQVMEAGCITGDRINGLVDGVS GTWYVKFDTFTPAASRGLPVTRVISRMTLQQ
Lfil	121	ILARAVGDDIIMNASNVVSVFDDGNKVTVELENGEKYEGDLLVGADGIWSKVRKQLFGLT
Ljap	1	-----
Gmax	121	ILAHAVGEDVIMNDSNVVDFVDHGD KVTVELENGQKYDGDLLVGADGIWSKVRKKLFGQT
Mtru	121	ILARAVGDDVIMNGSNVVD FIDHETKVTVVL DNGQKYDGDLLVGADGIWSKVRTKLFGST
Atha	121	ILARAVGEDVIRNESNVVDFEDSGDKVTV VLENGQRYEGDLLVGADGIWSKVRNNLFGRS
Lfil	181	EAVYSGYTCYTGIADFVPADIESVGYRVFLGHKQYFVSSDVGAGKMQWYAFHQEPPGGVD
Ljap	1	-----
Gmax	181	EATYSGYTCYTGIADFVPADIESVGYRVFLGHKQYFVSSDVGAGKMQWYGFHQEPAGGAD
Mtru	181	EATYSGYTCYTGIADFVPADIESVGYRVFLGHKQYFVSSDVGAGKMQWYAFHQEPAGGVD
Atha	181	EATYSGYTCYTGIADFIPADIESVGYRVFLGHKQYFVSSDVGCGKMQWYAFHEEPAGGAD
Lfil	241	IPNGRKERV LKIFEGWCDNVIELIVATEEEAILRRDIYDRTPTLTWGKGCVTLLGDSVHA
Ljap	1	-----
Gmax	241	IPNGKKERLLKIFKGWCDNVIDLIHATEEEAILRRDIYDRTPFTTWGKGHVTLLGDSIHA
Mtru	241	TPNGKKERLLKIFEGWCDNAIDLIVATEEEAILRRDIYDRTPTLTWGKGRVTLLGDSVHA
Atha	241	APNGMKKRLFEIFDGWCDNVLDLLHATEEEAILRRDIYDRSPGFTWGKGRVTLLGDSIHA
Lfil	301	MQPNIGQGGCMAIEDGYQLALELDNAWQESIKSGSPIDIDSSLKSYERERRLRVAIIHAM
Ljap	1	-----
Gmax	301	MQPNMGQGGCMAIEDSYQLALELDNAWQOSIKSGSPIDIDSSLKSYERERRLRVAIVHGM
Mtru	301	MQPNMGQGGCMAIEDGYQLAFELDNAWQOSAKSGSTIDIASLKSYERERRLRVTFVHGM
Atha	301	MQPNMGQGGCMAIEDSFQLALELDEAWKQSVETTTTPVDVVSLLKRYEESRRLRVAIHAM
Lfil	361	ARMAAIMASTYKPYLGVGLGPLEFLT KFRVPHPGRVGGRFFIDKMMPLMLNWWVLGGNSSK
Ljap	1	-----
Gmax	361	ARMAAMASTYKAYLGVGLGPLEFLT KFRIPHPGRVGGRRFFIDKMMPLMLNWWVLGGNSSK
Mtru	361	ARMAAIMASTYKAYLGVGLGPFEFLT KFRIPHPGRVGGRRFFIQKSMPLMLNWWVLGGNSSK
Atha	361	ARMAAIMASTYKAYLGVGLGPLSFLT KFRVPHPGRVGGRFFVDIAMP SMLDWWVLGGNSEK
Lfil	421	LEGRPLCCRLSDKANDQLHRWFEDDDALERTITGEWILLPCGDEAGLSKPICLRHDETKP
Ljap	1	-----
Gmax	421	LEGRPVCCRLSDKANDQLHRWFEDNDALERAINGEWILLPCGDEAGPTKPICLTQDEMKP
Mtru	421	LEGRPLCCRLSDKASDQLHTWFEDDDALERTINGEWILLPCGDVPGHVKPISLNQDDTKP
Atha	421	LQGRPPSCRLTDKADDRLREWFEDDDALERTIKGEWYLIPHGDDCCVSETICLTKDEDQP

```

Lfil  481  CIIGQEDYPGSSITITLPQVSQMHAQINYKDGAFFLTDLQSQHGTWITDNEGRRNRVPPN
Ljap   1  ----QEDYPGSSITITLPQVSQMHAQINYKDGAFFLTDLQSQHGTWITDNEGRRNRVPPN
Gmax  481  CIIGQKDHHPGSSIIPLPQVSQMHAQINYKDGAFFLTDLRSQHGTWITDNEGRRYRVPPN
Mtru  481  YIIGQEDYPGSLITIPLPQVSQIHARINFKDGAFFLTDLRSQHGTWITDNEGRRYMVSPN
Atha  481  CIVGDQDFPGMRIVIPSSQVYKLYACSCDLQRRSFLLDGSKRTRNLC-----

```

```

Lfil  541  CPARVRPSDLIEFG
Ljap   57  CPARVRPSDLIEFG
Gmax  541  YPARVRPSDVVEFG
Mtru  541  YPARIRPSHVIIEFG
Atha  -----

```

Appendix 6

Supplementary Figures S18 – S19 Amino acid alignments for non-differentially expressed flavonoid genes in *Lotus filicaulis*

Figure S18. Amino acid alignment for F3H-2

Lfil	1	MEVSGKPIRVQSLAHSGLSRVPPE-----YIQPPQ
Ljap	1	MEVSGEPIRVQSLAHSGLSRVPPE-----YIQPPQ
Gmax	1	MT-----TRVQSLAQSGLSRVPPQ-----YIQPPQ
Mtru	1	MELTGEPIRVQSI IQSGLSQVPPE-----YIQPPQ
Atha	1	MY-----LMSSFSSSLPPYPFHISKISTIKTSSLGELESKTTSDSHKSGFLSTSS
Lfil	31	NR---PVHHNQDPSTNIPTINLSGFDPNHR---ESIRACREWGAFHVTNHGVPTTLL
Ljap	31	NR---PVHHNQDPSTNIPTINLSGFDPNHRDSTRESIRACREWGAFHVTNHGVPTTLL
Gmax	26	TR---PVRHTAPEP-DSIPVIDLSSFDPTQRASTRDSIARACREWGAFHVTNHGVPPSLL
Mtru	31	NR---PIN---TDPTATIPADLFNFENTNHRNSTLESISHACRGWGAFHVTNHGIPPSLL
Atha	51	HRKLDLISH-YPGDAASIPTVDLSSSD-----SAREAIGDACRDWGAFHVINHGVP IHL
Lfil	84	DSLRRMGLAFFNECPMPEKRLRYACAAGSAASEGYGSRMLVSD--ENNDGAQVLDWRDYFD
Ljap	88	DSLRRSGLAFFNECPMPEKRLHYACAAGSAASEGYGSRMLVSD--ENNDGAQVLDWRDYFD
Gmax	82	ASLRRAGLSFFSDTPIPKLRYSCSA--AASEGYGSKMLATTTSQNDQAVQVLDWRDYFD
Mtru	85	DAVRNSGLTFFNNCPMSEKLYSCTAGTAASEGYGSRMLVSS--ND---HEVLDWRDYFD
Atha	105	DRMRSLGLSFFQDSPMEKRLRYACDSTSAASEGYGSRMLIGA--KD---DVVLDWRDYFD
Lfil	142	HHTLPLSRRNPNRWPEFSSGYRELVARYSDEMKA LAQKLLSLISESLGLRPSCIEDV VGE
Ljap	146	HHTLPLSRRNPNRWPEFSSGYRELVARYSDEMKA LAQKLLSLISESLGLRPSCIEDV VGE
Gmax	140	HHTLPLSRRNPNRWPEFPADYREL VATYSDEMKILAQKLLALISESLGLRASCIEDAVGE
Mtru	140	HHTFPLSRRNPINWPDFTSDYREIMVNYSDMKILAQKLLALISESLGLQSSCIEDAVGD
Atha	160	HHTFPPLSRRNP SHWPIHPSDYRQVVG EYGD EMMKKLAQM LGLISESLGLPCSSIEFVAVGE
Lfil	202	FYQNITISYPPCPQPD LTLGLQSHSDFGAITLLIQDDVGGLQVLKNGDGGGGGDSWVTV
Ljap	206	FYQNITISYPPCPQPD LTLGLQSHSDFGAITLLIQDDVGGLQVLKNGD--GGDSWITV
Gmax	200	FYQNITISYPPCP E PDLTLGLQSHSDMGAITLLIQDDVGGLQVLKG-----GNKQWTV
Mtru	200	IYQNITVSYPPCPQPD LTLGLQSHSDFGAVTLLIQDDVGGLQVLKGDGD---GGDKWTV
Atha	220	IYQNITVTYPPCPQPELT LTLGLQSHSDFGAITLLIQDDVEGLQLYKDA-----QWLTV
Lfil	262	QPLSDAILVLLADQTEIITNGKYRSCVHRAITNPDRARLSVATFHDPAKTVRISPASELI
Ljap	264	QPLSDAILVLLADQTEI VNKKT--L-----
Gmax	254	QPLSDAILVLLADQTEIITNGKYRSCVHRAITNPDRARLSVATFHDPAKTVKISPASELI
Mtru	257	KPLSDAILVLLGDQTEIITNGKYRSCVHRAVTNPDRARLSVATFHDPAKTVKIFPVSELI
Atha	273	PPISDAILIL IADQTEIITNGRYKSAQHRAVTNANRARLSVATFHDP SKTARIAPVSQL-
Lfil	322	DESSPAKFRDVVYG DYVTSWYTKGPEGKRNIDALVLES*
Ljap	286	-----P*
Gmax	314	NDSSLAKYRDVVYG DYVSSWYTKGPGGKRNIDALLLP*
Mtru	317	NDSSPAKYRGVVYG DYVSSWYTKGPDGKRNIDALLLES*
Atha	332	---SPPSYKEVVYGYVSSWYSKGPEGKRNLDALLY--*

Figure S19. Amino acid alignment for F3H

Lfil	1	MGEVDPAFIQDPEHRPKLSI-----IKAEGIPVIDLSPIISQTVSDPSAIEGLVKEIGE
Ljap	1	MGEVDPAFIQEPEHRPNLSI-----IQAEGIPVIDLSPIINKTISDPSAIEGLVKKVGR
Gmax	1	MGEVDPAFIQEPEHRPKLSP-----NQAEGIPVIDLSPIITNHAVSDPSAIENLVKEIES
Mtru	1	MGEVDPAFVQEPEHRPKLSI-----IEAKGIPEIDLSPILHHAVPNPSDIENLVKEIGS
Atha	1	MGELDEAFIQAPEHRPNTHLTNSGDFIFSDEIPTIDLSSLE-DTHHDKTA---IAKEIAE
Lfil	55	ACKEWGFFQVTNHGVPLSLRQSIDEASRKFFAQSIIEEKRKIYRDESTISGYDYDEHTKNV
Ljap	55	ACKEWGFFQVTNHGVPLSLRQSIDEASRKFFAQSMEEKRKIYRDESTISGYDYDEHTKNV
Gmax	55	ACKEWGFFQVTNHGVPLTLRQNIIEKASRLFFEQEQEKKKVSDEKSTTGYDYDEHTKNV
Mtru	55	ASKEWGFFQVTNHGVPLSLRQRLIEASRLFFAQSLIEKKKVARDEVNPTGYDYDEHTKNV
Atha	57	ACKRWGFFQVINHGLPSALRHRVEKTAEEFNLTTEEKRKVKRDEVNPMGYHDEEHTKNV
Lfil	115	RDWKEVDFDFVAREPTVVPLTSDDEDDRLTQWTNKSPEYPPNFRVILQEYIQELEKLAFK
Ljap	115	RDWKEVDFDFVATEPTVVPLTSD-EDDDRITQWTNKSPEYPPNFSVILQEYMQEMEKLAFK
Gmax	115	RDWKEVDFDLAKDPTFIPVTSD-EHDDRVTHTWNTVSPYPPNFRDIMEEYIQEVEKLSFK
Mtru	115	RDWKEVDFDLSQDPTLVPRSSD-EHDDGVIQWTNQSPQYPPQFRAILEEYIQEVEKLAYR
Atha	117	RDWKEIFDFFLQDSTIVPASPE-PEDTELRLKLTNQWPQNPSHFREVCQEYAREVEKLAFR
Lfil	175	LLELIALSIGLEAKRFEQFFTKDQTSFIRLNHYPPCPFPHLALGVGRHKDAGGLTILAQD
Ljap	174	LLELIALSIGLEAKRFEQFFTKDQTSFIRLNHYPPCPFPHLALGVGRHKDAGGLTILAQD
Gmax	174	LLELIALSIGLEAKRFEFFFMKDQTSFIRLNHYPPCPYPHLALGVGRHKDGGALTVLAQD
Mtru	174	LLELIALSIGLEAKRFEFF-KYQTSFIRFNHYPPCPYPHLALGVGRHKDAGALTILAQD
Atha	176	LLELVSIISLGLPCDRLTGFF-NEQTSFLRFNHYPPCPNPELALGVGRHKDGGALTVLAQD
Lfil	235	QVGGLEVKRKADQQWVRVEPTDAFIINVCDSIQVWSNDAYESVEHRVMVNSEKERFSIP
Ljap	234	EVGGLEVVKRKADQQWVRVEPTSDAFIINVCDSIQF-----
Gmax	234	EVGGLEVVKRKADQEWIRVKPTPDAYIINVGDLIQVWSNDAYESVEHRVMVNSEKERFSIP
Mtru	233	EVGGLEVVKRKSQQQWVLVKPTPDAYIINVGDIQVWSNDAYESVEHRVMVNSEKERFSIP
Atha	235	SVGGLOVSRSDGQWIPVKPISDALIINMGNCIQVWTNDEYWSAEHRVVVNTSKERFSIP
Lfil	295	YFFNPAHDCEVKPLDELTNEQNPSKYRPIYEWGKFFVNRMGNLKKKNVENIQIYHY----
Ljap	269	-----FVCRFGAM-----
Gmax	294	FFFNPAHDIIEVKPLEELTNEHNPSKYRPIYKWKFLVHRKGSNFKKQNVENIQIYHYKIT*
Mtru	293	FFFFPAHDTVVKPLEELTNEENPPKYRPIYNWGKFLVNRKSNFEKKKVENIQIYHYKIA*
Atha	295	FFFFPSHEANIEPLEELISEENPPCYKKYNWGKFFVSRNRSDFKKLEVENIQIDHFK-A*

Appendix 7

Supplementary Figures S20 – S28 Amino acid alignments for red-differentially expressed flavonoid genes in *Lotus filicaulis*

Figure S20. Amino acid alignment for ANS

Lfil	1	MAPTVERVESLSGSGIQSIPKEYVRPKEELANIGDVFEEEKK-VGPQVPTIDLKEIDSP
Ljap	1	MAPTVERVESLSGSGIQSIPKEYVRPKEELANIGDVFEEEKK-VGPQVPTIDLKDIDSP
Gmax	1	MG-TVAPRVESLASSGIKCIPKEYVRPQEELKSIGNVFEEEKK-EGLQVPTIDLREIDSE
Mtru	1	MG-TVAQRVESLALSGISSIPKEYVRPKEELANIGNIFDEEEKK-EGPQVPTIDLKEINSS
Atha	1	MV--AVERVESLAKSGIISIPKEYIRPKEELESINDVFLEEEKEDGPQVPTIDLKNIESD
Lfil	60	DEFVRAKCREKLKAAEEWGMHLVNHGIPDELINQLKSAGAEFFSLPVEEKEYANDQA
Ljap	60	DEFVRAKCREKLKAAEEWGMHLVNHGIPDELINQLKSAGAEFFSLPVEEKEYANDQA
Gmax	59	DEVVRGKCREKLKAAEEWGMHLVNHGIPDELIERVKKAGETFFGLAVEEKEYANDLE
Mtru	59	DEIVRGKCREKLKAAEEWGMHLVNHGISDDLINRLKAGETFFELPVEEKEYANDQS
Atha	59	DEKIRENCIEELKASLDWGMHLINHGIPADLMERVKKAGEEFFSLPVEEKEYANDQA
Lfil	120	AGNVQGYGSKLANNASGQLEWEDYFFHLIFPEDKRDL SIWPKTPSYYTEVTS DYARRLRV
Ljap	120	AGNVQGYGSKLANNASGQLEWEDYFFHLIFPEDKRDL SIWPKTPSYYTEVTS DYARRLRV
Gmax	119	SGKIQGYGSKLANNASGQLEWEDYFFHLAFPEDKRDL SFWPKKPADYIEVTSEYAKRLRG
Mtru	119	SGKIQGYGSKLANNASGQLEWEDYFFHCIFPEDKRDL SIWPKTPADYTKVTSEYAKELRV
Atha	119	TGKIQGYGSKLANNASGQLEWEDYFFHLAYPEEKRDLSIWPKTPSDYIEATSEYAKCLRL
Lfil	180	LASKILEVLSLELGLEEGRLEKEVGGMEELL LQMKINYYPKCPQPELALGVEAHTDISAL
Ljap	180	LASKILEVLSLELGLEEGRLEKEVGGMEELL LQMKINYYPKCPQPELALGVEAHTDISAL
Gmax	179	LATKILEALSIGLGLGRRLEKEVGGMEELL LQMKINYYPICPQPELALGVEAHTDVSSSL
Mtru	179	LASKIMEVLSLELGLEGGRLEKEAGGMEELL LQMKINYYPICPQPELALGVEAHTDVSSSL
Atha	179	LATKVFKALS VGLGLEPDRLEKEVGGLEELL LQMKINYYPKCPQPELALGVEAHTDV SAL
Lfil	240	TFL LHNMV PGLQLFYEGKWVTAKCVPDSILMHIGDTTEILSNGKFKSILHRGLVNKEKVR
Ljap	240	TFL LHNMV PGLQLFYEGKWVTAKCVPDSILMHIGDTTEILSNGKFKSILHRGLVNKEKVR
Gmax	239	TFL LHNMV PGLQLFYEGQWVTAKCVPDSILMHIGDTTEILSNGKYKSILHRGLVNKEKVR
Mtru	239	TFL LHNMV PGLQLFYEGKWVTAKCVPDSILMHIGDTTEILSNGKYKSILHRGLVNKEKVR
Atha	239	TFIL LHNMV PGLQLFYEGKWVTAKCVPDSIVMHIGDTLEILSNGKYKSILHRGLVNKEKVR
Lfil	300	ISWAVFCEPPKEKII LKPLPELVTETEPARFPPTFAQHIH HKLFRKDQEASAQ--SK*
Ljap	300	ISWAVFCEPPKEKII LKPLPELVTETEPARFPPTFAQHIH HKLFRKDQEASAQ--SK*
Gmax	299	ISWAVFCEPPKEKII LQPLPELVTETEPARFPPTFAQHIH HKLFRKDQEGLPN----*
Mtru	299	ISWAVFCEPPKEKII LKPLPELVTETEPARFPPTFAQHIH HKLFRKDDEEEKDDPKK*
Atha	299	ISWAVFCEPPKDKI VLKPLPEMVSVESPAKFPPTFAQHIH HKLFGKEQEELVSEKND*

Figure S21. Amino acid alignment for CHS-2

Lfil	1	MV-----SVAEIRKAQRAEGPATILAIGTANPPNCVDQSTYPDFYFKITNSEHMTTELKEK
Ljap	1	MV-----SVAEIRKAQRAEGPATILAIGTANPPNCVDQSTYPDFYFKITNSEHMTTELKEK
Gmax	1	MV-----SVAEIRQAQRAEGPATILAIGTANPPNRVDQSTYPDYFYFRITNSDHMTTELKEK
Mtru	1	MV-----SVSEIRKAQRAEGPATILAIGTANPNANCVEDQSTYPDFYFKITNSEHMTTELKEK
Atha	1	MVMAGASSLDEIRQAQRADGPAGILAIGTANPENHVLAQAEYPDYFYFRITNSEHMTDLKEK
Lfil	56	FQRMCDKSMIKKRYMYLNEEILKENPNLICAYMAPSLDARQDMVVVEVPRLGKEAATKAIK
Ljap	56	FQRMCDKSMIKKRYMYLNEEILKENPNLICAYMAPSLDARQDMVVVEVPRLGKEAATKAIK
Gmax	56	FQRMCDKSMIKTRYMYLNEEILKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAVKAIK
Mtru	56	FQRMCDKSMIKRRYMYLTTEEILKENPSVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAIK
Atha	61	FKRMCDKSTIRKRHMLTTEEFILKENPHMCAYMAPSLDTRQDI VVVEVPRLGKEAAVKAIK
Lfil	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPSVKRYMMYQQGCFAGGTVLRLAK
Ljap	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPSVKRYMMYQQGCFAGGTVLRLAK
Gmax	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKQLGLRPYVKRYMMYQQGCFAGGTVLRLAK
Mtru	116	EWGQPKSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLRLAK
Atha	121	EWGQPKSKITHVVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLRLAK
Lfil	176	DLAENNKGARVLVVCSEITAVTFRGPNNDTHLDSL VGQALFGDGAAAVIVGSDPVPEI-EK
Ljap	176	DLAENNKGARVLVVCSEITAVTFRGPNNDTHLDSL VGQALFGDGAAAVIVGSDPVPEI-EK
Gmax	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAVIVGSDPIPOV-EK
Mtru	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPVPEI-EK
Atha	181	DLAENNRGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFSDGAAALIVGSDPDTSVGEK
Lfil	235	PLFELVWTAQTIAPDSGAIDGHLREVGLTFHLLKDVP GIVSKNIEKALIEAFQPLGISD
Ljap	235	PLFELVWTAQTIAPDSGAIDGHLREVGLTFHLLKDVP GIVSKNIDKALVEAFQPLNISD
Gmax	235	PLYELVWTAQTIAPDSEGAIDGHLREVGLTFHLLKDVP GIVSKNIDKALFEAFNPLNISD
Mtru	235	PIFEMVWTAQTIAPDSEGAIDGHLREAGLTFHLLKDVP GIVSKNITKALVEAFEPLGISD
Atha	241	PIFEMVSAAQTILPDSGAIDGHLREVGLTFHLLKDVP GLISKNI VKSLDEAFKPLGISD
Lfil	295	YNSIFWIAHPGGPAILDQVEQKLSLKPEKMKATREVLSEYGNMSSACVLFILDEMRRKSA
Ljap	295	YNSIFWIAHPGGPAILDQVEQKLGKPEKMKATRNVLS DYGNMSSACVLFILDEMRRKSA
Gmax	295	YNSIFWIAHPGGPAILDQVEQKLGKPEKMKATRDVLSEYGNMSSACVLFILDEMRRKSA
Mtru	295	YNSIFWIAHPGGPAILDQVEQKLA LKPEKMNATREVLSEYGNMSSACVLFILDEMRRKST
Atha	301	WNSLFWIAHPGGPAILDQVEIKLGLKEKMRATRHVLSEYGNMSSACVLFILDEMRRKSA
Lfil	355	QDGLKTTGEGLEWGVLF GFGPGTLTIETVVLRSVA-I*
Ljap	355	ENGLKTTGEGLEWGVLF GFGPGTLTIETVVLRSVA-I-
Gmax	355	ENGHKTTGEGLEWGVLF GFGPGTLTIETVVLHSVA-I*
Mtru	355	IDGLKTTGEGLEWGVLF GFGPGTLTIETVVLRSVA-I*
Atha	361	KDGVATTGEGLEWGVLF GFGPGTLTIETVVLHSPV-L*

Figure S22. Amino acid alignment for CHS-4

Lfil	1	MV-----SVAEIRKAQRAEGPATIFAIGTANPPNCVDQSTYPDFYFRVTNSEHKTELKEK
Ljap	1	MV-----SVAEIRKAQRAEGPATIFAIGTANPPNCVDQSTYPDFYFRVTNSEHKTELKEK
Gmax	1	MV-----SVAEIRQAQRAEGPATILAIGTANPPNRVDQSTYPDYFRITNSDHTMTELKEK
Mtru	1	MV-----SVSEIRQAQRAEGPATILAIGTANPPNCEQSTYPDFYFKITNSEHKTELKEK
Atha	1	MVMAGASSLDEIRQAQRADGPAGILAIGTANPENHVLAQAEYPDYFRITNSEHMTDLKEK
Lfil	56	FQRMCDKSMIKKRYMHLTEDLLKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAVKAI-
Ljap	56	FQRMCDKSMIKKRYMHLTEDLLKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAITKAIK
Gmax	56	FQRMCDKSMIKTRYMYLNEEILKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAVKAIAK
Mtru	56	FQRMCDKSMIKRRYMYLTEEILKENPSSVEYMAPSLDARQDMVVVEVPRLGKEAAVKAIAK
Atha	61	FKRMCDKSTIRKRHMLTEEFLLKENPHMCAYMAPSLDTRQDIIVVVEVPRLGKEAAVKAIAK
Lfil	114	-----
Ljap	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLR LAK
Gmax	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKQLGLRPYVKRYMMYQQGCFAGGTVLR LAK
Mtru	116	EWGQPKSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLR LAK
Atha	121	EWGQPKSKITHVVVCTTSGVDMPGADYQLTKLLGLRPSVKRLIMMYQQGCFAGGTVLRITAK
Lfil	114	-----
Ljap	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPIPEI-EK
Gmax	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPIPV-EK
Mtru	176	DLAENNKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPVPEI-EK
Atha	181	DLAENNRGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFSDGAAALIVGSDPDTSVGEK
Lfil	114	-----
Ljap	235	PIFEMVWTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGIVSKNIEKALIEAFQPLGISD
Gmax	235	PLYELVWTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGIVSKNIDKALFEAFNPLNISD
Mtru	235	PIFEMVWTAQTIAPDSEGAIDGHLREAAGLTFHLLKDVPGIVSKNINKALVEAFEPLGISD
Atha	241	PIFEMVSAAQTILPDSDDGAIDGHLREVGLTFHLLKDVPGLIISKNIVKSLDEAFKPLGISD
Lfil	114	-----
Ljap	295	YNSIFWIAHPGGPAILDQVEQKLGLKPEKMNATRDVLSNYGNMSSACVLFILDEMRRKSA
Gmax	295	YNSIFWIAHPGGPAILDQVEQKLGLKPEKMKATRDVLSEYGNMSSACVLFILDEMRRKSA
Mtru	295	YNSIFWIAHPGGPAILDQVEQKLALKPEKMKATREVLSEYGNMSSACVLFILDEMRRKSA
Atha	301	WNSLFWIAHPGGPAILDQVEIKLGLKEEKMRA TRHVLSEYGNMSSACVLFILDEMRRKSA
Lfil	114	-----
Ljap	355	QDGLKTTGEGLEWGVLF GFGPGLTIETVVLRSVAI*
Gmax	355	ENGHKTTGEGLEWGVLF GFGPGLTIETVVLHSVAI*
Mtru	355	QDGLKTTGEGLEWGVLF GFGPGLTIETVVLRSVAI*
Atha	361	KDGVATTGEGLEWGVLF GFGPGLTVETVVLHSVPL*

Figure S23. Amino acid alignment for CHS-5

Lfil	1	-----
Ljap	1	MV----SVAEIRKAQRAEGPATIEAIGTANPPNCVDQSTYPDFYFRVTNSEHKTELKEK
Gmax	1	MV----SVAEIRQAQRAEGPATILAIGTANPPNRVDQSTYPDYFRITNSDHMTTELKEK
Mtru	1	MV----SVSEIRKAQRAEGPATILAIGTANPPNCVEQSTYPDFYFRITNSEHKTELKEK
Atha	1	MVMAGASSLDEIRQAQRADGPAGILAIGTANPENHVLQAEYPDYFRITNSEHMTDLKEK

Lfil	1	-----
Ljap	56	FQRMCDKSMIKKRYMHLTEDILKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAVKAIK
Gmax	56	FQRMCDKSMIKTRYMYLNEEILKENPNMCAYMAPSLDARQDMVVVEVPKLGKEAAVKAIK
Mtru	56	FQRMCDKSMIKRRYMYLTEEILKENPSVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAIK
Atha	61	FKRMCDKSTIRKRHMHLEEFILKENPHMCAYMAPSLDTRQDI VVVEVPKLGKEAAVKAIK

Lfil	1	-----
Ljap	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLR LAK
Gmax	116	EWGQPKSKITHLIFCTTSGVDMPGADYQLTKQLGLRPYVKRYMMYQQGCFAGGTVLR LAK
Mtru	116	EWGQPKSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLR LAK
Atha	121	EWGQPKSKITHVVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLR LAK

Lfil	1	-MAENNKGARVLVVCSELTAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPVPEV-EK
Ljap	176	DLAENNKGARVLVVCSELTAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPVPEV-EK
Gmax	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPIPV-EK
Mtru	176	DLAENNKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPIPEI-EK
Atha	181	DLAENNRGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFSDGAAALIVGSDPDTSVGEK

Lfil	59	PLFELVWTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPG-----
Ljap	235	PLFELVWTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGIVSKNIEKALIEAFQPLGISD
Gmax	235	PLYELVWTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGIVSKNIDKALFEAFNPLNISD
Mtru	235	PIFEMVWTAQTIAPDSEGAIDGHLREAAGLTFHLLKDVPEIVSKNINKALVDAFQPLGVSD
Atha	241	PIFEMVSAAQTILPDSGGAIDGHLREVGLTFHLLKDVPGLI SKNIVKSLDEAFKPLGISD

Lfil	97	-----
Ljap	295	YNSIFWIAHPGGPAILDQVEQKLSLKPEKMRATREVLSEYGNMSSACVLFILDEMRRKSA
Gmax	295	YNSIFWIAHPGGPAILDQVEQKLGKPEKMKATRDVLSEYGNMSSACVLFILDEMRRKSA
Mtru	295	YNSIFWIAHPGGPAILDQVEQKLALKPEKMKATREVLSEYGNMSSACVLFILDEMRRKSA
Atha	301	WNSLFWIAHPGGPAILDQVEIKLGLKEKMRATRHVLSEYGNMSSACVLFILDEMRRKSA

Lfil	97	-----
Ljap	355	QDGLKTTGEGLEWGVLF GFGPGLTIETVVLRSVTI*
Gmax	355	ENGHKTTGEGLEWGVLF GFGPGLTIETVVLHSAI*
Mtru	355	QDGLKTTGEGHEWGVLF GFGPGLTIETVVLRSVAI*
Atha	361	KDGVATTGEGLEWGVLF GFGPGLTVETVVLHSPV*

Figure S24. Amino acid alignment for DFR

Lfil	1	MGSVPETVCVTGAAGFIGSWLVMRLMERGYMVRATVRDPANMKKVKHLLLELPEAKTKLTL
Ljap	1	MGSVPETVCVTGAAGFIGSWLVMRLMERGYMVRATVRDPANMKKVKHLLLELPEAKTKLTL
Gmax	1	MGSASESVCVTGASGFIGSWLVMRLIERGYTVRATVRDPVNMKKVKHLVELPGAKSKLSL
Mtru	1	MGSMAETVCVTGASGFIGSWLVMRLMERGYMVRATVRDPENLKKVSHLLELPGAKGKLSL
Atha	1	MVSQKETVCVTGASGFIGSWLVMRLIERGYFVRATVRDPGNLKKVQHLLDLPNAKTLTL
Lfil	61	WKADLAEEGSFDEAIKGTGVFHVATPMDFESKDPENEVIKPTINGVLDIMKASQKAKTV
Ljap	61	WKADLAEEGSFDEAIKGTGVFHVATPMDFESKDPENEVIKPTINGVLDIMKACQKAKTV
Gmax	61	WKADLAEEGSFDEAIKGTGVFHVATPMDFESKDPENEVIKPTINGVLDIMKACIKAKTV
Mtru	61	WKADLGEEGSFDEAIKGTGVFHVATPMDFESKDPENMIKPTIKGVLDIMKACIKAKTV
Atha	61	WKADLSEEGSYDDAINGCDGVFHVATPMDFESKDPENEVIKPTVNGMLGIMKACVKAKTV
Lfil	121	RRLVFTSSAGTLNVIEHQKQMFDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWKFAKE
Ljap	121	RRLVFTSSAGTLNVIEHQKQMFDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWKFAKE
Gmax	121	RRLIFTSSAGTLNVIERQKPVFDDTCWSDVEFCRRVKMTGWMYFVSKTLAEKEAWKFAKE
Mtru	121	RRFIIFTSSAGTLNVTEDQKPLWDESCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWKFAKE
Atha	121	RRFVFTSSAGTVNVEEHQKNVYDENDWSDIEFIMSKMTGWMYFVSKTLAEKAAWDFAEKE
Lfil	181	HGIDFITIIPPLVVGSLFLMPTMPPSLITALSPITGNEAHYSIIKQGQYVHLDDLCLAHIF
Ljap	181	HGIDFITIIPPLVVGSLFLMPTMPPSLITALSPITGNEAHYSIIKQGQYVHLDDLCLAHIF
Gmax	181	QGLDFIT IIPPLVVGPFLLMPTMPPSLITALSPITGNEAHYSIIKQGQFVHLDDLCLAHIF
Mtru	181	HNMDFITIIPPLVVGPFLLPTMPPSLITALSPITGNEAHYSIIKQGQFVHLDDLCEAHIF
Atha	181	KGLDFISLIPTLVVGPFITTSMPPSLITALSPITRNEAHYSIIRQGQYVHLDDLCLNAHIF
Lfil	241	LFEHPESSEGRYICSASEATIHDIAKLINSKYPEYNIPTKFKNIPDELELVRFSKKIKDM
Ljap	241	LFEHPESSEGRYICSASEATIHDIAKLINSKYPEYNIPTKFKNIPDELELVRFSKKIKDM
Gmax	241	LFEHPEVEGRYICSACDATIHDIAKLINQKYPEYKVP TKFKNIPDQLELVRFSKKITD
Mtru	241	LFEHMEVEGRYLCSACEANIHDIAKLINTKYPEYNIPTKFNINIPDELELVRFSKKIKDL
Atha	241	LYEQAAAKGRYICSSHDATILTISKFLRPKYPEYNVPSTFEGVDENLKSIEFSSKKLTDM
Lfil	301	GFEFKYSLEDMYTGAIDTCKEKGLLPKAAENPSN-----
Ljap	301	GFEFKYSLEDMYTGAIDTCKEKGLLPKAAENPSN-----
Gmax	301	GFKFKYSLEDMYTGAIDTCRDKGLLPKPAEKGLF-----
Mtru	301	GFEFKYSLEDMYTEAIDTCIEKGLLPKFVK--ST-----
Atha	301	GFNFKYSLEEMFIESIETCRQKGLFPVLSYQSISEIKTKNENIDVKTGDGLTDGMKPCN
Lfil	335	-----G-----K*
Ljap	335	-----G-----K*
Gmax	335	-----TKPAETPVNAI-MHK*
Mtru	333	-----N-----K*
Atha	361	KTETGITGERTDAPMLAQQMCA*

Figure S25. Amino acid alignment for F3H

Lfil	1	MASFKPKTLTTLAQQNTLESSFVRDEDERPKVAYNNFSNEIPVISLAGIDEVDGRRSEIC
Ljap	1	MASFKPKTLTTLAQQNTLESSFVRDEDERPKVAYNNFSNEIPVISLAGIDEVDGRRSEIC
Gmax	1	MAPT-AKTLYLAQEKTLLESSFVRDEERPKVAYNEFSDEIPVISLAGIDEVDGRRREIC
Mtru	1	MAP--AQTLTYLAQEKTLLESSFVRDEDERPKVAYNNFSNEIPTISLDGIDAGGRRAEIC
Atha	1	-----
Lfil	61	NKIVEACENWGIFQVVDHGVDTLVSHTTTLAKEFFALPPEEKLRFDMTGGKKGGFIVSS
Ljap	61	NKIVEACENWGIFQVVDHGVDTLVSHTTTLAKEFFALPPEEKLRFDMTGGKKGGFIVSS
Gmax	60	EKIVEACENWGIFQVVDHGVDTLVAEMTRLAKEFFALPPDEKLRFDMSGAKKGGFIVSS
Mtru	59	NKIVEACENWGIFQVVDHGVDSKLISEMTRFAKGFFDLPPPEEKLRFDMSGGKKGGFIVSS
Atha	1	-----MTRLARDFALPPEDKLRFDMSGGKKGGFIVSS
Lfil	121	HLQGESVQDWREIVTYFSYPIRNRDYSRWPDTAGWKA VTEEYSEKLMGLACKLLEVLSE
Ljap	121	HLQGESVQDWREIVTYFSYPIRNRDYSRWPDTAGWKA VTEEYSEKLMGLACKLLEVLSE
Gmax	120	HLQGESVQDWREIVTYFSYPKRERDYSRWPDTPEGWRSVTEEYS DKVMGLACKLMEVLSE
Mtru	119	HLQGEAVKDWREIVTYFSYPIRORDYSRWPDKPEGWKEVTEQYSEKLMNGLACKLLEVLSE
Atha	34	HLQGEAVQDWREIVTYFSYPIRNRDYSRWPDKPEGWVKVTEEYSERLMSGLACKLLEVLSE
Lfil	181	AMGLEKEALTKACVDMEQKVVVNYYPKCPQPDLTGLKRHTDPGTITLLLQDQVGGGLQAT
Ljap	181	AMGLEKEALTKACVDMQKVVVNYYPKCPQPDLTGLKRHTDPGTITLLLQDQVGGGLQAT
Gmax	180	AMGLEKEGLSKACVDMQKVVVNYYPKCPQPDLTGLKRHTDPGTITLLLQDQVGGGLQAT
Mtru	179	AMGLEKDALTKACVDMQKVVVNYYPKCPQPDLTGLKRHTDPGTITLLLQDQVGGGLQAT
Atha	94	AMGLEKESLTNACVDMQKIVVNYYPKCPQPDLTGLKRHTDPGTITLLLQDQVGGGLQAT
Lfil	241	RDNGKTWITVQPLEGAFVVNLGDHGHYLSNGRFKNADHQAVVNSNSSRLSIATFQNPAPD
Ljap	241	RDNGKTWITVQPVGAFVVNLGDHGHYLSNGRFKNADHQAVVNSNSSRLSIATFQNPAPD
Gmax	240	RDNGKTWITVQPVGAFAFVVNLGDHAHYLSNGRFKNADHQAVVNSNHSRLSIATFQNPAPN
Mtru	239	KDNGKTWITVQPVGAFVVNLGDHGHYLSNGRFKNADHQAVVNSNYSRLSIATFQNPAPD
Atha	154	RDNGKTWITVQPVGAFVVNLGDHGHFLSNGRFKNADHQAVVNSNSSRLSIATFQNPAPD
Lfil	301	ATVYPLKVREGEKSVMEEPITFAEMYRRKMSKDIELARMKKLAKEK-KLQDLE-----
Ljap	301	ATVYPLKVREGEKSVMEEPITFAEMYRRKMSKDIELARMKKLAKEK-KLQDLE-----
Gmax	300	ATVYPLKIREGEKPVMEEPITFAEMYRRKMSKDIEIARMKKLAKEK-HLQDLENEKHLQE
Mtru	299	ATVYPLKIRIDGEKSVMEEPITFAEMYRRKMSKDIEIARMKKLAKEEKEKLQDLE-----
Atha	214	ATVYPLKVREGEKATILEEPITFAEMYRRKMGRLLELARKKLAKEE-RD-----
Lfil	353	---KAKLEPKPMNEIFA*
Ljap	353	---KAKLEPKPMNEIFA*
Gmax	359	LDQKAKLEAKPLKEILA*
Mtru	352	---KAKIEAKPLNEILA*
Atha	262	----HKEVDKPVQIFA*

Figure S26. Amino acid alignment for F3'H

Lfil	1	M-FPWWIIIGFATITFLIFIHRRVVKF-STR-PSLPLPPGPKWPPIIGNFPHMGVPVPHHSLA
Ljap	1	M-FPWWIIIGFATITFLIFIHRRVVKF-STR-PSLPLPPGPKWPPIIGNFPHMGVPVPHHSLA
Gmax	1	M-SP-LIIVALATIAAAIILIYRIIKF-ITR-PSLPLPPGPKWPPIVGNLPHMGVPVPHHSLA
Mtru	1	M-SLWFI-AIASFTLCILIYRFMKF-AKRSSSLPLPPGPKWPPIIGNMPLGPAPHQSTIA
Atha	1	MATLFLTILLATVLFLLIL--RIFSHRRNRSHNNRLPPGPNWPPIIGNLPHMGTKPHRTLS
Lfil	58	ALARAHGPLMHLKLGFDVVVAASAASVAEQFLKVHDANFSSRPPNAGAKYIAYNYQDLVF
Ljap	58	ALARAHGPLMHLKLGFDVVVAASAASVAEQFLKVHDANFSSRPPNAGAKYIAYNYQDLVF
Gmax	57	ALARAHGPLMHLKLGFDVVVAASASVAEQFLKTHDSNFSSRPPNAGAKYIAYNYQDLVF
Mtru	58	ALAKIHGPLMHLKLGFDVIVAASGSVAEQFLKVHDANFSSRPPNTGAKYIAYNYQDLVF
Atha	59	AMVTTYGPILHLRLGFDVVVAASKSVAEQFLKTHDANFASRPPNSGAKHMAIYNYQDLVF
Lfil	118	APYGARWRYLRKITNLHLFSGKALDNFKHLRQEEVSRLTRNISK--SNSKAVNLGQLLNV
Ljap	118	APYGARWRYLRKITNLHLFSGKALDNFKHLRQEEVSRLTRNISK--SNSKVNLGQLLNV
Gmax	117	APYGPRWRLLRKITSVHLFSGKAMNEFRHLRQEEVARLTCNLAS--SDTKAVNLGQLLNV
Mtru	118	APYGPRWRLRKISSVHLFSNKVMEEFKHLRQEEVARLTSNLASNYSDTKAVNLGQLLNV
Atha	119	APYGHRWRLLRKISSVHLFSAKALEDFKHVRQEEVGTLTREIVR--VGTKPVNLGQLVNM
Lfil	176	CTTNALS RVMIGRRVFNDG DGGCDPRADEFKAMVVELMVLAGVFNIGDFIPSLEWLDLQG
Ljap	176	CTTNALS RVMIGRRVFNDG DGGCDPRADEFKAMVVELMVLAGVFNIGDFIPSLEWLDLQG
Gmax	175	CTTNALARAMIGRRVFNDGNGGCDPRADEFKAMVMEVMVLAGVFNIGDFIPSLEWLDLQG
Mtru	178	CTTNALARVMI GRRVFNDGNGGCDPKADEFKEMVLELMVLAGVFNISDFIPSLEWLDLQG
Atha	177	CVVNALGREMIGRRIFG---ADADHKADEFRSMVTEMMALAGVFNIGDFVPSLDWLDLQG
Lfil	236	VQAKMKKLHNRFD EFLTSIIEEH-NTSSKSENHKDLLSTLLSLKDVP-DDDGNRLNDIEI
Ljap	236	VQAKMKKLHNRFD EFLTSIIEEH-NTSSKSENHKDLLSTLLSLKDVP-DDDGNKLNNDIEI
Gmax	235	VQAKMKKLHKRFDAFLTSIIEEHNSSSKNENHNKFLSILLSLKDVR-DDHGNHLTDTEI
Mtru	238	VQAKMKKLHKKRFDAFLTNIIDERENS NFKSEKHKDLLSTLLLLKEET-DVDGNKLTYTEI
Atha	234	VAGKMKRLHKRFDAFLSSILKEH-EMNGDQKHTDMLSTLISLKGTDLDG DGSGLTDTEI
Lfil	294	KALLLLNMFTAGTDTASTTEWAI AELIRSPRILAQVQQELDTVVGRERNVREDDLPHPY
Ljap	294	KALLLLNMFTAGTDTASTTEWVIAELIRSPRILAQVQQELDTVVGRERNVREDDLPHPY
Gmax	294	KALLLLNMFTAGTDTSSSTTEWAI AELIKNPQILAKLQQELDTVVGRDRSVKEEDLAHPY
Mtru	297	KALLLLNMF AAGTDTSSSTTEWAI AELIRNPRILAQVQQELDNVVGDRNVKEDDIPNLPY
Atha	293	KALLLLNMFTAGTDTASTVDWAIAELIRHPDIMVKAQEELDIVVGRDRPVNESDIAQLPY
Lfil	354	LQAVVKETFRLLHPSTPLSLPRVASESCEVLGYHIPKGSTLLVNVWAIARDPKEWAEPLF
Ljap	354	LQAVVKETFRLLHPSTPLSLPRVASESCEVLGYHIPKGSTLLVNVWAIARDPKEWAEPLF
Gmax	354	LQAVIKETFRLLHPSTPLSVPRAAAESCEIFGYHIPKGATLLVNIWAIARDPKEWNDPLF
Mtru	357	LQAVIKETFRLLHPSTPLSLPRIASESCEIFGYHIPKGSTLLVNVWAIARDPKEWVDPLEF
Atha	353	LQAVIKENFRLLHPPTPLSLPHIASESCEINGYHIPKGSTLLTNIWAIARDPDQWSDPLAF
Lfil	414	KPERF-LEGDKVDVDVKGNDFFKVI PFGAGRRICAGMSLGLRMVQLLTATLVHSFNWELEN
Ljap	414	KPERF-LEGDKVDVDVKGNDFFKVI PFGAGRRICAGMSLGLRMVQLLTATLVHSFNWELEN
Gmax	414	RPERFLLGGEKADVDVRGNDFFVI PFGAGRRICAGISLGLQMVQLLTAALAHSFDEWELE
Mtru	417	KPERFLPGGEKCDVDVKGNDFFVI PFGAGRRICPGMSLGLRMVQLLTATLVHSFDWELEN
Atha	413	KPERFLPGGEKSGVDVKGSDFEIIPFGAGRRICAGISLGLRTIQFLTATLVQGFDEWELAG

```
Lfil 473 GLNHEKLNMD EAYGLTLQRAVPLSVYSRPRLSPHVYAASH-*
Ljap 473 GLNHEKLNMD EAYGLTLQRAVPLSVYSRPRLSPHVYAASH-*
Gmax 474 CMNPEKLNMD EAYGLTLQRAVPLSVHPRPRLAPHVYSMSS-*
Mtru 477 GLNAGKMNMD EGYGLTLQRAVPLSVHPKPRRLSPHVYSSCF-*
Atha 473 GVTPEKLNME ES YGLTLQRAVPLV VHPKPR LAPNVYGLGSG*
```

Figure S27. Amino acid alignment for MAX1

Lfil	1	MVFMDFEWLFQIPS-VPWSSAMFTLLATIG-GFSVYLYGPYWGVRKVPGPSPVPLIGHLP
Ljap	1	MVFMDFEWLFQIPS-VPWSSAMFTLLATIG-GFLVYLYGPYWGVRKVPGPSPVPLIGHLP
Gmax	1	MVSIVLEWLFPEPC-VAMF---TTLMLLIG-GLLGYLYGPYWGIRKVPGPPTLPLVGHLH
Mtru	1	MVFMDLEWLFPIPISSVFA---STLLALAG-GWLIYLYEPYWRVRKVPGPSPVPLVGHLH
Atha	1	MKTQHQQWWEVLDPF-LTQHEALIAFLTFAAVVIVIIYLYRPSWSVCNVPGPTAMPLVGHLH
Lfil	59	LLAKYGPDVFSVLAKQYGPIYRFHMGRQPLIIADAELCKEAGIKKFKDITNRSIPSPIS
Ljap	59	LLAKYGPDVFSVLAKQYGPIYRFHMGRQPLIIADAELCKEAGIKKFKDITNRSIPSPIS
Gmax	56	LLAKYGPDVFSILAKQYGPIYRFHMGRQPLIIVADPELCKEVEGIKKFKDIPNRSIPSPIS
Mtru	57	LLAKHGPDVFSVLAKQYGPIYRFHMGRQPLIIVADAELCKEVEGIKKFKDIPNRSIPSPIS
Atha	60	LMAKYGPDVFSVLAKQYGPIIFRFQMGGRQPLIIIAEAELCREVEGIKKFKDLNRSIPSPIS
Lfil	119	ASPLHQKGLFFTKDSQWSTMRNTILSLYQPSHLSRLVPTMQSFIESATQNLDQKEDFIIF
Ljap	119	ASPLHQKGLFFTKDSQWSTMRNTILSLYQPSHLSRLVPTMQSFIESATQNLDQKEDFIIF
Gmax	116	ASPLHQKGLFFTRDSRWSTMRNTILSVYQPSHLASLVPTMQSFIESATQNLDTPNEDIIF
Mtru	117	ASPLHQKGLFFSRDSQWSTMRNTILSVYQPSHLSRLVPTMQSFIESATQNLDQKEDFIIF
Atha	120	ASPLHKKGLFFTRDKRWSKMRNTILSLYQPSHLTSLIPTMHSFITSATHNLDSPKPRDIVF
Lfil	179	SNLSLSLATDVIGQAAFGVDFGLSKPQPVCDEIKSVN---KEVRDSSTGNEVSDFINQHI
Ljap	179	SNLSLSLATDVIGQAAFGVDFGLSRPQSVRDESG--N---KEVRGSGAGNEVSDFINQHI
Gmax	176	SNLSLRLATDVIGEAAGFVNFGLSKPHSVCSIKSVSV--NNVR--NDDDEVSDFINQHI
Mtru	177	SNLSLKLATDVIGQAAFGVNFGLSQSHSVHNESKNVATDNKDLMNASGSNEVTDFINQHI
Atha	180	SNLFLKLTTDTIIGQAAFGVDFGLSGKKPIKD-----VEVTDFINQHIV
Lfil	236	YSTTQLKMDLSGSFSIILGLLVPILQEPFRQVLKRIPGTMDWKIERTNKRKLSGRLDEIVE
Ljap	234	YSTTQLKMDLSGSFSIILGLLVPILQEPFRQILKRIPGTMDWKIECTNKRKLSGRLDEIVE
Gmax	232	YSTTQLKMDLSGSFSIILGLLAPILQEPFRQILKRIPGTMDSKIESTNEKLSGRLDEIVK
Mtru	237	YSTTQLKMDLSGSFSIILGLLVPILQEPFRQILKRIPGTMDWKIERTNEKLGRLDEIVE
Atha	222	YSTTQLKMDLSGSLSIILGLLIPILQEPFRQVLKRIPGTMDWRVEKTNARLSGQLNEIVS
Lfil	296	KRMKDRVRSSKDFLSLILNARESKTVSENVFTPDYISAVTYEHLLAGSATTSTFTLSSIVY
Ljap	294	KRMKDKVRSSKDFLSLILNARESKTVSENVFTPDYISAVTYEHLLAGSATTSTFTLSSIVY
Gmax	292	RRMEDKNRTISKNFLSLILNARESKKVSSENVFSPDYISAVTYEHLLAGSATTSTFTLSSIVY
Mtru	297	KRTKDRTRSSKDFLSLILNARESKAVSENVFTPEYISAVTYEHLLAGSATTSTFTLSSIVY
Atha	282	KRAKEAETDSKDFLSLILKARESDPFAKNIFTSDYISAVTYEHLLAGSATTSTFTLSSIVY
Lfil	356	LVAGHPEVEKKMLQEIDGFGPVDQTPTSQDLQEKFPYLDQVIKEAMRYTTSPLVARETS
Ljap	354	LVAGHPEVEKKMLQEIDGFGPVDQTPTSQDLQEKFPYLDQVIKEAMRYTTSPLVARETS
Gmax	352	LVAGHREVEKKILQEIDGFGPPDRIPTAQDLHDSFPYLDQVIKEAMRYTTSPLVARETS
Mtru	357	LVAHPEVEKKMLEEIDGYGSLDQIPTSQDLHDKFPYLDQVIKEAMRYTTSPLVARETS
Atha	342	LVSGLLDVEKRLQEIDGFGNRDLIPTAHLQHKFPYLDQVIKEAMRYTTSPLVARETA
Lfil	416	NEVEIGGYLLPKGTWVWLALGVVAKDPRNFPEPEKFKPERFDPKCEEMKRRHPYAFIPFG
Ljap	414	NEVEIGGYLLPKGTWVWLALGVVAKDPRNFPEPEKFKPERFDPKCEEMKRRHPYAFIPFG
Gmax	412	NEVEIGGYLLPKGTWVWLALGVAKDPRNFPEPEKFKPERFDPKCEEMKRRHPYAFIPFG
Mtru	417	NEVEIGGYLLPKGTWVWLALGVAKDHNKFNPEPEKFKPERFDPNCEEMKRRHPYAFIPFG
Atha	402	KEVEIGGYLLPKGTWVWLALGVAKDPKNFPEPEKFKPERFDPNGEEKRRHPYAFIPFG

Lfil	476	IGPRACIGQKFSLQEIKLSLIHLYRKYLFRHSPNMENPLELEYGIVLNFKHGVKVRRAIKR
Ljap	474	IGPRACIGQKFSLQEIKLSLIHLYRKYLFRHSPNMENPLELEYGIVLNFKHGVKVRRAIKR
Gmax	472	IGPRACIGQKFSLQEIKLTLIHLYQKYVFRHSVDMKEPVEMEYGMVLNFKHGIKLRVIRR
Mtru	477	IGPRACIGQKFSMQEIKLSLIHLYKKYLFRHSADMESPLELEYGIVLNFKHGVKFSVIKR
Atha	462	IGPRACVGQRFALQEIKLTLHLYRNYIFRHSLEMEIPLQLDYGIILSFKNGVKLRITIKR
Lfil	536	TERSC*
Ljap	534	TERSC*
Gmax	532	T----*
Mtru	537	TEMSC*
Atha	522	F----*

Figure S28. Amino acid alignment for UFGT73C5

Lfil	1	MVFQEN---QPHFILFPLMAQGHIIIPMVDIARLLAQRGVIVTIFTTPKNASRFTSVLSRA
Ljap	1	-----
Gmax	1	MVFQTN--NNPHFILFPLMAQGHIIIPMMDIARLLAHRGVIVTIFTTPKNASRFNSVLSRA
Mtru	1	MVLPANINDVPHFVLFPLIAQGHIIIPMIDIAKLLAQRGVIVTIFTTPKNASRFTSVLSRA
Atha	1	MVSETTKSSPLHFVLFPFMAQGHMIPMVDIARLLAQRGVITITIVTTPHNAARFKNVLNRA
Lfil	58	ISSGLQIRLVELQFPYKQAGLPEGCECFDMATSKGMFSKIFRGITMLQNSAEELFEKLSP
Ljap	1	-----EELFEKLSP
Gmax	59	ISSGLQIRLVQLHFPSKEAGLPEGCECFDMVTSIDMVYKMFNVINMLHKQAEFFFEALTTP
Mtru	61	VSSGLQIKIVTLNFPKQAGLPEGCECFDMVDSIDMRMNLFAITLLQKPAEELFDALTTP
Atha	61	IESGLPINLVQVKFPYLEAGLQEGQENIDSLDTMERMIPFFKAVNFLEEPVQKLIIEEMNP
Lfil	118	RPSCIISDFCIPWTAQVAAKYNIPRISFHGFSCFCLHCLLKVQT-SKVTESITSETEYFT
Ljap	10	RPSCIISDFCIPWTAQVAAKYNIPRISFHGFSCFCLHCLLKVQT-SKVTESITSETEYFT
Gmax	119	KPSCIISDFCIPWTAQVAQKHCIPRISFHGFACFCLHCLMLVHT-SNVCESTASESEYFT
Mtru	121	KPSCIISDFEIPWTIQIAENHKIPRISFHGFSCFCLHCLMKIQT-SKILERVDSSEYFT
Atha	121	RPSCIISDFCIPYTSKIAKKNIPKILFHGMGCFCLLCMHVLRKNREILDNLKSDKELET
Lfil	177	VPGIPDQIQVNKSQIPGPM---S---DDLKDYGEQMYEAEKSSYGVIIINTFEELNAYVK
Ljap	69	VPGIPDQIQVNKSQIPGPM---S---DDLKDYGEQMYEAEKSSYGVIIINTFEELNAYVK
Gmax	178	IPGIPDQIQVTKEQIPMMI---SNSDEEMKHFRQMRDADIKSYGVIIINTFEELKAYVR
Mtru	180	VPGIPDQIQVTKEQIPGIL---K---GELKEFGKMHDAEMKSYGEIINTFEELKAYVK
Atha	181	VPDFPDRVEFTRTQVPVETYPVA---GDWKDIFDGMVEANETSYGVIVNSFQELEPAYAK
Lfil	231	DYKKERNDKVWCIGPVSLCNKDGLDKAQRGNKASINEHHCLKWLDLQQPKSVVYVCLGSL
Ljap	123	DYKKERNDKVWCIGPVSLCNKDGLDKAQRGNKASINEHHCLKWLDLQQPKSVVYVCLGSL
Gmax	235	DYKKVRNDKVWCIGPVSLCNQDNLDKVQRGNHASINEHHCLKWLDLQPPKSAVYVCFGLSL
Mtru	234	DYKKEKNGKVWFVGPVSLCNKDGLDKAQRGIITASISEHHCLKWLDLHQPKSVVYACLGLSL
Atha	238	DYKEVRSGKAWTIGPVSLCNKVGADKAERGNKSDIDQDECLKWLD SKKHG SVLYVCLGSI
Lfil	291	CNLIPSQLMELALALEATERPFIWVIREGNQF-QELE-KWMSEEGFEERTKGRGLVIRGW
Ljap	183	CNLIPSQLMELALALEATERPFIWVIREGNQF-QELE-KWMGEKGFEEERTKGRGLVIRGW
Gmax	295	CNLIPSQLVELALALEDTKKPFVWVIREGNKF-QELEKKWISSEEGFEERTKGRGLIIRGW
Mtru	294	CNLIPSQLMELALALEATNRPFIWVIREGNKSSEELE-KWISE----ERNKGRGLIIRGW
Atha	298	CNLPLSQLKELGLGLEESQRPFIWVIRGWKEY-KELV-EWFSSEGFEDRIQDRGLLIKGW
Lfil	349	APQVLILSHPAIGGFLTHCGWNSTLEGISAGVPMVTWPLFADQFLNEKLVTQVLRIGVSL
Ljap	241	APQVLILSHPAIGGFLTHCGWNSTLEGISAGVPMVTWPLFADSF-EENLSRPPGRIPVSX
Gmax	354	APQVLILSHPSIGGFLTHCGWNSTLEGISAGVPMITWPLFADQFLNEKLVTQVLKIGVSV
Mtru	349	APQVLILSHPSIGGFLTHCGWNSTLEGISAGIPIVTWPLFADQFLNEKLVTQVLRIGVSL
Atha	356	SPQMLILSHPSVGGFLTHCGWNSTLEGITAGLPLLTWPLFADQFCNEKLVEVLKAGVRS
Lfil	409	GVEVPLKKGEEEEKTGVLVKKEDIKKGICLLMDDESKESEERRQKARELSEIAKKAVENG
Ljap	300	RHMLPLSIRKIE-----
Gmax	414	GMEVPMKFGEEEEKTGVLVKKEDIKRAICIVMDDGGEESKDRRERATKLSEIAKRAVEKEG
Mtru	409	GVEVPMRLGVESLGVLVKKEGIKEAICMVMD-EGEESKERRERASKLSEMAKRAVEKGG
Atha	416	GVEQPMKKGEEEEKIGVLVDKEGVKKAVEELMG-ESDDAKERRRRRAKELGDSAHKAVEEGG

Lfil	469	SSYHNM	TL	LLIQDIMQ	-----
Ljap	311	-----	-----	-----	-----
Gmax	474	SSHLD	MT	LLIQDIMQ	QSSSKEEVRTSLKESQFSCNA*
Mtru	468	SSHLD	MT	LLIQDIMQ	-----*
Atha	475	SSHSN	ISFL	LLIQDIME	LAEP-----NN*

Appendix 8

Supplementary Figures S29 – S30 Amino acid alignments for yellow-differentially expressed flavonoid genes in *Lotus filicaulis*

Figure S29. Amino acid alignment for CHS-1

Lfil	1	MV-----TVEEIRNAQRSSGPATILAFGTATPSHCVMQADYPDYYFRITNSEHMTDLKEK
Ljap	1	MV-----TVEEIRNAQRSSGPATILAFGTATPSHCVMQADYPDYYFRITNSEHMTDLKEK
Gmax	1	MV-----TVEEIRNAQRSHGPATILAFGTATPSNCVSOADYPDYYFRITNSEHMTDLKEK
Mtru	1	MV-----TVEEIRKAQRSNGPATILAFGTATPSHCVTQAEYPDYYFRITNSEHMTDLKEK
Atha	1	MVMAGASSLDEIRQAQRADGPAGILAI GTANPENHVLQAEYPDYYFRITNSEHMTDLKEK
Lfil	56	FKRMCEKSMIRKRYMHLTEETILKENPAMCAYMAPSLDARQDLVVVEVPKLGKEAAAKAIK
Ljap	56	FKRMCEKSMIRKRYMHLTEETILKENPAMCAYMAPSLDARQDLVVVEVPKLGKEAAAKAIK
Gmax	56	FKRMCEKSMIKKRYMHLTEETFLKENPNMCEYMAPSLDVQRDVVMEVPKLGKQAATKAIK
Mtru	56	FKRMCEKSMIKKRYMHLTEETFLKENPNMCAYMAPSLDARQDLVVVEVPKLGKDAAKKAIA
Atha	61	FKRMC DKSTIRKRHMHLTEETFLKENPHMCAYMAPSLDTRQDI VVVEVPKLGKEAAVKAIK
Lfil	116	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLKPSVKRLMMYQQGCFAGGTVLRLAK
Ljap	116	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLKPSVKRLMMYQQGCFAGGTVLRLAK
Gmax	116	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLRLAK
Mtru	116	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLKPSVKRLMMYQQGCFAGGTVLRLAK
Atha	121	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLRLAK
Lfil	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAIIVGADPDLAV-ER
Ljap	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAIIVGADPDLAV-ER
Gmax	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAALIIGSDPDPAV-ER
Mtru	176	DLAENNNARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAMIIGADPDLTV-ER
Atha	181	DLAENNRGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFSDGAAALIVGSDPDTSVGEK
Lfil	235	PIFQIVSAGQTILPDSG AIDGHLREVGLTFHLLKDVPGIISKHIEKSLSEAFAPIGISD
Ljap	235	PIFQIVSAGQTILPDSG AIDGHLREVGLTFHLLKDVPGIISKHIEKSLSEAFAPIGISD
Gmax	235	PIFEMI SAAQTILPDSG AIDGHLREVGLTFHLLKDVPGIISKHIEKSLSEAFAPIGISD
Mtru	235	PIFEIVSAAQTILPDSG AIDGHLREVGLTFHLLKDVPGIISKHIEKSLSEAFAPIGISD
Atha	241	PIFEMVSAAQTILPDSG AIDGHLREVGLTFHLLKDVPGIISKHIVKSLDEAFKPLGISD
Lfil	295	WNSLFWIAHPGGPAILDQVEAKLRLKEEKL RSTRHVLGEYGNMSSACVLFILDEVRRRSK
Ljap	295	WNSLFWIAHPGGPAILDQVEAKLRLKEEKL RSTRHVLGEYGNMSSACVLFILDEVRRRSK
Gmax	295	WNSIFWIAHPGGPAILDQVEEKRLRLKEEKL RSTRHVLSEYGNMSSACVLFILDEMRRKRSK
Mtru	295	WNSIFWVAHPGGPAILDQVEEKRLRLKEEKL RSTRHVLSEYGNMSSACVLFILDEMRRRSK
Atha	301	WNSLFWIAHPGGPAILDQVEIKLGLKEEKL RSTRHVLSEYGNMSSACVLFILDEMRRKSA
Lfil	355	EEGKETTTGDGLEWGVLF GFGPGLTVETTVVLH SVPLEAX
Ljap	355	EEGKETTTGDGLEWGVLF GFGPGLTVETTVVLH SVPLEAX
Gmax	355	EEGKSTTGEGLWGVLF GFGPGLTVETTVVLH SVPLEGX
Mtru	355	EEGKITTTGEGLWGVLF GFGPGLTVETTVVLH SVPVQGX
Atha	361	KDGVATTGEGLEWGVLF GFGPGLTVETTVVLH SVPVIX--

Figure S30. Amino acid alignment for PAP-1

```

Lfil  1 MDLETLYSPCFM---SNSNWVFQESAHNTEWSREDNKRFEALAIY-DKDTTPDRWLNVA
Ljap  1 MDLETLYSPCFM---PNSNWVFQESAHSTEWREDNKRFEALAIY-DKDTTPDRWLNVA
Gmax  1 MELETLYPPCFM---PNSNWVFQES-HSTEWREDNKRFEALAIY-DNDTPDRWEKVAA
Mtru  1 MDLETLYSPYFM---LDTNWLVEES-YQTEWSREDNKRFEALAIY-DKDTTPDRWLKVAE
Atha  1 MSSSTMYRGVNMFSNNTINWIFQEV-REATWTAEENKRFEKALAYLDDKDNLESWSKIAD

Lfil  57 MIPGKTVLDDVIKQYRELEEDVGEIEAGHVPVPGYHSSS-----FTFEVVENQNFD---
Ljap  57 MIPGKTVLDDVIKQYRELEEDVGEIEAGHVPVPGYHSSS-----FTFEVVENQNFD---
Gmax  56 MIPGKTVFDVIKQYRELEEDVSEIEAGRVPIPGYLASS-----FTFEIVDNNHNYD---
Mtru  56 MIPGKTVFDVIKQYRELVEDVSEIEAGNVPIPGYLASS-----FTFEVVEKQNYD---
Atha  60 LIPGKTVADVIRYKELEDVDSIEAGLIPIPGYGGDASSAANSDFGLENSSYGYDYV

Lfil  107 --GLKRKPGTT--LR--GSDHERKKGVPTWEEHKKRFLMGLLKYGKGDWNRNIARNFVMTK
Ljap  107 --GLKRKPGTT--LR--GSDHERKKGVPTWEEHKKRFLMGLLKYGKGDWNRNIARNFVMTK
Gmax  106 --GCRRL-AP--VR--GSDQERKKGVPTWEEHRRFLMGLLKYGKGDWNRNISRNFFVTK
Mtru  106 --GNRRRH-VT--VR--GSDHERKKGVPTWEEHRRFLMGLLKYGKGDWNRNISRNFFVTK
Atha  120 VGGKRSSPAMTDCFRSPMPEKERKKGVPTWEEHRLFLMGLIKKYGKGDWNRNIAKSEVTTTR

Lfil  161 TPTQVASHAQKYYIRQKV-SGGKDKRRPSIHDITTVNLTETSTTSENNKLPSFK-----
Ljap  161 TPTQVASHAQKYYIRQKV-SGGKDKRRPSIHDITTVNLTETSTTSENNKLPSFK-----
Gmax  159 TPTQVASHAQKYYIRQKV-SGGKDKRRPSIHDITTVNLTETSASDKN-KPQLFNASPVLA
Mtru  159 TPTQVASHAQKYYIRQKVS-SGGKDKRRPSIHDITTVTLTETSSPSEN-KSLLVNVSPMQQ
Atha  180 TPTQVASHAQKYFLRQL--TDGKDKRRSSIHDITTVNIPDADASATA-TTADVALSP---

Lfil  214 -----STSNYNSGSLMVF-NPNCDDLMPSSSDIITSKTLKLGQDLYD
Ljap  214 -----STSNYNSGSLMVF-NPNCDDLMPSSSDIITSKTLKLGQDLYD
Gmax  217 PQQKLNSISKVQLGWTSSHYNDGSFMVF-NPNSDALFV--SSSPDVTSMALKMQGQDLYD
Mtru  218 -----KMGWSTSHYNDGS-----PQGQDLYD
Atha  234 -----TPAN--SFDVFLQPN-----

Lfil  258 CSLHEAYAKLIKIPSFRAAPRNFNKEAVFGIHVL*
Ljap  258 CSLHEAYAKSKIPSFRAAPRNFNKEAVFGIHVL*
Gmax  274 CALHEAYAKVKVPGFSMAPRDFNNEAVFGIHAL*
Mtru  239 CSFHEAYAKLKVSGFATASRDFNKGAVFGIHAL*
Atha  247 --PHYSFASASASSYNAFPQWS-----*

```

Appendix 9

Supplementary Figures S31 – S34 Amino acid alignments for non-differentially expressed lignin genes in *Lotus filicaulis*

Figure S31. Amino acid alignment for CCR

Lfil	1	-----
Ljap	1	-----
Gmax	1	MPT--DTSSVSG--EIVCVTGAGGF ¹ IASWL ² VKILL ³ LEKGYTVRGTVRNPD ⁴ DPKNGHLKELE
Mtru	1	MPAYDNTSSVSGGDQTV ¹ CVTGAGGF ² IASWL ³ VKILL ⁴ ERGYTVRGTVRNPD ⁵ DPKNGHLKELE
Atha	1	MPV--DVAS ¹ PAG--KT ² VCVTGAGGY ³ IASW ⁴ IVKILL ⁵ ERGYTVKGTVRNPDDPKNTHLRELE

Lfil	1	-----
Ljap	1	-----
Gmax	57	GGKERLT ¹ LHKVD ² LFDIAS ³ IKAA ⁴ LHGCHGVFHTASPVTDN ⁵ PEEMVEPAVKGT ⁶ KNV ⁷ IIAAAE
Mtru	61	GARERLT ¹ LHKVD ² LLDLQSI ³ QSVVHGCHGVFHTASPVTDN ⁴ PDEM ⁵ LEPAVNGTKNV ⁶ IIASAE
Atha	57	GGKERLT ¹ LCKAD ² LQDYEA ³ LKAAT ⁴ DGCDGVFHTASPVTD ⁵ DPEQMVEPAVNGAKFV ⁶ INAAAE

Lfil	1	-----
Ljap	1	-----
Gmax	117	AKVRRVVF ¹ TSSIGTVYMDPN ² TSRDALVDESF ³ WSDLEYCKNTKNWYCYGKT ⁴ VAEQA ⁵ AWDVA
Mtru	121	AKVRRVVF ¹ TSSIGTVYMDPN ² TSRDVVVDESYWSDLEHCKNTKNWYCYGKT ³ VAEQSAWDIA
Atha	117	AKV ¹ KRVVIT ² TSSIGAVYMDPN ³ RDPEAVVDESCWSDLD ⁴ FCNTKNWYCYGKM ⁵ VAEQA ⁶ AWETA

Lfil	1	-----
Ljap	1	-----
Gmax	177	KERGVDLVV ¹ NPVLV ² IGPLLQPTINASTI ³ HLKYL ⁴ TGSAKTYVNATQAYVHVRDVALAHI
Mtru	181	KENQVDLVV ¹ NPVVLGPLLQPTINASTI ² HLKYLNGAAKTYVNATQSYVHVKDVALAHL
Atha	177	KEKGV ¹ DLVVLN ² PVLVLGPP ³ LQPTINASLYHVLKYL ⁴ TGSAKTYANLTQAYVDVRDVALAHV

Lfil	1	-----PRVKPYIFSNQKLK
Ljap	1	-----CSDEKNPRVKPYIFSNQKLK
Gmax	237	LVYETPSASGRFICAESS ¹ LHRGELVEILAK ² FFPEYPIPTKCSDEKNPRVKPYIFSNQKLK
Mtru	241	LVYETNSASGRYICCE ¹ TALHRGEVVEILAKY ² FFPEYPLPTKCSDEKNPRVKPYK ³ FSNQKLK
Atha	237	LVYEAPSASGRYLLAES ¹ ARHRGEVVEILAKL ² FFPEYPLPTKCK ³ DEKNPRAKPYK ⁴ FTNQKIK

Lfil	15	DLGLEFTPVKQCLYE ¹ TVKNLQEKGHLPVLIN---Q ² QEL-----I*
Ljap	21	DLGLEFTPVKQCLYE ¹ TVMNLQEKGHLPVLIN---Q ² QEL-----I-
Gmax	297	DLGLEFTPVKQCLYD ¹ TVKNLQEN ² GHLPVPPK---Q ³ KDS-----Y*
Mtru	301	DLGLEFTPVKQCLYD ¹ TVRS ² LQEKGHLP ³ IPPM---Q ⁴ EDS-----A*
Atha	297	DLGLEFTSTKQSLYD ¹ TVKSL ² LQEKGHLAPPPPPPSASQ ³ ESVENGIGKIGS*

Figure S32. Amino acid alignment for F5H

Lfil	1	MDSLPTWQTIANLRHEQLPMAFLFMVPMILLGLASRIRKRPPYPPGPKGLPIIGNMLMM
Ljap	1	MDSLPTWQTIANLRHEQLPMAFLFMVPMILLGLASRIRKRPPYPPGPKGLPIIGNMLMM
Gmax	1	M-----ANLDDLDPFQTSILILVPIALLVALLSRTRRRAPYPPGPKGLPIIGNMLMM
Mtru	1	MDSLILKSPIMENLKEEPFLMAIMFIVPLILLGLVSRILKRPRYPGPIGLPIIGNMLMM
Atha	1	MESSI-SQTL SKLSDP--TTSLVIVVSLFIFISFITR-RRRPPYPPGPRGWPIIGNMLMM
Lfil	61	DQLTHRGLATLAKQYGGIFHLRMGFIHMAVSDPDAARQVLQVHDNIFSNRPATIAISYL
Ljap	61	DQLTHRGLATLAKQYGGIFHLRMGFIHMAVSDPDAARQVLQVHDNIFSNRPATIAISYL
Gmax	52	EQLTHRGLANLAKHYGGIFHLRMGFLHMAISDPVAARQVLQVQDNIFSNRPATIAISYL
Mtru	61	DQLTHRGLANLAKKYGGIFHLRMGFLHMAISDAARQVLQVQDNIFSNRPATIAIKYL
Atha	57	DQLTHRGLANLAKKYGGILCHLRMGFLHMYAVSSPEVARQVLQVQDSVFSNRPATIAISYL
Lfil	121	TYDRADMAFAHYGPFWRQMRKLCVMKLF SRKRAESWQSVRDEVDAAVRDVDQNTGTTPVNI
Ljap	121	TYDRADMAFAHYGPFWRQMRKLCVMKLF SRKRAESWQSVRDEVDAAVRDVDQNTGTTPVNI
Gmax	112	TYDRADMAFAHYGPFWRQMRKLCVMKLF SRKRAESWQSVRDEVDAAVRAVASSVGKPVNI
Mtru	121	TYDRADMAFAHYGPFWRQMRKLCVMKLF SRKHAEWQSVRDEVDAVAVRTVSDNIGNPVNI
Atha	117	TYDRADMAFAHYGPFWRQMRKV CVMKVFSRKRAESWASVRDEVDKMVRVSVSCNVGKPINV
Lfil	181	GEVVFNLTKNIIYRAAFGSSSKEGQDEFIGILQEFSKLF GAFNIADFIPYI GAIDPQGLN
Ljap	181	GEVVFNLTKNIIYRAAFGSSSKEGQDEFIGILQEFSKHFGAFNIADFIPYI GAIDPQGLN
Gmax	172	GELVFNLTKNIIYRAAFGSSSQEGQDE-----LN
Mtru	181	GELVFNLTKNIIYRAAFGSSSREGQDEFIGILQEFSKLF GAFNISDFVPCFGAIDPQGLN
Atha	177	GEQIFALTRNITYRAAFGSACEKGQDEFIRILQEFSKLF GAFNVADFIPYFGWIDPQGIN
Lfil	241	ARLVKARAALDSFIGKIIDEHVQKKRGGGGGV DGD-GETDMVDELLAFYTEEEVAGSKGS
Ljap	241	ARLVKARAALDSFIGKIIDEHVQKKRGGGGGV DGD-GETDMVDELLAFYTEEEVAGSKGS
Gmax	201	SRLARARGALDSFSDKIIDEHVHKMKNDKSSEIVD-GETDMVDELLAFYSEEAK-----L
Mtru	241	ARLVKARKELD SFIDKIIDEHVQKKKS-----VVD-EETDMVDELLAFYSEEAK-----L
Atha	237	KRLVKARNDL DGFIDDIIDEHMKKKENQNAVD DGDVVD TDMVDDLAFYSEEAK-----L
Lfil	300	SNESED LQNSIKLTKDNIKAIIMDV MFGGTETVASAIEWAISELMKTPEELKRVQQELAD
Ljap	300	SNESED LQNSIKLTKDNIKAIIMDV MFGGTETVASAIEWAMSELMKTPEELKRVQQELAD
Gmax	255	NNESDDLQNSIRLTKDNIKAIIMDV MFGGTETVASAIEWAMAELMRSPEDQKRVQQELAD
Mtru	290	NNESDDL NNSIKLTKDNIKAIIMDV MFGGTETVASAIEWAMAELMKSPEDLKKVQQELAE
Atha	292	VSETAD LQNSIKLTRDNIAIIMDV MFGGTETVASAIEWALTELLRSPEDLKRVQQELAE
Lfil	360	VVGLDRVVEESDFEKLTHLKCALKETLR LHPPILLHESAEADVTVGGYFIPKKARVMIN
Ljap	360	VVGLDRVVEESDFEKLTHLKCALKETLR LHPPILLHETAEDVTIGGYFIPKKARVMIN
Gmax	315	VVGLDRRAEESDFEKLTYLKCALKETLR LHPPILLHETAEDATVGGYLVPPKKARVMIN
Mtru	350	VVGLSRQVEESDFEKLTYLKCALKETLR LHPPILLHETAETAEATVNGYFIPKQARVMIN
Atha	352	VVGLDRRVEESDIEKLTYLKTTLKETLRMHPPILLHETAEDTSIDGFFIPKKS RVMIN
Lfil	420	VWAIGRDKNCWEEPDSFKPSRFLKPGVPDFKGSNFEFIPFGSGRRSCPGMQLGLYALDLA
Ljap	420	VWAIGRDKNCWEEPNSFKPSRFLKPGVNFKGSNFEFIPFGSGRRSCPGMQLGLYALDLA
Gmax	375	AWAIGRDKNSWEEPESFKPARFLKPGVPDFKGSNFEFIPFGSGRRSCPGMVGLGLYALELT
Mtru	410	AWAIGRDANCWEEPQSFKPSRFLKPGVPDFKGSNFEFIPFGSGRRSCPGMQLGLYALDLA
Atha	412	AFAIGRDPTS WTDPDTFRPSRFLKPGVPDFKGSNFEFIPFGSGRRSCPGMQLGLYALDLA

```

Lfil  480 VAHLLHCFTWELPDGMKPSEMDMSDVFGLTAPRASRLVAIPTKRVICPLS*
Ljap  480 VAHLLHCFTWELPDGMKPSEMDMSDVFGLTAPRASRLVAIPTKRVCPLS*
Gmax  435 VAHLLHCFTWELPDGMKPSEMDMGDVFGLTAPRSTRLLAVPTKRVCPLF*
Mtru  470 VAHLLHCFTWELPDGMKPSEMDMSDVFGLTAPRASRLVAIPTKRVLCPLD*
Atha  472 VAHLLHCFTWKLDPGMKPSELDMNDVFGLTAPKATRLFAVPTTRLICAL-*

```

Figure S33. Amino acid alignment for POD-2

Lfil	1	-----
Ljap	1	MANSMSLLM-LLSLLAFAPLCLCHKK--MGGYLYPQFYDYSCPQAKEIVKSIIVANAVARE
Gmax	1	MANSMSFFL-LLSLLAFAPLCLCHYN--QEGYLYPQFYDYSCPQVQHIVKSVLAKYVAEQ
Mtru	1	MANSLSFLM-LLSLLAFAPFCLCHKK--MGSYLYPQFYDYSCPQAQNIIVKSILANAVAKE
Atha	1	MAKSLNIIIAALSLIAFSPFCLCSKAYGSGGYLFPQFYDQSCPKAQEIVQSIVAKAFEHD

Lfil	1	-----K
Ljap	58	TRTAASIIRLHFHDCFVKGCDAIILLDGSSTSEKVSNNPNRNSARGFEVIDEIKQALEK
Gmax	58	PRIAASIIRLHFHDCFVKGCDAIILLDSVNIISEKGSNNPNRNSARGFEVDAIKAELEK
Mtru	58	PRIAASIIRLHFHDCFVKGCDAIILLDNSGSIISEKGSNNPNRNSARGFEVIDEIKYALEK
Atha	61	PRMPASIIRLHFHDCFVKGCDAIILLDSGTIISEKRSNNPNRNSARGFELIEEIKHALEK

Lfil	2	ECPQVVSCADILALAARDSTVLTTGGPSWEVPLGRRDSRSASLSGSNNNIPAPNNTFQTIL
Ljap	118	ECPQVVSCADILALAARDSTVLTTGGPSWEVPLGRRDSRSASLSGSNNNIPAPNNTFQTIL
Gmax	118	KCPSTVSCADILTLAARDSVVLTTGGPSWEVPLGRRDSLGAISGSNNNIPAPNNTFQTIL
Mtru	118	ECPHTVSCADILATAARDSVTLGGPNWEVPLGRRDSLGAISGSNNNIPAPNNTFQTIL
Atha	121	ECPETVSCADILALAARDSTVITGGPSWEVPLGRRDARGASLSGSNNNDIPAPNNTFQTIL

Lfil	62	TKFKLQGLNIVDLVALSGSHT-----
Ljap	178	TKFKLQGLNIVDLVALSGSHTIGNSRCASFRQRLYNQITGNGKADFTLDPNYAAELRTQCP
Gmax	178	TKFNLQGLDIVDLVALSGGHTIGNARCTTFKQRLYNQSGNGEPDSTLDQYYAATLRNRCP
Mtru	178	TKFKLQGLDIVDLVALSGSHTIGKSRCTSFQRRLYNQITGNGKQDFTLDQYYAELRTQCP
Atha	181	TKFKRQGLDIVDLVSLSGSHTIGNSRCTSFQRRLYNQSGNGKPDMTLSQYYATLLRQRCP

Lfil	82	-----
Ljap	238	KSGGDQNLVFLDPVTPTKFDNTYFKNLLAYKGLLSDEILLTANQESAQLVKLYAERNDL
Gmax	238	SSGGDQNLFFLDYATPYKFDNSYFTNLLAYKGLLSDDQVLFTMNQESAELVKLYAERNDI
Mtru	238	RSGGDQNLFFLDYVTPTKFDNNYFKNLLAYKGLLSDEILLTKNQESAELVKLYAERNDL
Atha	241	RSGGDQTLFFLDFATPFKFDNHYFKNLIIMYKGLLSDEILFTKNKQSKELVELYAENQEA

Lfil	82	-----
Ljap	298	FFEQFAKSMIMMGNISPLTGSKGEIRKNCRVIN---*
Gmax	298	FFEQFAKSMIKMGNISPLTNSKGEIRENCRRINA--*
Mtru	298	FFEQFAKSMIKMGNISPLTGSRGNIIRTNCRVINTW-*
Atha	301	FFEQFAKSMVKMGNISPLTGAKGEIRRICRVNHAY*

Figure S34. Amino acid alignment for POD-1

Lfil	1	MVNSIKCFL-LLS-LAFVPFCHCKKEQ--GVNLYPQFYDQSCPKAQEIVKSIVAKAFKAD
Ljap	1	MVNSIKCLL-LLS-LAFVPFCHCKKEQ--GVNLYPQFYDQSCPKAQEIVKSIVAKAFKAD
Gmax	1	MVKSINFL-LLSLLAFVPSCHCKKKI--GGYLYPQFYDQSCPKAQEIVQSIVAKAVAKE
Mtru	1	MVNSINFL-LLSLLVFAPCCHCKTKL--GDYLYPQFYDESCPKVEEIVKSVVAKAVTKE
Atha	1	MAKSLNILLTAALSLTAFSPFCLCSKAYSGGYLFPQFYDQSCPKAQEIVQSIVAKAFEHD
Lfil	57	PRMAASLLRLHFHDCFVKGCDGSVLLDSSGTIISEKRSNPNRNSARGFEVIDEIKSAIEK
Ljap	57	PRMAASLLRLHFHDCFVKGCDGSVLLDSSGTIISEKRSNPNRNSARGFEVIDEIKSAIEK
Gmax	58	PRMAASLLRLHFHDCFVKGCDAVLLDSSGTIISEKRSNPNRD SARGFEVIDEIKSAIEK
Mtru	58	PRMAASLLRLHFHDCFVKGCDAVLLDSSGTIISEKRSNPNRNSARGFEVIEEIKSAVEK
Atha	61	PRMPASLLRLHFHDCFVKGCDAVLLDSSGTIISEKRSNPNRNSARGFELIEEIKHAEHQ
Lfil	117	ECPTHVSCADILATIAARDSTVLTGGPSWGVPLGRRDSLGLASLGSNQNIAPANNTFQTIL
Ljap	117	ECPTHVSCADILATIAARDSTVLTGGPSWGVPLGRRDSLGLASLGSNQNIAPANNTFQTIL
Gmax	118	ECPTHVSCADILALAARDSTVLTGGPSWGVPLGRRDSLGLASTSGSNNNIPANNTFQTIL
Mtru	118	ECPTVSCADILTIAARDSTVLTGGPSWDVPLGRRDSLGLASTSGSNNNIPANNTFQTIL
Atha	121	ECPETVSCADILALAARDSTVITGGPSWEVPLGRRDARGASLGSNNNDIPANNTFQTIL
Lfil	177	TKFKLKGLDIVDLVALSGSHTIGDSRCTSFQRQLYNQTGNGKADFTLEQSYAAKLRTQCP
Ljap	177	TKFKLKGLDIVDLVALSGSHTIGDSRCTSFQRQLYNQTGNGKADFTLDQSYAAKLRTQCP
Gmax	178	TKFKLKGLDIVDLVALSGSHTIGNSRCTSFQRQLYNQTGNGKADFTLDQVYAAELRTRCP
Mtru	178	TKFKLKGLNIVDLVALSGSHTIGDSRCTSFQRQLYNQTGNGKSDFTLDQNYAAQLRTRCP
Atha	181	TKFKRQGLDIVDLVSLSGSHTIGNSRCTSFQRQLYNQSGNGKPDMTLSQYYATLLRQRCP
Lfil	237	RSGGDQNLFVLDFVTPVKFDNNYFKNLLAFKGLLSDEILLTDNKE SADLVKKYAERNDL
Ljap	237	RSGGDQNLFVLDFVTPVKFDNNYFKNLLAFKGLLSDEILLTDNKE SADLVKKYAERNDL
Gmax	238	RSGGDQNLFVLDFVTPVKFDNFFYKNLLANKGLLSDEILLTKNQVSADLVKQYAENNDL
Mtru	238	RSGGDQNLFVLDFVTPVKFDNNYFKNLLANKGLLSDEILLTKNQVSADLVKKYAESNDL
Atha	241	RSGGDQTLFFLDFATPFKFDNHYFKNLIMYKGLLSDEILFTKNKQSKELVELYAEHQEA
Lfil	297	FFEQFAKSMVKMGNITPLTGSERGEIRRICRRINN--*
Ljap	297	FFEQFAKSMVKMGNITPLTGSERGEIRRICRRINN--*
Gmax	298	FFEQFAKSMVKMGNITPLTGSERGEIRKNCRGINK--*
Mtru	298	FFEQFAKSMVKMGNITPLTGSERGEIRKRCKRINN--*
Atha	301	FFEQFAKSMVKMGNISPLTGAKGEIRRICRRVNHAY*

Appendix 10

Supplementary Figures S35 Amino acid alignments for red-differentially expressed lignin genes in *Lotus filicaulis*

Figure S35. Amino acid alignment for CAD

Lfil	1	MSNE-EGKIVCVGTGASGYIASWIVKFLLEHGYTVRATVRDLSPNPKVEHLVKLDVAKERL
Ljap	1	MSNE-EGKIVCVGTGASGYIASWIVKFLLEHGYTVRATVRDLSPNPKVEHLVKLDVAKERL
Gmax	1	MSNNAGKVVVGTGASGYIASWIVKFLLRGYTVRATVRYPNLLKKVDHLVKLEGAKEERL
Mtru	1	MMSG-EGKVVCVGTGANFYIASWIVKFLLRQGYTVRATVRHPSNSEKVDHLVKLDGAKERL
Atha	1	MSSE-EEKTVCVGTGASGYIASWIVKLLLRGYTVKASVRDPNDPRKTEHLLEALEGAEEERL
Lfil	60	QLFKADLLEEGSFDSVIQGCHGVFHVASPVLMFVEDPQAEIDPAVKGTNLVLKSCAKSP
Ljap	60	QLFKADLLEEGSFDSVIQGCHGVFHVASPVLMFVEDPQAEIDPAVKGTNLVLKSCAKSP
Gmax	61	QLFKADLLEEGSFDSVVEGCHGVFHTASPVRFVNDPQAEIDPAVKGTNLVLKSCAKSP
Mtru	60	QLFKADLLEEGSFDSVVEGCDGVFHTASPVRFVNDPQAEIDPAIKGTNLVLQSCAKSP
Atha	60	KLFKANLLEEGSFDSAIIDGCEGVFHTASPFYHDKDPQAEIDPAVKGTINVLSSCLKTS
Lfil	120	SVKRVVLTSSTSAVLFNGRPKSPPEVVDETWFSDPDFLRESKFFHKWY-----
Ljap	120	SVKRVVLTSSTSAVLFNGRPKSPPEVVDETWFSDPDFLRESKV---CV-----
Gmax	121	SVKRVVLTSSISAVAFNRRPKTPQVVDETWFSDPDVCRELEL---WYTLSKTLAEDAAN
Mtru	120	SVKRVILTSSISAVVFDTRPKNPGVIVDETWFSPNPDLCRESKL---WYTLSKTLAEAAAN
Atha	120	SVKRVVLTSSIAAVAFNGMPRTPETIVDETWFADPDYCRASKL---WYVLSKTLAENAAAN
Lfil	168	-----TLSKTMA--EDAA--
Ljap	165	-----PFPNLSIGWVDVKDVAKA
Gmax	178	KFVNENSIDMISINPTMVAGPLLQPEINESVEPIILNLLINGK-PFPNKSFGWVDVKDVANA
Mtru	177	KFVNENSIDMVAINPTMVAGPLLQPELNGSVEPIILNLLISGI-PFPNKAYGWCNVKDVANA
Atha	177	KFAKENNLQLVSINPAMVIGPLLQPTLNTSAAAVLSLIKGAQTFPNATFGWVNVKDVANA
Lfil	179	-----WKFAK-
Ljap	183	HIQAYEITSASGRYCLVERVAHFSEILLEFYMICIQHCKFQISVWTMNHMSQH--FNFHKK
Gmax	237	HILAYEIASASGRYCLVERVIHYSELATILR--GLYPTLQIPDKCEVDEPYIPTYQISTE
Mtru	236	HILAYETASASGRYCLAERVVHYSELAMILR--DLYPTLQISDKCEDDGPYMPPTYQISKE
Atha	237	HIQAFENPDADGRYCLVERVAHYSEVVNILH--DLYPDFQLPEKCADEKIYIPTYKVSKE
Lfil	184	-EN-KIDMVVILPSMVVGPLLRLPEVNFS-VEPVLNIIINGVPE-P-
Ljap	241	EPR-AXESSLLXKXVF---RKPXKAS-EKRR-----LST--
Gmax	295	KAKKDLGIEFTPLEVSL---RETVESFREKKI-----VNENP*
Mtru	294	KAK-SLGIEFTSLEVTI---KETVESFREKKI-----VNF--*
Atha	295	KAE-SLGEFVPLEVSI---KETVESLRDKGF-----IRF--*

Appendix 11

Supplementary Figures S36 – S40 Amino acid alignments for yellow-differentially expressed lignin genes in *Lotus filicaulis*

Figure S36. Amino acid alignment for C3'H

Lfil	1	MSLLLAIPISLITLILSYTLYQ---RLKFKLPPGPRPWPVVGNYDIKPVRFRCFAEWAQ
Ljap	1	MSLLLAIPISLITLILSYTLYQ---RLKFKLPPGPRPWPVVGNYDIKPVRFRCFAEWAQ
Gmax	1	MALLLIVPISLVTLWLGYTLYQ---RLRFKLPPGPRPWPVVGNYDIKPVRFRCFAEWAQ
Mtru	1	MSPLIIVFATIAAAIILYRILNLITKPSLPLPPGPSWPPIVGNLPHMGVPPHHAVALAL
Atha	1	MSWFLIA-VATIAAVVSYKLIQ---RLRYKFPFGPSPKPIVGNLYDIKPVRFRCYYEWAQ
Lfil	58	SYGPIISVWFGSTLNVVVSNSELAJEVLKEKDQQLADHRHSRSAAKFSRDGDLIWADYG
Ljap	58	SYGPIISVWFGSTLNVVVSNSELAJEVLKEKDQQLADHRHSRSAAKFSRDGDLIWADYG
Gmax	58	SYGPIISVWFGSTLNVIVNSELAJEVLKEHDQQLADHRHSRSAAKFSRDGKDLIWADYG
Mtru	61	KHGPIIMHRLRGYVDVIVAASASVAEQFLKVHDTNFSRPTNSGAKYLAYNYQDLVFAPYG
Atha	57	SYGPIISVWIGSILNVVVS SAELAJEVLKEHDQKLADHRNRSTEAFSRNGQDLIWADYG
Lfil	118	PHYVKVRKVCTLELFSPKRLEALRPIREDEVTAMVESVFHDCTNS ENLGKGLLVKKYLGT
Ljap	118	PHYVKVRKVCTLELFSPKRLEALRPIREDEVTAMVESVFHDCTNS ENLGKGLLVKKYLGT
Gmax	118	PHYVKVRKVCTLELFTPKRLES LRPIREDEVTIMVESVYNHCTTTGNLGKAILVRKHLGS
Mtru	121	PRWRLLRKISYVHMFSSKALDEFRRHRIQEEVARLIRNLASSGSKAVNLGQ-----MLNV
Atha	117	PHYVKVRKVCTLELFTPKRLES LRPIREDEVTAMVESVFRDCNLPENRAKGLQLRKYLGA
Lfil	178	VAFNNITRLAFGKRFVNSE---GVIDEQGVEFKAIVANGLKLGASLAMA EHIPWLRWMFP
Ljap	178	VAFNNITRQAFGKRFVNSE---GVIDEQGVEFKAIVANGLKLGASLAMA EHIPWLRWMFP
Gmax	178	VAFNNITRLAFGKRFVNSE---GVMDEQGVEFKAIVENGLKLGASLAMA EHIPWLRWMFP
Mtru	175	CTTNALARVMI GRVFNEGGCECDPRADFEKSMVVVELMVLAGVFNIGDFLP AFEWLDL
Atha	177	VAFNNITRLAFGKRFMNAE---GVVDEQGLEFKAIVSNGLKLGASLSIAEHIPWLRWMFP
Lfil	235	L-EEEF AFAKHGARRDRLTRAIMEEHTQARQKSGGAKQHFDALLTLQDKYDLS EDT----
Ljap	235	L-EEEF AFAKHGARRDRLTRAIMEEHTQARQKSGGAKQHFDALLTLQDKYDLS EDT----
Gmax	235	L-EEG AFAKHGARRDRLTRAIMEEHTQARQKSGGAKQHFDALLTLQDKYDLS EDT----
Mtru	235	QGVQGMKKLHKRFDTFLTSTIEDHRISKSE---KHNDLITSLLSLKEKLPED EDTLNDT
Atha	234	A-DEK AFAEHGARRDRLTRAIMEEHTLARQKSSGAKQHFDALLTLKDQYDLS EDT----
Lfil	290	-IIGLLWDMITAGMDTTAISVEWAMAELIKNPRVQQKAQEELDRVIGFERVLTETDFS NL
Ljap	290	-IIGLLWDMITAGMDTTAISVEWAMAELIKNPRVQQKAQEELDRVIGFERVLTETDFS NL
Gmax	290	-IIGLLWDMITAGMDTTAISVEWAMAELIRNPRVQQKVEELDRVIGLERVMTEADFS SL
Mtru	292	EIKALLLNIFTAGTDTTASTTEWATSELIKNPKLMACIQNELDTVVGRDRLVTEQD LTNL
Atha	289	-IIGLLWDMITAGMDTTAITAEWAMAEMIKNPRVQQKVEEFDRVVGLDRILTEADFS RL
Lfil	349	PYLQCVAKEAMRLHPPTPLMLPHRSNANVKIGGYDIPKGSNVHVN VWAVARDPAVWKNPL
Ljap	349	PYLQCVAKEAMRLHPPTPLMLPHRSNANVKIGGYDIPKGSNVHVN VWAVARDPAVWKNPL
Gmax	349	PYLQCVIKEAMRLHPPTPLMLPHRANANVKVGGYDIPKGSNVHVN VWAVARDPAVWKDPL
Mtru	352	PYLEAVVKETLRLHPSTPLSLPRVATQSCEIFNYHIPKGATLLVNVWALSRDPKEWSNPL

Atha 348 PYLQCVVKE¹SFRLHPPTPLMLPHRSNAD²VKIGGYDIPKGSNVHVN³VWAVARDPAVWKNPF

Lfil 409 EFRPERFLE¹---EDVDMKGHDFRLLPFGAGRRVCPGAQLGINLVTSMLGHLLHHFCWAP²

Ljap 409 EFRPERFLE¹---EDVDMKGHDFRLLPFGAGRRVCPGAQLGINLVTSMLGHLLHHFCWAP²

Gmax 409 EFRPERFLE¹---EDVDMKGHDFRLLPFGAGRRVCPGAQLGINLVTSMLGHLLHHFCWTP²

Mtru 412 EF¹KPERFLPGGEKF²DVD³VRGN⁴DFEVI⁵PFGAGRR⁶ICAGMS⁷LG⁸LRMV⁹QLLTATLAHAYDWEL¹⁰

Atha 408 EFRPERFLE¹---EDVDMKGHDFRLLPFGAGRRVCPGAQLGINLVTSMM²SHLLHHFVWTP³

Lfil 465 PEGVKPEEIDMVENPGLVTYMRTPVQAVATPRLPSHLYKRVADI¹*

Ljap 465 PEGVKPEEIDMVENPGLVTYMRTPVQAVATPRLPSHLYKRVADI¹*

Gmax 465 PEG¹MKPEEIDMGENPGLVTYMRTP²IQAIA³SPRLPSHLYKRVPAEI⁴*

Mtru 472 ENGVSPEKINMDEAYGLTLQRAVPILAHPRPRLSPHLYL-----*

Atha 464 PQGT¹KPEEIDMSEN²PGLVTYMRTPVQAVATPRLPS³DLYKRVPYDM⁴*

Figure S37. Amino acid alignment for CAD

Lfil	1	MGSLE--AERTTVGWAARDPSGILSPYTF	TLRNTGPDDVYIKVHYCGVCHTDVHQVKNDLG
Ljap	1	MGSLE--AERTTVGWAARDPSGILSPYTF	TLRNTGPDDVYIKVHYCGVCHTDVHQVKNDLG
Gmax	1	MGSLE--AERTTVGLAARDPSGILSPYTYN	LRNTGPDDVYIKVHYCGICHSDLHQIKNDLG
Mtru	1	MGSIEVAERTTVGLAAKDPGILTPYTYTLRNTGPDDVYIKI	HYCGVCHSSDLHQIKNDLG
Atha	1	MGIME--AERKTTGWAARDPSGILSPYTYTLRE	TGPELVNIRIICCGICHTDLHQTKNDLG
Lfil	60	MSNYPMVPGHEVVGEVLEVGSDVTRFTVGEIVGAGLLVGCCCKKCTACQSDIEQYCKKKIW	
Ljap	60	MSNYPMVPGHEVVGEVLEVGSDVTRFTVGEIVGAGLLVGCCCKKCTACQSDIEQYCKKKIW	
Gmax	60	MSNYPMVPGHEVVGEVLEVGSDVSRFRVGEIVGVGLLVGCCCKNCQPCQQDIENYCSKKIW	
Mtru	61	MSNYPMVPGHEVVGEVLEVGSNVTRFKVGEIVGVGLLVGCCCKSCRACDSIEQYCNKKIW	
Atha	60	MSNYPMVPGHEVVGEVLEVGSDVSKFTVGDIVGVGCLVGCCGGCSPCERDLEQYCPKKIW	
Lfil	120	NYNDVYVDGKPTQGGFAETIIVVEQKFVVKIPEGMAPEQVAPLLCAA	VTVYSPLSHFGLKE
Ljap	120	NYNDVYVDGKPTQGGFAETIIVVEQKFVVKIPEGMAPEQVAPLLCAA	VTVYSPLSHFGLKE
Gmax	120	SYNDVYVDGKPTQGGFAETMVVEQKFVVKIPEGIAPEQVAPLLCAGVTVYSPLVHFGLKE	
Mtru	121	SYNDVYTQGGFAETIIVVEQKFVVKIPEGIAPEQVAPLLCAGVTVYSPLSHFGLKT	
Atha	120	SYNDVYINGOPTQGGFAKATVNVHOKFVVKIPEGMAVEQAAPLLCAGVTVYSPLSHFGLKQ	
Lfil	180	SGLRGGILGLGGVGHMGVITAKAMGHVTVISSDRKKKEAIEDLGADAYLVSSDTTSMQ	
Ljap	180	SGLRGGILGLGGVGHMGVITAKAMGHVTVISSDRKKKEAIEDLGADAYLVSSDTTSMQ	
Gmax	180	SGLRGGILGLGGVGHMGVKIAKALGHVTVISSDCKKKEALEHLGADQYLVSSDVTAMQ	
Mtru	181	PGLRGGILGLGGVGHMGVIAKAFGHVTVISSDCKKKEALEDLGADSYLVSSDVTVMQ	
Atha	180	PGLRGGILGLGGVGHMGVKIAKAMGHVTVISSNKKKEEALQDLGADDYVIGSDQAKMS	
Lfil	240	EAADSLDYIIDTVPVGHLEPYLSLLKLDGKLILMGVINTPLQFITPMVMLGRRSITGSF	
Ljap	240	EAADSLDYIIDTVPVGHLEPYLSLLKLDGKLILMGVINTPLQFITPMVMLGRRSITGSF	
Gmax	240	EAADSLDYIIDTVPVGHLEPYLSLLKLDGKLILMGVINTPLQFVSPMVMLGRRSITGSF	
Mtru	241	EAADSLDYIIDTVPVGHLEPYLSLLKLDGKLILMGVINTPLQFVTPMVMLGRKSITGSF	
Atha	240	EIADSLDYVIDTVPVHHALEPYLSLLKLDGKLILMGVINNPLOFLTPLMLGKVKITGSF	
Lfil	300	IGSIKETEEMLFEWKEKGLTSMIEIVNMDYINKAFERLEKNDVRYRFVVDVKGSKLIDQ*	
Ljap	300	IGSIKETEEMLFEWKEKGLTSMIEIVNMDYINKAFERLEKNDVRYRFVVDVKGSKLIDQ*	
Gmax	300	IGSMKETEEMLFEWKEKGLSSMIEIVNMDYINKAFERLEKNDVRYRFVVDVKGSKLVD*	
Mtru	301	VGSVKETEEMLFEWKEKGLSSMIEIVTMDYINKAFERLEKNDVRYRFVVDVKGSKFED*	
Atha	300	IGSMKETEEMLFCKEKGLSSIIEVVKMDYVNTAFERLEKNDVRYRFVVDVEGSNLDA*	

Figure S38. Amino acid alignment for CCoA-OMT

Lfil	1	-----
Ljap	1	-----
Gmax	1	MTL-----IKEL EQPNQ IA GHKE LA HKS LLQSDALYQYILETSVYPREHESLKE
Mtru	1	MAI-----NNG- EEQQN QLIGDA DLA HKTLLQSDALYQYILDTSVFPREHPC LKE
Atha	1	MATTTTEATKTSSTNGE DQKQS Q NLRHQ EV GHKS LLQSDDLYQYILETSVYPRE PES MKE

Lfil	1	-----
Ljap	1	-----
Gmax	51	LRE LT TEKHPWN LMATPP DE QQL IG MLIK LINAKNTMEIGVFTGYSLL STALALPS SDG---
Mtru	50	LRE MT TEKHPRN HMAIPG DE QQL IS MLV KLINAKNTLEIGVYTGYSLL STALALPS SDG KVL
Atha	61	LRE VT AKHPWN NIMTTSA DE QQL FN MLIK LVNAKNTMEIGVYTGYSLL LATALALPE DG KIL

Lfil	1	-----IL GLPI IEKAGVAHKIDFREGPAL PVLDE LE DE NNKGT FD FV FVDADKDN
Ljap	1	-----S FLWML -----I
Gmax	108	-----KAGVAHKIDFREGPAL PLLDQ L IK DEKNKGAFDFIY VVDADKDN
Mtru	110	ALDKSREYYEL GLPM IEKAGVAHKIDF QGL AL PLLD ELLQDEN NKGT FD FV FVDADK NN
Atha	121	AMDVNRENYEL GLPI IEKAGVAHKIDFREGPAL PVLDE IV AD EN NHGT YDFI FVDADKDN

Lfil	52	YLN YH KRVIELVKVGGLIG FD N TL WNGSVV SAP DAP MDY LKH YR GYVIELNKY LAQDSR
Ljap	8	RM NYH KRVIELVKVGGLIG FD N TL WNGSVV SAP DAP MDY LKH YR AYVIELNKY LAQDSR
Gmax	151	YLN YH KRVIELVKVGGL V GYD N TLWNGSVV AP PDAP MDY VKY YR DFV ME LNKAL ALDSR
Mtru	170	YLN YH V RM IDLVKVGGLIGYD N TL WS GS V A AP PDAP MD FIR N LRGYVIE FN KKH LAQDSR
Atha	181	Y IN YHKRLIDLVK I GGVIGYD N TLWNGSVV AP PDAP MR KYV RYR DFV LE LNKAL AA D PR

Lfil	112	IEICQLPVGDGITLCRR I *
Ljap	68	IEICQLPVGDGITLCRR I -
Gmax	211	VEICQLPVGDGITLCRR I *
Mtru	230	IEICQL S IADGITLCRR I K*
Atha	241	IEIC M LPVGDGIT I CR R IS*

Figure S39. Amino acid alignment for CCR-2

Lfil	1	MPAAAAES---PSPTGETICVTGAGGFASWMVKLLLQKGYTVRGTVRNPDDPKNGHLR
Ljap	1	MPAA-----AGETICVTGAGGFASWMVKLLLQKGYTVRGTVRNPDDPKNGHLR
Gmax	1	MPSAEA-----SSETICVTGAGGFASWMVKLLLEKGYTVRGTLRNPDDPKNGHLK
Mtru	1	MPAATAAAAAESSVSGETICVTGAGGFASWMVKLLLEKGYTVRGTLRNPDDPKNGHLK
Atha	1	MPVDVA-----SPAGKTVCVTGAGGYIASWIVKILLERGYTVKGTVRNPDDPKNTHLR

Lfil	58	ELEGASDRLTLIKVDLLDLNSVRAAVHGS
Ljap	50	ELEGASDRLTLIKVDLLELNSVRAAVHGS
Gmax	52	EEFEGASERLTLHKVDLLHLDVSRSVINGCHGVFHTASPVTDNPEEMVEPAVNGAKNVI IA
Mtru	61	KLEGAKERLTLVKVDLLDLNSVKEAVNGCHGVFHTASPVTDNPEEMVEPAVNGAKNVI IA
Atha	54	ELEGCKERLILCKADLQDYEAALKAIDGCDGVFHTASPVTDDPEQMVPEPAVNGAKFVINA

Lfil	118	AAEAKVRRVFTSSIGAVYMDPKRSVDSVVDSCWSDLEFCKDTKNWYCYGKAVAEQAAW
Ljap	110	AAEAKVRRVFTSSIGAVYMDPKRSVDSVVDSCWSDLEFCKDTK-----
Gmax	112	AAEAKVRRVFTSSIGAVYMDPSRSIDLVDSCWSDLEYCKNTKNWYCYGKAVAEQAAW
Mtru	121	GAEAKVRRVFTSSIGAVYMDPNRSVDVEVDSCWSDLEFCKTKTNWYCYGKAVAEQAAW
Atha	114	AAEAKVKRVITSSIGAVYMDPNRDPEAVVDSCWSDLD FCKNTKNWYCYGKMVAEQAAW

Lfil	178	DTAKEKGVDLVVVNPVLVLGPLLQPTINASTVHILKYLTGSAKTYANATQAYADVVRDVAL
Ljap	155	-----ANATQAYADVVRDVAL
Gmax	172	DTAKENGVDLVVVNPVLVLGPLLQPTINASTIHILKYLTGSAKTYANATQAYVHVRDVAL
Mtru	181	DVAKEKGVDLVVVNPVLVLGPLLQPTINASTIHILKYLTGSAKTYANATQAYVHVRDVAL
Atha	174	ETAKEKGVDLVVLNPVLVLGPPLQPTINASLYHVLKYLTGSAKTYANLTQAYVDVRDVAL

Lfil	238	AHILVYEKPEASGRYICSESSLHRGELVEILAKHFPEYPMPTKCKDEKNPRAKPYIFSNO
Ljap	170	AHILVYERPEASGRYICSESSLHRGELVEILAKHFPEYPMPTKCKDEKNPRAKPYIFSNO
Gmax	232	AHILVYEKPSASGRYLCAESSLHRGELVEILAKYFPEYPMPTKCSDEKNPRAKPYTFSNO
Mtru	241	AHILVYEKPSASGRYLCAETSLHRGELVEILAKYFPEYPIPTKCSDEKNPRVKPHIFSNO
Atha	234	AHVLVYEAPSASGRYLIAESARHRGEVVEILAKLFPEYPLPTKCKDEKNPRAKPYKFTNO

Lfil	298	KLKDLGLEFTPVSHSLYETVKSLQEKGHLT--IPK---QEDSLAVKS*
Ljap	230	KLKDLGLEFTPVSHSLYETVKSLQEKGHLT--IPK---QEDSLAVKS*
Gmax	292	KLKDLGLEFTPVSQCLYETVKSLQEKGHLT--VPAKQEEEDSTTVKS*
Mtru	301	KLKDLGLEFTPVSECLYETVKSLQDQGHLS--IPN---KEDSLAVKS*
Atha	294	KIKDLGLEFTSTKQSLYDTHRKNPWKMALRSGLEK---LINSISKYPP*

Figure S40. Amino acid alignment for CCR-1

Lfil	1	-----
Ljap	1	-----
Gmax	1	MPT--DTSSVSG--EIVCVTGAGGFIASWLVKLLLEKGYTVRGTVRNPDDPKNGHLKELE
Mtru	1	MPAYDNTSSVSGGDQTVCVTGAGGFIASWLVKLLLERGYTVRGTVRNPEDPKNGHLKELE
Atha	1	MPV--DVASPAK--KTVCVTGAGGYIASWIVKILLERGYTVKGTVRNPDDPKNTHLRELE

Lfil	1	-----
Ljap	1	-----MIEPAVNGARNVVMAAAE
Gmax	57	GGKERLTLHKVDLFDIASIKAAIHGCHGVFHTASPVTDNPEEMVEPAVKGTKNVIAAAE
Mtru	61	GARERLTLHKVDLLDLQSIQSVVHGCHGVFHTASPVTDNPEEMLEPAVNGTKNVIIASAE
Atha	57	GGKERLILCKADLQDYEAALKAAIDGCDGVFHTASPVTDDPEQMVEPAVNGAKFVINAAAE

Lfil	1	-----
Ljap	19	TKVQRVVFTSSIGTVYMDPNRSRDEVVDESCWSDLEYCKNTKNWYCYGKTVAEQVAWEAA
Gmax	117	AKVRRVVFTSSIGTVYMDPNTSRDAIVDESFWSDLEYCKNTKNWYCYGKTVAEQAAWDVA
Mtru	121	AKVRRVVFTSSIGTVYMDPNTSRDVVDESFWSDLEHCKNTKNWYCYGKTVAEQSAWDIA
Atha	117	AKVKRVVITSSIGAVYMDPNRDEAVVDESCWSDLDFCKNTKNWYCYGKMVAEQAAWETA

Lfil	1	-----VDVKDAALAHV
Ljap	79	KEKGVDLVVVNPVLVIGPLLQPTINASTIHILKYLTGSAKTYVNATQAYVDVKDAALAHV
Gmax	177	KERGVDLVVVNPFVLVIGPLLQPTINASTIHILKYLTGSAKTYVNATQAYVHVVDVALAHI
Mtru	181	KENQVDLVVVNPVVLGPLLQPTINASTIHILKYLNGAAKTYVNATQSYVHVVDVALAHL
Atha	177	KEKGVDLVVLNPVLVLGPPQLQPTINASLYHVLKYLTGSAKTYANLTQAYVDVRDVALAHV

Lfil	12	LVEYTPSASGRYICAESSLHRGELVEILAKYFPEYPIPTKCSDEKNPRVKPYIFSNQKLK
Ljap	139	LVEYTPSASGRYICAESSLHRGELVEILAKYFPEYPIPT-----
Gmax	237	LVEYTPSASGRFICAESSLHRGELVEILAKFFPEYPIPTKCSDEKNPRVKPYIFSNQKLK
Mtru	241	LVEYTNASGRYICETALHRGEVVEILAKYFPEYPLPTKCSDEKNPRVKPYKFSNQKLK
Atha	237	LVYEAPSASGRYLLAESARHRGEVVEILAKLFFPEYPLPTKCKDEKNPRAPKYKFTNQKIK

Lfil	72	DLGLEFTPV-----
Ljap		-----
Gmax	297	DLGLEFTPVKQCLYDTVKNLQENGHLVPPKQKDS-----Y*
Mtru	301	DLGLEFTPVKQCLYDTVRSLQEKGHLPIPPMQEDS-----A*
Atha	297	DLGLEFTSTKQSLYDTVKSLQEKGHLAPPPPPPSASQESVENGIGKS*

Appendix 12

Supplementary Figures S41 – S47 Amino acid alignments for yellow-differentially expressed carotenoid genes in *Lotus sessilifolius*

Figure S41. Amino acid alignment for CRTISO

Lses	1	VKGARVLVLEKYVIPGGSSGFYQRDGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCOMQ
Ljap	1	VKGARVLVLEKYVIPGGSSGFYQRDGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCKMQ
Gmax	1	VKGARVLVLEKYVIPGGSSGFYQRDGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCRMQ
Mtru	1	VKGAKVLVLEKYVIPGGSSGFYHREGYTFDVGSSVMFGFSDKGNLNLITQALEAVGCRMQ
Atha	1	VKEARVLVLEKYVIPGGSSGFYERDGYTFDVGSSVMFGFSDKGNLNLITQALKAVGRKME
Lses	61	VVPDPTTVHFHLPNNLSVRVHREYDKFIEELTCYFPHEKEGILKFYGECKWIFNALNSLE
Ljap	61	VVPDPTTVHFHLPNNLSVRVHREYDKFIEELTCYFPHEKEGILKFYGECKWIFNALNLLE
Gmax	61	VVPDPTTVHFHLPNNLSVRVHREYDKFIEELTSYFPHEKEGILKFYGECKWIFNALNSLE
Mtru	61	VIPDPTTVHFHLPNHLSVRVHREYDKFIEELTSYFPHEKDGILKFYGECKWIFNALNSLE
Atha	61	VIPDPTTVHFHLPNNLSVRTHREYDDFIAELTSKFPHEKEGILGFYDCWKIFNSLSNSLE
Lses	121	LKSLEEPLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Ljap	121	LKSLEEPLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Gmax	121	LKSLEEPLYLFGQFFRRPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Mtru	121	LKSLEEPLYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIQDPQLLSFIDAECFIVSTVN
Atha	121	LKSLEEPITYLFGQFFQKPLECLTLAYYLPQNAGAIARKYIKDPQLLSFIDAECFIVSTVN
Lses	181	ALQTPMINAAMVLCDRHFGGINYPGGVGGIKSLAKGLADQGSEILYKANVTSVIIIEQG
Ljap	181	ALQTPMINAAMVLCDRHFGGINYPGGVGGIKSLAKGLADQGSEILYKANVTSVIIIEQG
Gmax	181	ALQTPMINAAMVLCDRHFGGINYPGGVGGIKSLAKGLVDQGSEIVYKANVTSIIIEQG
Mtru	181	ALQTPMINASMVLCDRHFGGINYPGGVGGIKSLAKGLVDQGSQILYKANVTSIIIEQG
Atha	181	ALQTPMINASMVLCDRHYGGINYPGGVGGIKSLAEGLVDQGSEIQYKANVKSIIIDHG
Lses	241	KAVGVRLSDGREFFAKTIIISNATRWDT-----
Ljap	241	KAVGVRLSDGREFFAKTIIISNATRWDTFG-----
Gmax	241	KAVGVRLSNGREFFAKTIIISNATRWDTFGKLLKGVPLPKEEENFQKVYVKAPSFLSIHMG
Mtru	241	KAVGVRLSDGREFFAKTIIISNATRWDTFGKLVKGESLPKEEENFQKVYVKAPSFLSIHMG
Atha	241	KAVGVRLADGREFFAKTIIISNATRWDTFGKLLKGEKLPKEEENFQKVYVKAPSFLSIHMG
Lses		-----
Ljap	270	-----
Gmax	301	VKAEVLPPDTCCHHFVLESNWSKLEEPYGSIFLSIPTVLDSSLAPEGRHILHIFTTSSME
Mtru	301	VKAEVLPPDTCCHHFVLENNWTSLEEPYGSIFLSIPTILDSSLAPEGRHILHIFTTSSMD
Atha	301	VKAEVLPPDTCCHHFVLEDDWKNLEEPYGSIFLSIPTILDSSLAPDGRHILHIFTTSSIE
Lses		-----
Ljap	270	-----
Gmax	361	DWEGLSRVEYEAKKQLVADEITSRLENKLFPGLRSSIDFMEVGTPKTHRRYLARDEGTYG
Mtru	361	DWQGLSRIDYEAKKQVVADEIISRLEKLFPGLRSSIDFIEVGTPKTHRRYLARDEGTYG

```
Atha 361 DWEGLPKKEYEAKKEDVAARIIQRLKKLFPGLSSSITFKKEVGTPRTHRRFLARDKGTYG

Lses -----
Ljap 270 -----
Gmax 421 PMPRRI PKGLLGMPFNNTG IDGLYCVGDSCFPGQGVIAVAFSGVMCAHRVAADIGLEKKS
Mtru 421 PMPRSI PKGLLGMPFNNTS IDGLYCVGDSCFPGQGVIAVAFSGVMCAHRVAADIGLEKKS
Atha 421 PMPRGT PKGLLGMPFNNTA IDGLYCVGDSCFPGQGVIAVAFSGVMCAHRVAADIGLEKKS

Lses -----
Ljap 270 -----WFR
Gmax 481 PVLD SMLLRLLGWLR
Mtru 481 PVLDGMLLGLLGWLR
Atha 481 RVLDVGLLGLLGWLR
```

Figure S42. Amino acid alignment for CYB

Lses	1	RASSALLELVPEFKKENLDFELPLYDSTKGAVVDLAVVGGGPAGLAVAQQVSEAGLSVC
Ljap	1	RASSALLELVPEVKKENLDFELPLYDSMKGAVVDLAVVGGGPAGLAVAQQVSEAGLSVC
Gmax	1	RASSALLELVPEFKKENLDFELPLYDSSKGAMVDLAVVGGGPAGLAVAQQVSEAGLSVC
Mtru	1	RASSALLELVPEFKKENLDFELPLYDSTKGTVVVDLAVVGGGPAGLAVAQQVSEAGLSVC
Atha	1	VSGSAALLDLVPETKKENLDFELPLYDTSKSQVVDLAVVGGGPAGLAVAQQVSEAGLSVC
Lses	61	AIDPNPRLIWPNNYGVWVDFEAMDLLDCLDTTWSGAVVYIDDKTKKYLD RPYGRVNRKL
Ljap	61	AIDPNPRLIWPNNYGVWVDFEAMDLLDCLDTTWSGAVVYIDDKTKKYLD RPYGRVNRKL
Gmax	61	AIDPNPRLIWPNNYGVWVDFEAMDLLDCLDTIWSGAVVFIDDKTKKDL RPYGRVNRKL
Mtru	61	AIDPNPRLIWPNNYGVWVDFEAMDLLDCLDKTWSGAVVYIDDKTKKDL RPYGRVNRKL
Atha	61	SIDPSPKLIWPNNYGVWVDFEAMDLLDCLDTTWSGAVVYVDEGVKKDLS RPYGRVNRKQ
Lses	121	LKSKMLQKCISNGVKFHQAKVIKVIHEESKSM LICNDGVTIQATVVLDATGFSRSLVQYD
Ljap	121	LKSKMLQKCISNGVKFHQAKVIKVIHEESKSM LICNDGVTIQATVVLDATGFSRSLVQYD
Gmax	121	LKSKMLQKCISNGVKFHQAKVIKVIHEEAKSLICNDGVTVHATVVLDATGFSRCLVQYD
Mtru	121	LKSKMLQKCIANGVKFHQAKVIKVIHEESKSM LICNDGVTVQATVVLDATGFSRCLVQYD
Atha	121	LKSKMLQKCIANGVKFHQSKVTNVVHEEANSTVVCS DGVKIQASVVLDATGFSRCLVQYD
Lses	181	KPYNPGYQVAYGILAEVEEHPFDVDKMLFMDWRD SHLDNDMELKERN SKIPTFLYAMPFS
Ljap	181	KPYNPGYQVAYGILAEVEEHPFDVDKMLFMDWRD SHLDNNMELKEKNSRIPTFLYAMPFS
Gmax	181	KPYNPGYQVAYGILAEVD EHPFDVDKMLFMDWRD SHLDNDMELKQRNSRIPTFLYAMPFS
Mtru	181	KPYNPGYQVAYGILAEVEEHPFDVDKMLFMDWRD SHLDNDLIKERN SKIPTFLYAMPFS
Atha	181	KPYNPGYQVAYGIVAEVDGHPFDVDKMFMDWRDKHLD SYP ELKERN SKIPTFLYAMPFS
Lses	241	STKIFLEETSLVARPGLRMDDIQDRMVARLKHLG IKVKSIEEDEHCVIPMGGLPVLPQR
Ljap	241	STKIFLEETSLVARPGLRMDDIQDRMVARLKHLG IKVKSIEEDEHCVIPMGGLPVLPQR
Gmax	241	STKIFLEETSLVARPGLRMDDIQERMAARLKHLG IRVKSIEEDEHCVIPMGGLPVLPQR
Mtru	241	STKIFLEETSLVARPGLRMDDIQDRMVARLKHLG INVKSI EEDEQCVIPMGGLPVLPQR
Atha	241	SNRI FLEETSLVARPGLRMEDIQERMAARLKHLG INVKRI EEDERCVIPMGGLPVLPQR
Lses	301	VVGIGGTAGMVHPSTGYMVARTLAAAPIVAN SIVQYLGSDRGISGDHL SAQVWKDLWPIE
Ljap	301	VIGIGGTAGMVHPSTGYMVARTLAAAPIVANA I VQYLGSDRGISGDGLSSIVWKDLWPIE
Gmax	301	VVGIGGTAGMVHPSTGYMVARTLAAAPIVAN SIVQCLGSDRGFSGDEISAQVWKDLWPIQ
Mtru	301	VVGIGGTAGMVHPSTGYMVARTLAAAPIVANA I VQYLGSDRGILLGDEV SARVWKDLWPIE
Atha	301	VVGIGGTAGMVHPSTGYMVARTLAAAPIVANA I VRYLGSPNSLRGDQLSAEVWRDLWPIE
Lses	361	RRRQREFFCFGMDILLKLDLAGTKRFFDAFFDLE P RYWHGFLSSRLFLPELLVFGLSLFS
Ljap	361	RRRQREFFCFGMDILLKLDLPGTRRFFDAFFDLE PHYWHGFLSSRLFLPELLVFGLSLFS
Gmax	361	RRRQREFFCFGMDILLKLDLPGTKRFFDAFFDLE PHYWHGFLSSRLFLHELLVFGLSLFS
Mtru	361	RRRQREFFCFGMDVLLKLDLPGTRRFFDAFFN LEPHYWHGFLSSRLYLPELFTFGLSLFS
Atha	361	RRRQREFFCFGMDILLKLDLDATRRFFDAFFDLQ PHYWHGFLSSRLFLPELLVFGLSLFS
Lses	421	YASNTSRIEIMAKGTVP
Ljap	421	YASNTSRIEIMAKGTVP
Gmax	421	YASNTSRVEIMAKGTLP
Mtru	421	YASNTSRLEIMAKGTLP
Atha	421	HASNTSRLEIMTKGTVP

Figure S43. Amino acid alignment for E-CHY

Lses	1	-----
Ljap	1	SLSLPLSKSSINKTKPSSWVSPDWLTSLSRSLTAGKNDDSGIPIASAKLDDVSDLLGGA
Gmax	1	SLSLPLSKSSINKTKPSSWVSPDWLTSLSRSLTAGNDVSGIPVASAKLDDVSDLLGGA
Mtru	1	SLPLSHLSKSSINKTKSTSWVSPDWLTSLSKSLTTSKNDDSNIPIASAKLDDVSDLLGGA

Lses	1	-----
Ljap	61	LFLPLFKWMKEYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYGKGLVAEVSEFLFGSG
Gmax	60	LFLPLFKWMQDYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYAKGLVAEVSEFLFGSG
Mtru	61	LFLPLFKWMNEYGPIYRLAAGPRNFVVVSDPAIAKHVLKNYGKYGKGLVAEVSEFLFGDG

Lses	1	-----KLQPAALNGTAVNMEDKF
Ljap	121	FAIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRAERLVEKLQPDALNGTAVNMEDKF
Gmax	120	FAIAEGPLWTARRRAVPSLHKRYLSVIVDRVFCRAERLVEKLQPDALNGTAVNMEAKF
Mtru	121	FAIAEGPLWTARRRAVPSLHKRYLSIMVDRVFKCAERLVEKLQADAVNGTAVNMEDKF

Lses	19	SQTLTDVIGLSVFNYDFDSLNDSPVIGAVYTALKEAEARSTDLLPYWQGEFLRKINPRQ
Ljap	181	SQTLTDVIGLSVFNYDFDSLNDSPVIGAVYTALKEAEARSTDLLPYWQGEFLRKIIPRQ
Gmax	180	SQTLTDVIGLSVFNYNFDSLNTDSPVIEAVYTALKEAEARSTDLLPYWKFKFLCKIIPRQ
Mtru	181	SQTLTDVIGLSVFNYNFDALNSDSPVIEAVYTALKEAEARSTDLLPYWKIDFLCKIIPRQ

Lses	79	-----
Ljap	241	IKAENAVTIIIRKTVEDLIEKCKEIVESEGERIDGDEYVNDSDPSILRFLLASREEVSSDQ
Gmax	240	IKAEEAVSVIRKTVEDLIEKCKEIVESEGERIDVEEYVNDSDPSILRFLLASREEVSSVQ
Mtru	241	IKAENAVTVIRKTVEDLIEQCKEIVESEGERIDADEYVNDADPSILRFLLASREEVSSVQ

Lses		-----
Ljap	301	LRDDLLSMLVAGHETTGSVLTWTLYLLSKDSSSLAKVQEEVDRVLOGRRPTFEDMKNLKF
Gmax	300	LRDDLLSLVAGHETTGSVLTWTLYLLSKDSSSLAKAQEEVDRVLOGRRPTYEDIKDLKF
Mtru	301	LRDDLLSMLVAGHETTGSVLTWTLYLLSKDSSSLAKAQEEVDRVLOGRRPTYEDMKDLKF

Lses		-----
Ljap	361	LTRCIITESRLYPHPVLIIRRAQVPDELPGAYKVNAGQDIMISVYNIHHSSEVWDRAEEF
Gmax	360	LTRCIITESRLYPHPVLIIRRAQVPDELPGGYKLDAGQDIMISVYNIHRSSEVWDRAEEF
Mtru	361	LNRCIITESRLYPHPVLIIRRSQIPDELPGDYKIDAGQDIMISVYNIHSSKVWDRAEEF

Lses		-----
Ljap	421	LPERFDLDGPMNPNETNTDFRFIPFSGGPRKCVGDQFALLEATVSLAIFLQHMNFELVPDQ
Gmax	420	VPERFDLDGVPNPNETNTDFRFIPFSGGPRKCVGDQFALMEAIIVALAIFLQHMNFELVPDQ
Mtru	421	LPERFDLDGVPNPNETNTDFRFIPFSGGPRKCVGDQFALLEATVALAVFLQHMNFELVPDQ

Lses		-----
Ljap	481	NISMTTGATIHTTNGLYMKLSQRVK
Gmax	480	NISMTTGATIHTTNGLYMKLSRRLK
Mtru	481	NIGMTTGATIHTTNGLYMKLSQRLK

Figure S44. Amino acid alignment for PSY

Lses	1	VLLWVNCGKGNISISLLGLGGRGSRSQRRFRLCSGISFASFSSAVAEPSPRSSEERVYEVVL
Ljap	1	-----LGGRGSRSQRRYRLCSGISFASFSSAVAEPSPRSSEERVYEVVL
Gmax	1	VLLWVSCGKENPISIVGLGGRGGRSQRRFGLCNGLISFASFSPAVADPSRSSEERVYEVVL
Mtru	1	VLLWVNCGKENAISLMGLVGRDVRITQRRFRHCSGISFACFSSAVVEPTRSSEERVYEVVL
Atha	1	AVLWVATSNPDPMNNGLV-RVLESSRLFSPCQNQRLSLVASPSGEIALSSEEKVVNVVL
Lses	61	KQAALVKEQKRNTRKDLNLDIEDDFTNKDLLNAAYDRCEVCAEYAKTFYLGTLMTTEER
Ljap	44	KQAALVKEQKRDTRKRLNLDIEDDFTNKDLLSGAYDRCEVCAEYAKTFYLGTLMTTEER
Gmax	61	KQAALVKEQDKGTAKRALNLDIEGDLTNGDLLSDAYDRCEVCAEYAKTFYLGTLMTQER
Mtru	61	KQAALVKEQKDKIKRGLNLDIEGDFTNGLLSAYDRCDVCAEYAKTFYLGTLMTQER
Atha	60	KQAALVNKQLRSSSYDLQVQDVVLPSSLGELAYDRCEVCAEYAKTFYLGTLMTTPER
Lses	121	RKAIWAIYVWCRRTDELVDGPNSPHITPKALDRWEQRLSDVFEGRPYDMYDAALSDTVTK
Ljap	104	RKAIWAIYVWCRRTDELVDGPNSPHITPKALDRWEQRLSDVFEGRPYDMYDAALSDTVTK
Gmax	121	RKAIWAIYVWCRRTDELVDGPNASHITPKALDRWEQRLSDVFEGRPYDMYDAALSDTVSK
Mtru	121	RKAIWAIYVWCRRTDELVDGPNASHITPKALDRWEQRLTDVFEGRPYDMYDAALSDTVTK
Atha	120	RKAIWAIYVWCRRTDELVDGPNASHITPMALDRWEARLEDFRGRPFDMLEDAALADTVAR
Lses	181	YPVDIQPFKDMIEGMRLDLRKSRYDNFDELYLYCYVAGTVGLMSVPVMGIAPESKASTE
Ljap	164	YPVDIQVNLMLLTG-----
Gmax	181	YPVDIQPFKDMIEGMRLDLRKSRYNSFDELYLYCYVAGTVGLMSVPVMGIAPESKATTE
Mtru	181	YPVDIQPFKDMIEGMRLDLRKSRYNNFDELYLYCYVAGTVGLMSVPVMGIEPQSKASTE
Atha	180	YPVDIQPFKDMIEGMRLDLRKSRYQNFDDLYLYCYVAGTVGLMSVPVMGIDPKSKATTE
Lses	241	SVYNAALALGIANQLTNILRDVGEDARRGRVYLPQDELAQAGLSDEDFISGRVTDKWRNF
Ljap		-----
Gmax	241	SVYNAALALGIANQLTNILRDVGEDARRGRVYLPQDELAQAGLTDDDFIRGKVTDKWRNF
Mtru	241	SIYNAALALGIANQLTNILRDVGEDARRGRVYLPQDELAQAGLSDDDFIRGRVTDKWRNF
Atha	240	SVYNAALALGIANQLTNILRDVGEDARRGRVYLPQDELAQAGLSDEDFAGKVTDKWRNF
Lses	301	MKGQIRRRARMFFDEAEKGVSELSSASRWPVWASLLLYRQILDSIEANDYNNFTKRAYVGK
Ljap		-----
Gmax	301	MKGQIQRRARMFFDEAEKGVSELNSASRWPVWASLLLYGQILDSIEANDYNNFTKRAYVGK
Mtru	301	MKGQIKRRARMFFDEAEKGVSELSSASRWPVWASLLLYRQILDSIEANDYNNFTKRAYVGK
Atha	300	MKMQLKRRARMFFDEAEKGVTELSAASRWPVWASLLLYRRILDEIEANDYNNFTKRAYVGK
Lses	361	AKKLLSLPAA
Ljap		-----
Gmax	361	VKKLLSLPAA
Mtru	361	AKKLLSLPVA
Atha	360	VKKIIAALPLA

Figure S45. Amino acid alignment for ZEP-1

Lses	1	-----
Gmax	1	MATTLCYNSLNPSTTVFSRTHFSVPGCRTKQQRKKVMHVKTSPSKKQLRIILVAGGGIGGLV
Mtru	1	MVSTLSHKCLSPSMTTTFSRTHFSNPGNRRTTKQRRKLMQVKTPQKKQLKVLVAGGGIGGLV
Atha	1	MGSTPFCYSINPSKLDFTTRTHVFSVPGVSGFRSRRALLGVKTEKKKKSRLVAVAGGGIGGLV

Lses	1	-----IT
Gmax	61	FALAAKRKGFEVVMVFEKDLSAIRGEGQYRGPIQIQSNALAALEAIDSEVADEVMRVGCIT
Mtru	61	FALAAKRKGFEVVMVFEKDLSAIRGEGQYRGPIQIQSNALAALEAIDMNVVADEVMRVGCIT
Atha	61	FALAAKKKGFDVLVFEKDLSAIRGEGKYRGPIQIQSNALAALEAIDIEVAEQVMEAGCIT

Lses	3	GDRINGLVDGVSGSWYVKFDTFTPAVERGLPVTRVISRMVLQGILARAVGDDIIMNASNV
Gmax	121	GDRINGLVDGVSGSWYVKFDTFTPAVERGLPVTRVISRMVLQEIARAVGEDIIIMNASNV
Mtru	121	GDRINGLVDGVSGSWYIKFDTFTPAVERGLPVTRVISRMALQEIARAVGDDVIMNGSNV
Atha	121	GDRINGLVDGISGTWYVKFDTFTPAASRGLPVTRVISRMTLQQILARAVGEDVIRNESNV

Lses	63	VNFVDDGNKVTVELENGEKYEGDLLVGADGIWSKVRKQLFGLTEAVYSGYTC-----
Gmax	181	VNFVDDGNKVTVELENGQKYEGDVLVGADGIWSKVRKQLFGLTEAVYSGYTCYTGIADFV
Mtru	181	VDFIDHETKVTVVLNNGQKYDGDLLVGADGIWSKVRTKLFGSTAATYSGYTCYTGIADFV
Atha	181	VDFEDSGDKVTTVLENGQRYEGDLLVGADGIWSKVRNNLFGRSEATYSGYTCYTGIADFI

Lses		-----
Gmax	241	PADIEIVGYRVFLGHKQYFVSSDVGAGKMOWYAFHKEPPGGVDEPNGKKERLLRIFEGWC
Mtru	241	PPDIESVGYRVFLGHKQYFVSSDVGAGKMOWYAFHQEPAGGVDPNGKKERLLKIFEGWC
Atha	241	PADIESVGYRVFLGHKQYFVSSDVGGGKMOWYAFHEEPAGGADAPNGMKKRLFEIFDGWC

Lses		-----
Gmax	301	DNAVDLILATEEEAILRRDIYDRIPTLTWGKGRVTLLGDSVHAMQPNMGQGGCMAIEDSY
Mtru	301	DNAIDLIVATEEEAILRRDIYDRIPTLTWGKGRVTLLGDSVHAMQPNMGQGGCMAIEDGY
Atha	301	DNVLDLLHATEEEAILRRDIYDRSPGFTWGKGRVTLLGDSIHAMQPNMGQGGCMAIEDSF

Lses		-----
Gmax	361	QLAWELENAWEQSIKSGSPIDIDSSLRSYERERRLRVAIIHGMAARMAALMASTYKAYLGV
Mtru	361	QLAFELDNAWQSAKSGSTIDIASSLKSYERERRLRVTFVHGMAARMAALMASTYKAYLGV
Atha	361	QLALELDEAWKQSVETTTTPVDVVSSLKRYEESRRLRVVAIIHAMARMAAIIASTYKAYLGV

Lses		-----
Gmax	421	GLGPLEFLTKFRIPHPGRVGGRRFFVDIMMPSMLSQVVLGGNSDKLEGRPLSCRLTDK
Mtru	421	GLGPFEFLTKFRIPHPGRVGGRRFFIQKSMPLMLNWLGGNSSKLEGRPLCCRLSDK
Atha	421	GLGPLSFLTKFRVPHHPGRVGGRRFFVDIAMPSMLDQVVLGGNSEKLGQRPPSCRLTDK

Figure S46. Amino acid alignment for ZEP-2

Lses	1	CRLSDKANDQLQRWFEDDDALERA	VNGEWILLPCGDEAGLSK	PICLSQDEM	KPCMIGQED
Ljap	1	-----	-----	-----	QED
Gmax	1	CRLSDKANDQLHRWFEDN	DALERAINGEWILLPCGDEAGP	TKPICLTQDEM	KPCIIGQKD
Mtru	1	CRLSDKASDQLHTTW	FEDDDALERTINGEWILLPCGDV	PGHVKPISLNQDD	TKPYIIGQED

Lses	61	YPGCSITITPLPQVSQM	HARINYKDGAFFLTDLQ	SQHGTWITDNG	GRRYRVFPND	PARVRP
Ljap	4	YPGSSITITLPQVSQM	HAQINYKDGAFFLTDLQ	SQHGTWITDNEGR	RNRVFPNC	PARVRP
Gmax	61	HPGSSITITPLPQVSQM	HARINYKDGAFFLTDLRS	LHGTWITDNEGR	RYRVPPNY	PARVRP
Mtru	61	YPGSLITITPLPQVSQ	LHARINFKDGAFFLTDLRS	SQHGTWITDNEGR	RYMVSPNY	PARTRP

Lses	121	SDLIEFGSNKASYRVKVTR
Ljap	64	SDLIEFGSNKVIDDP----
Gmax	121	SDVVEFGSDKASYRVKVTR
Mtru	121	SHVIEFGCNQASYRVKVTR

Figure S47. Amino acid alignment for ZEP-3

Lses	1	KKKLRILVAGGGIGGLVFALAAKRKGFEVVFEEKDMSAIRGEGKYRGPIQIQSNALAALE
Ljap	1	-----LIIAG-----
Gmax	1	KKKLRVLVAGGGIGGLVFALAAKRKGFEVVFEEKDMSAIRGEGQYRGPIQIQSNALAALE
Mtru	1	KKQLKVLVAGGGIGGLVFALAAKRKGFEVVFEEKDLSAIRGEGQYRGPIQIQSNALAALE
Atha	1	KKKSRLVAGGGIGGLVFALAAKKKGFVVLVFEKDLAISRGEGKYRGPIQIQSNALAALE
Lses	61	AIDLKVADEVMRIGCITGDRINGLVDGISGSWYIKFDTFTPAERGLPVTRVISRMALQE
Ljap	6	-----
Gmax	61	AIDLEVAEEVLRVGCITGDRINGLVDGISGSWYIKFDTFTPAERGLPVTRVISRMALQE
Mtru	61	AIDMNVAEVLRVGCITGDRINGLVDGVSWSWYIKFDTFTPAERGLPVTRVISRMALQE
Atha	61	AIDIEVAEQVMEAGCITGDRINGLVDGISGTWYVKFDTFTPAASRGLPVTRVISRMTLQQ
Lses	121	ILACAVGEDVILNGSNVVDFFIDHGNKVTVELENGQKYDGDLLVGADGIWSKVRKKLFGES
Ljap	6	-----
Gmax	121	ILAHAVGEDVIMNDSNVVDFFVDHGDKVTVELENGQKYDGDLLVGADGIWSKVRKKLFGQT
Mtru	121	ILARAVGDDVIMNGSNVVDFFIDHETKVTVVLNNGQKYDGDLLVGADGIWSKVRTKLFGST
Atha	121	ILARAVGEDVIRNESNVVDFFEDSGDKVTVVLNNGQRYEGDILLVGADGIWSKVRNNLFGRS
Lses	181	EAIYSGYTCYTGIADFVPPDIESVGYRVFLGHKQYFVSSDVGAGKMQWYAFHQEPPGGVD
Ljap	6	-----VSCK-----
Gmax	181	EATYSGYTCYTGIADFVPPDIESVGYRVFLGHKQYFVSSDVGAGKMQWYGFHQEPAGGAD
Mtru	181	EATYSGYTCYTGIADFVPPDIESVGYRVFLGHKQYFVSSDVGAGKMQWYAFHQEPPAGGVD
Atha	181	EATYSGYTCYTGIADFIPADIESVGYRVFLGHKQYFVSSDVGCGKMQWYAFHEEPAGGAD
Lses	241	IPNGRKRRVLKIFKGWCDNVTELIVATEEEAAILRRDIYDRTPTLTWGKGRVTLLGDSVHA
Ljap	10	-----
Gmax	241	IPNGKKERLLKIFKGWCDNVLDLIHATEEEAAILRRDIYDRTPFTTWGKGHVTLTGDSIHA
Mtru	241	TPNGKKERLLKIFEGWCDNAIDLIVATEEEAAILRRDIYDRTPTLTWGKGRVTLLGDSVHA
Atha	241	APNGMKKRLFEIFDGWCDNVLDLLHATEEEAAILRRDIYDRSPGFTWGKGRVTLLGDSIHA
Lses	301	MQPNIGQGGGMAIEDGFQLALELDTAWQQSMKSGSSIDIDSSLKSYERERRLRVAIVHAM
Ljap	10	-----
Gmax	301	MQPNMGQGGCMAIEDSYQLALELDNAWQQSIKSGSPIDIDSSLKSYERERRLRVAIVHGM
Mtru	301	MQPNMGQGGCMAIEDGYQLAFELDNAWQQSAKSGSTIDIASLKSYSERERRLRVTFVHGM
Atha	301	MQPNMGQGGCMAIEDSFQLALELDEAWKQSVETTTTPVDVVSLLKRYEESRRLRVAIIHAM
Lses	361	ARMAALMASTYKAYLGVLGPLEFLTMRFRVPHPGRVGGRYFIDKMMPLMLNWWLGGNSSK
Ljap	10	-----
Gmax	361	ARMAAMMASTYKAYLGVLGPLEFLTFRIPHPGRVGRFFIDKMMPLMLNWWLGGNSSK
Mtru	361	ARMAALMASTYKAYLGVLGPFEFLTFRIPHPGRVGRFFIQKSMPLMLNWWLGGNSSK
Atha	361	ARMAALMASTYKAYLGVLGPLSFLTKFRVPHPGRVGRFFVDIAMPMLDWWLGGNSEK
Lses	421	LEGRPLCCRLSDKANDQLRTWFEDDDALESAINGEWILLPCGDETHGSEPIRISQNEMKP
Ljap	10	-----
Gmax	421	LEGRPVCCRLSDKANDQLHRWFEDNDALERAINGEWILLPCGDEAGPTKPICLTQDEMKP
Mtru	421	LEGRPLCCRLSDKASDQLHTWFEDDDALERTINGEWILLPCGDVPGHVKPIISINQDDTKP

Atha 421 LQGRFPSCRLTDKADDRLREWFEDDDALERTIKGEWYLIPHGDDCCVSETLCLTKDEDQP

Lses 481 CIIGSGTTVTIPSPQVSPMHARINYKDGGFFLIDLRSEHGTWISDIEGKRYRVPPNYPAR
Ljap 10 -----ALQLQYLCPRVSPMHARINYKDGGFFLIDLRSEHGTWISDIEGKRYRVPPNYPAR
Gmax 481 CIIGSGSSIIIPLPQVSQMHARINYKDGAFFLTDLRSLHGTWITDNEGRRYRVPPNYPAR
Mtru 481 YIIGN-----
Atha 481 CIVGSGMRIVIPSSQVYKLYACSCDLQRRSFLLD-----

Lses 541 IHPSDVLEFGSKK
Ljap 65 IHPSDVLEFGSRK
Gmax 541 VRPSDVVEFGSDK
Mtru -----
Atha 515 -----GSSK

Appendix 13

Supplementary Figures S48 – S54 Amino acid alignments for red-differentially expressed flavonoid genes in *Lotus sessilifolius*

Figure S48. Amino acid alignment for ANS

Lses	1	MAPTVERVESLSTSGIQSIPKEYVRPKEELENIGNVFEEKK-DGPQVPTIDLKEIDSP
Ljap	1	MAPTVERVESLSGSGIQSIPKEYVRPKEELANIGDVFEKK-VGPQVPTIDLKDIDSP
Gmax	1	MG-TVAPRVESLASSGIKCIPKEYVRPEKELKSIGNVFEEKK-EGPEVPTIDLREIDSE
Mtru	1	-----KEYVRPKEELANIGNIFDEKK-EGPQVPTIDLKEINSS
Atha	1	MV--AVERVESLAKSGIISIPKEYIRPKEELESINDVFLEKKEDGQPQVPTIDLKNIESD
Lses	60	DEFVRANCREKLKKAEEWGVMHLVNHGIPDELINQLKTAGAEFFSLPVEEKEYANDQT
Ljap	60	DEFVRAKCREKLKKAEEWGVMHLVNHGIPDELINQLKSAGAEFFSLPVEEKEYANDQA
Gmax	59	DEVVRGKCRQKLKKAEEWGVMNLVNHGIQDELIERVKKAGEEFFGLAVEEKEYANDQE
Mtru	39	DEIVRGKCREKLKKAEEWGVMHLVNHGISDDLINRLKKAGETFFELPVEEKEYANDQS
Atha	59	DEKIRENCIEELKKASLDWGMHLINHGIPADLMERVKKAGEEFFSLPVEEKEYANDQA
Lses	120	TGNVQGYGSKLANNASQLEWEDYFFHLIFPEDKRDLSIWPKTPSYYTEVTSYARRLRV
Ljap	120	AGNVQGYGSKLANNASQLEWEDYFFHLIFPEDKRDLSIWPKTPSYYTEVTSYARRLRV
Gmax	119	SGKIQQYGSKLANNASQLEWEDYFFHLVFPEDKRDLSIWPKKPDYITEVTSEYAKRLRG
Mtru	99	SGKIQQYGSKLANNASQLEWEDYFFHCIFPEDKRDLSIWPKTPADYTKVTSEYAKELRV
Atha	119	TGKIQQYGSKLANNASQLEWEDYFFHLAYPEEKRDLSIWPKTPSDYIEATSEYAKCLRL
Lses	180	LASKVLEVLSELGLEEGRLEKEVGGMEELLQMKINYYPKCPQPELALGVEAHTDVSSL
Ljap	180	LASKILEVLSELGLEEGRLEKEVGGMEELLQMKINYYPKCPQPELALGVEAHTDTSAL
Gmax	179	LATKMLEALSIGLGLEGGRLKEVGGMEELLQIKINYYPICPQPELALGVEAHTDVSSL
Mtru	159	LASKIMEVLSELGLEGGRLKEAGGMEELLQMKINYYPICPQPELALGVEAHTDVSSL
Atha	179	LATKVFKALSIVGLGLEPDRLEKEVGGIEELLQMKINYYPKCPQPELALGVEAHTDVSAI
Lses	240	TFLVHNMPVGLQLFYQDKWITAKCVPDSILMHIGDTIEILSNGKFKSILHRGLVNKEKVR
Ljap	240	TFLLNHMPVGLQLFYEGKWVTAKCVPDSILMHIGDTIEILSNGKFKSILHRGLVNKEKVR
Gmax	239	TFLLNHMPVGLQLFYQGQWFTAKCVPNSILMHIGDTIEILSNGKYKSILHRGLVNKEKVR
Mtru	219	TFLLNHMPVGLQLFYEGKWVTAKCVPDSILMHIGDTIEILSNGKYKSILHRGLVNKEKVR
Atha	239	TFILHNMPVGLQLFYEGKWVTAKCVPDSIVMHIGDTIEILSNGKYKSILHRGLVNKEKVR
Lses	300	ISWAVFCEPPKEKIILKPLPELVTETEPALFPPRTFAQHLHHKLFRKDQEAQAQ--SKX
Ljap	300	ISWAVFCEPPKEKIILKPLPELVTETEPARFPPRTFAQHIHHKLFRKDQEAQAQ--SKX
Gmax	299	ISWAMFCEPPKEKIILQPLPELVTETEPARFPPRTFAQHIHHKLFRKDQEGLPN----X
Mtru	279	ISWAVFCEPPKEKIILKPLPELVTETEPARFPPRTFAQHIHHKLFRKDQEEKKDDPKKX
Atha	299	ISWAVFCEPPKDKIVLKPLPEMVSVESPAKFPPRTFAQHIEHKLFGKEQEELVSEKNDX

Figure S49. Amino acid alignment for CHS-1

Lses	1	-----TVEEIRNAQRSHGPATILAFGTATPSHCVMQADYPDYYFRITNSEHMTDLKEK
Ljap	1	MV-----TVEEIRNAQRSSGPATILAFGTATPSHCVMQADYPDYYFRITNSEHMTDLKEK
Gmax	1	MV-----TVEEIRNAQRSHGPATILAFGTATPSNVCVSQADYPDYYFRITNSEHMTDLKEK
Mtru	1	MV-----TVEEIRKAQRSNNGPATILAFGTATPSHCVTQAEYPDYYFRITNSEHMTDLKEK
Atha	1	MVMAGASSLDEIRQAQRADGPAGILAIIGTANPENHVLQAEYPDYYFRITNSEHMTDLKEK
Lses	54	FKRMCEKSMIRKRYMHLTEETILKENPNMCAYMAPSLDARQDLVVVEVPKLGKDAAAKAIAK
Ljap	56	FKRMCEKSMIRKRYMHLTEETILKENPAMCAYMAPSLDARQDLVVVEVPKLGKEAAAKAIAK
Gmax	56	FKRMCEKSMIKKRYMHLTEETFLKENPNMCEYMAPSLDVQRQDVVMEVPKLGKQAATKAIAK
Mtru	56	FKRMCEKSMIKKRYMHITETFLKENPNMCAYMAPSLDARQDLVVVEVPKLGKDAAKKAIAA
Atha	61	FKRMC DKSTIRKRHMHLEETFLKENPHMCAYMAPSLDTRQDITVVVEVPKLGKEAAVKAIAK
Lses	114	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLKPSVKRLMMYQQGCFAGGTVLRLAK
Ljap	116	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLKPSVKRLMMYQQGCFAGGTVLRLAK
Gmax	116	EWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLRLAK
Mtru	116	EWGQPKSKITHVVFCTTSGVDMPGADYQLTKLLGLKPSVKRLMMYQQGCFAGGTVLRLAK
Atha	121	EWGQPKSKITHVVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAGGTVLRLIAK
Lses	174	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAMIVGSDPDLAV--ER
Ljap	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAIIVGADPDLAV--ER
Gmax	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAALIIVGSDPDPAV--ER
Mtru	176	DLAENNKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAMIIIGADPDITV--ER
Atha	181	DLAENNRGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFSDGAAALIVGSDPDTSVGEK
Lses	233	PIFELVSAAQITILPDSGAIDGHLREVGLTFHLLKDVPGIISKHIEKSLTEAFAPIGVSD
Ljap	235	PIFQLVSAAGITILPDSGAIDGHLREVGLTFHLLKDVPGIISKHIEKSLSEAFAPIGISD
Gmax	235	PIFEMVSAAQITILPDSGAIDGHLREVGLTFHLLKDVPGIISKHIEKSLVEAFEPIGISD
Mtru	235	PIFEIVSAAQITILPDSGAIDGHLREVGLTFHLLKDVPGIISKHIEKSLVEAFAPIGISD
Atha	241	PIFEMVSAAQITILPDSGAIDGHLREVGLTFHLLKDVPGLISKNIIVKSLDEAFKPLGISD
Lses	293	WNSLFWIAHPGGPAILDQVEVKLRRLKEEKL RSTRHVLSEYGNMSSACVLFILDEVRRRSK
Ljap	295	WNSLFWIAHPGGPAILDQVEAKLRRLKEEKL RSTRHVLGEYGNMSSACVLFILDEVRRRSK
Gmax	295	WNSLFWIAHPGGPAILDQVEEKLRRLKPEKLRSTRHVLSEYGNMSSACVLFILDEMRRKRSK
Mtru	295	WNSLFWIAHPGGPAILDQVEEKLRRLKEEKL RSTRHVLSEYGNMSSACVLFILDEMRRKRSK
Atha	301	WNSLFWIAHPGGPAILDQVEIKLGLKEEKL RSTRHVLSEYGNMSSACVLFILDEMRRKSA
Lses	353	EEGKETTTGEGLEWGVLF GFGPGLTVETVVLHSVP----
Ljap	355	EEGKETTTGEGLEWGVLF GFGPGLTVETVVLHSVPLEAX
Gmax	355	EEGKETTTGEGLEWGVLF GFGPGLTVETVVLHSVPLEGX
Mtru	355	EEGKITTTGEGLEWGVLF GFGPGLTVETVVLHSVPVQGX
Atha	361	KDGVATTGEGLEWGVLF GFGPGLTVETVVLHSVPL--X

Figure S50. Amino acid alignment for CHS-6

Lses	1	MVSVAEIRKAQRAEGPATILAIGTANPPNCVDQSTYPDFYFRITNSEHMTTELKEKFQRM
Ljap	1	MVSVAEIRKAQRAEGPATIIFAIGTANPPNCVDQSTYPDFYFRVITNSQHKTELKEKFQRM
Gmax	1	MVSVAEIRQAQRAEGPATILAIGTANPPNCVDQSTYPDYFRITNSEHMTTELKEKFQRM
Mtru	1	MVSVSEIRKAQRAEGPATILAIGTANPANCVETQSTYPDFYFKITNSEHMTTELKEKFQRM
Atha	1	MVMACASSLDEIRQAQRADGPAGILAIGTANPENHVLQAEYYPDYFRITNSEHMTDLKEK

Lses	61	DKSMIKKRYMHLTEDLLKENPNMCA-----
Ljap	61	DKSMIKKRYMHLTEDLLKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP
Gmax	61	DKSMIKKRYMYLNEEILKENPNMCAYMAPSLDARQDMVVVEVPKLGKEAAVKAKEWGQP
Mtru	61	DKSMIKKRYMYLTEEILKKNPSVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP
Atha	61	FKRMCDKSTIRKRHMHITEEFLKENPHMCAYMAPSLDTRQDIVVVEVPKLGKEAAVKAIK

Lses		-----
Ljap	121	KSKITHLIFCTTSGVDMPGADYQLTKLLGLRPYVKRNKGALQVARCFVWPKTWLRRTTKVL
Gmax	121	KSKITHLIFCTTSGVDMPGADYQLTKQLGLRPYVKRYMMYQQGCFAAGTTLRLAKDLAEN
Mtru	121	KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAAGTTLRLAKDLAEN
Atha	121	EWGQPKSKITHVVFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCFAAGTTLRIAK

Lses		-----
Ljap	181	VCXMFVNLNSLQLPFVALVTLTXTALLGKHCLEMEQLHSLLPVITQCLKLRNLCSSXYGLRK
Gmax	181	NKGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGAAAVIVGSDPI PQVEKPLYELV
Mtru	181	NKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAAALIVGSDPI PEIERPIFEMV
Atha	181	DLAENNRGARVLVVCSEITAVTFRGPSDTHLDSL VGQALFSDGAAALIVGSDPDTSVGEK

Lses		-----
Ljap	241	LLLQIVKEPLMVISVKLDXPFISLKMFPGLFQRTL RKHXSRPSNRXAYLITTYQYFGLHTQ
Gmax	241	WTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGLIVSKNIDKALFEAFNPLNISDYNISIFW
Mtru	241	WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGLIVSKNITKALVEAFEPLGISDYNISIFW
Atha	241	PIFEMVSAAQTILPDSGAIDGHLREVGLTFHLLKDVPGLISKNIKSLDEAFKPLGISD

Lses		-----
Ljap	301	AAQQFLIRLSRTNSXSGQA-----
Gmax	301	IAHPGGPAILDQVEQKLGLKPEKMKATRDLVSEYGNMSSACVLFILDEMRRKSAENGLKT
Mtru	301	IAHPGGPAILDQVEQKLAIKPEKMNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT
Atha	301	WNSLFWIAHPGGPAILDQVEIKLGLKEEKMRATRHLVSEYGNMSSACVLFILDEMRRKSA

Lses		-----
Ljap		-----
Gmax	361	TGEGLEWGVLFVFGFGPGLTIETVVLRSVAI-----X
Mtru	361	TGEGLEWGVLFVFGFGPGLTIETVVLRSVAI-----X
Atha	361	KDGVATTGEGLEWGVLFVFGFGPGLTVETVVLHVSPLX

Figure S51. Amino acid alignment for CHS-7

Lses	1	-----
Ljap	1	MVSVAEIRKAQRAEGPATIEFAIGTANPANCVDQSTYPDFYFRVTNSEHKTELKEKFQRM
Gmax	1	MVSVAEIRQAQRAEGPATIIAIGTANP EN RVDQSTYPDYFRITNS DH MTELKEKFQRM
Mtru	1	MVSVS E IRKAQRAEGPATIMAIGTANPANC E QSTYPDFYFKITNSEHK V ELKEKFQRM

Lses	1	-----
Ljap	61	DKSMIK K RYM H LTE D ILKENPNMCAYMAPSLDARQDMVVVEVPRLGKEAAVKAIKEWGQP
Gmax	61	DKSMIKTRYMYL N EEILKENPNMCAYMAPSLDARQDMVVVEVP K LGKEAAVKAIKEWGQP
Mtru	61	DKSMIK R RYMYL T EEILKENP S V C EYMAPSLDARQDMVVVEVPRLGKEAAVKAIKEWGQP

Lses	1	-----
Ljap	121	KSKITHLIFCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLR M AKDLAEN
Gmax	121	KSKITHLIFCTTSGVDMPGADYQLTK Q LGLRPYVKRYMMYQQGCFAGGTVLR L AKDLAEN
Mtru	121	KSKITHLIFCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLR L AKDLAEN

Lses	1	-----
Ljap	181	NKGARVLVVCSEITAVTFRGPSDTHLDSLVGQALFGDGAAALIVGSDP V PEIEKPLFELV
Gmax	181	NKGARVLVVCSEITAVTFRGPSDTHLDSLVGQALFGDGAAA V IVGSDPIP Q VEKPLYELV
Mtru	181	NKGARVLVVCSEVITAVTFRGPSDTHLDSLVGQALFGDGAAALIVGSDPIPEIEKPI F FEMV

Lses	1	-----HLLKDVP
Ljap	241	WTAQTIAPDSEGAIDGHLREVGLTFHLLKDVP
Gmax	241	WTAQTIAPDSEGAIDGHLREVGLTFHLLKDVP
Mtru	241	WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVP

Lses	36	IAHPGGPAILDQVEQKLSLKPEKMRATREVLSEYGNMSSACVLFILDEMRRKSAQDGLKT
Ljap	301	IAHPGGPAILDQVEQKLSLKPEKMRATREVLSEYGNMSSACVLFILDEMRRKSAQDGLKT
Gmax	301	IAHPGGPAILDQVEQKLGLKPEKMKATRDVLSEYGNMSSACVLFILDEMRRKSA EN GHKT
Mtru	301	IAHPGGPAILDQVE E KLGLKPEKMKATREVLSEYGNMSSACVLFILDEMRRKSAQ E GLKT

Lses	96	TGEGLEWGVLF G FGPGLTIETVVLRSVAIX
Ljap	361	TGEGLEWGVLF G FGPGLTIETVVLRSVAIX
Gmax	361	TGEGLEWGVLF G FGPGLTIETVVLH S VAIX
Mtru	361	TGEG L DWGVLF G FGPGLTIETVVLH S VAIX

Figure S52. Amino acid alignment for DFR-2

Lses	1	-----ETVCVTGAAGFIGSWLVMRLIERGYTVRATVRDPANMKKVKHLLELPNAKTKL
Ljap	1	MSS--ESETVCVTGAAGFIGSWLVMRLIERGYTVRATVRDPANMKKVKHLLELPDAKTKL
Gmax	1	MGSSSASESVCVTGASGFIGSWLVMRLIERGYTVRATVRDPANMKKVKHLVELPGAKTKL
Mtru	1	MGS--VSETVCVTGASGFIGSWLVMRLMERGYTVRATVRDPANMKKVKHLLELPGANSKL
Atha	1	MVS--QKETVCVTGASGFIGSWLVMRLIERGYFVRATVRDPGNLKKVQHLLDLPNAKTL
Lses	54	TLWKADLAEEGSFDEAIKCTGVFHVATPMDFESKDPENEVIKPTINGLLDILKACEKAK
Ljap	59	SLWKADLAEEGSFDEAIRGCTGVFHVATPMDFESKDPENEVIKPTINGLLDILKACEKAK
Gmax	61	SLWKADLAQEGSFDEAIKCTGVFHVATPMDFESKDPENEVIKPTINGLLDIMKACVKAK
Mtru	59	SLWKADLGEEGSFDEAIKCTGVFHVATPMDFESKDPEKEVINPTINGLLDIMKACKKAK
Atha	59	TLWKADLSEEGSYDDAINGCDGVFHVATPMDFESKDPENEVIKPTVNGMLGIMKACVKAK
Lses	114	SVRRLVFTSSAGTVDVTEQOKPVIDETCWSDIEFCLRVKMTGWMYFVSKTRAEQEAWKYA
Ljap	119	TVRRLVFTSSAGTVDVTEHPKPVIDETCWSDIEFCLRVKMTGWMYFVSKTRAEQEAWKYA
Gmax	121	TVRRLVFTSSAGTVDVTEHPNPVIDENCWSVDVDFCTRVKMTGWMYFVSKTLAEQEAWKYA
Mtru	119	TVRRLVFTSSAGTLDVTEQQNSVIDETCWSDVEFCRRVKMTGWMYFVSKTLAEQEAWKFS
Atha	119	TVRREVFVFTSSAGTVNVEEHQKNVYDENDWSDLIEFIMSKMTGWMYFVSKTLAEKAAWDEFA
Lses	174	KEHNIDFVSVIPPLVVGPFPLMPTMPPSLITALSLITGNEAHYSIIKQGQYVHLDDLCLAH
Ljap	179	KEHNIDFVSVIPPLVVGPFPLMPTMPPSLITALSLITGNEAHYSIIKQGQYVHLDDLCLAH
Gmax	181	KEHNIDFISVIPPLVVGPFPLMPTMPPSLITALSLITGNEASHYHIKQGQFVHLDDLCLGH
Mtru	179	KEHNIDFVSIIPPLVVGPFIMPSPMPPSLITALSLITGYEAHYSIIKQGQYIHLDDLCLAH
Atha	179	EEKGLDFISIIPTLVVGPFITTSMPPSLITALSPITRNEAHYSIIRQGQYVHLDDLCLNAH
Lses	234	IFLFENPKAQGRYMCSAYEATIHEVARMINKKYPEFNVPTKFKDIPDELDIKFSSKKIT
Ljap	239	IFLFENPKAQGRYMCSAYEATIHEVARMINKKYPEFNVPTKFKDIPDELDIKFSSKKIT
Gmax	241	IFVFENPKAEGRYICCSHEATIHDIAKLLNQKYPEYNVLTkfkNIPDELDIKFSSKKIT
Mtru	239	IFLFENPKAHGRYICCSHEATIHEVAKLINKKYPEFNVPTKFKDIPDLEIKFSSKKIT
Atha	239	IFLYEQAAAKGRYICSSHDATILTISKFLRPKYPEYNVPSTFEGVDENLKSIEFSSKKLT
Lses	294	DLGFKFKYSLEDMYTGAIETCREKGLLPKTTS-----
Ljap	299	DLGFKFKYSLEDMYTGAVETCREKGLLPKTA-----
Gmax	301	DLGFKFKYSLEDMFTGAVETCREKGLLPKPE-----
Mtru	299	DLGFIIFKYSLEDMFTGAIETCREKGLLPKVT-----
Atha	299	DMGFNFKYSLEEMFIESIETCRQKGFLPVSLSYQSISEIKTKNENIDVKTGDGLTDGMKP
Lses	325	---ETPVTN-----GTTQKX
Ljap	330	---ETPATN-----GTTQKX
Gmax	332	---ETTVNNELLPKPAETTVNDTMQKX
Mtru	330	---ETPV-N-----DTMKKX
Atha	359	CNKTETGITGERTDAP---MLAQQMCA

Figure S53. Amino acid alignment for F3H

```

Lses      1  -----LTTLAQQNTLESSFVRDEDERPKVAYNNFSNEIPVISLAGIDEVDGRRSEIC
Ljap      1  MASFKPKTTLTTLAQQNTLESSFVRDEDERPKVAYNNFSNEIPVISLAGIDEVDGRRSEIC
Gmax      1  MAPT-AKTLTYLAQEKTLLESSFVRDEERPKVAYNEFSDEIPVISLAGIDEVDGRRREIC
Mtru      1  MAP--AQTLTYLAQEKTLLESSFVRDEDERPKVAYNNFSNEIPTISLDGIDDAGGRRRAEIC
Atha      1  MAP---GTLTELAGESKLNSKFVRDEDERPKVAYNVFSDEIPVISLAGIDDVDGKRGEIC

Lses      53  NKIVEACENWGIFQVVNHGVDTELVSHTTTLAKEFFALPPEEKLRFDMSGGKKGGFIVSS
Ljap      61  NKIVEACENWGIFQVVDHGVDTELVSHTTTLAKEFFALPPEEKLRFDMSGGKKGGFIVSS
Gmax      60  EKIVEACENWGIFQVVDHGVDQQLVAEMTRLAKEFFALPPDEKLRFDMSGAKKKGGFIVSS
Mtru      59  NKIVEACENWGIFQVVDHGVD SKLI SEMTRFAKGFFDL PPEEKLRFDMSGGKKGGFIVSS
Atha      58  RQIVEACENWGIFQVVDHGVD TNLVADMTRLARDF FALPPED KLRFDMSGGKKGGFIVSS

Lses      113  HLQGESVQDWREIVTYFSYPIRQDYSRWPDTPAGWKAVTEEYSEKLMGLACKLLEVLSE
Ljap      121  HLQGESVQDWREIVTYFSYPIRNRDYSRWPDTPAGWKAVTEEYSEKLMGLACKLLEVLSE
Gmax      120  HLQGESVQDWREIVTYFSYPKRERDYSRWPDTPEGWRSVTEEYS DKVMGLACKLMEVLSE
Mtru      119  HLQGEAVKQDWREIVTYFSYPIRQDYSRWPDKPEGWKEVTEQYSEKLMNGLACKLLEVLSE
Atha      118  HLQGEAVQDWREIVTYFSYPIRNRDYSRWPDKPEGWVKVTEEYSERLMSGLACKLLEVLSE

Lses      173  AMGLEKEALTKACVDMQKVVVNYYPKCPQPDLTGLGKRHTDPGTITLLLQDQVGGGLQAT
Ljap      181  AMGLEKEALTKACVDMQKVVVNYYPKCPQPDLTGLGKRHTDPGTITLLLQDQVGGGLQAT
Gmax      180  AMGLEKEGLSKACVDMQKVVVNYYPKCPQPDLTGLGKRHTDPGTITLLLQDQVGGGLQAT
Mtru      179  AMGLEKDALTKACVDMQKVVVNYYPKCPQPDLTGLGKRHTDPGTITLLLQDQVGGGLQAT
Atha      178  AMGLEKESLTNACVDMQKITVVVNYYPKCPQPDLTGLGKRHTDPGTITLLLQDQVGGGLQAT

Lses      233  RDNGKTWITVQPVGEAFVVNLGDHGHYLSNGRFKNADHQAVVNSNSSRLSIATFQNPAPD
Ljap      241  RDNGKTWITVQPVGEAFVVNLGDHGHYLSNGRFKNADHQAVVNSNSSRLSIATFQNPAPD
Gmax      240  RDNGKTWITVQPVGEAFVVNLGDHGHYLSNGRFKNADHQAVVNSNHSRLSIATFQNPAPN
Mtru      239  RDNGKTWITVQPVGEAFVVNLGDHGHYLSNGRFKNADHQAVVNSNYSRLSIATFQNPAPD
Atha      238  RDNGKTWITVQPVGEAFVVNLGDHGHYLSNGRFKNADHQAVVNSNSSRLSIATFQNPAPD

Lses      293  ATVYPLKVREGEKSVLDEPITFAEMYRRKMSKDIELARMKKLAKEK-KLQDLE-----
Ljap      301  ATVYPLKVREGEKSVMEEPITFAEMYRRKMSKDIELARMKKLAKEK-KLQDLE-----
Gmax      300  ATVYPLKIREGEKPVMEEPITFAEMYRRKMSKDIELARMKKLAKEK-HLQDLENEKHLQE
Mtru      299  ATVYPLKIRDGEKSVMEEPITFAEMYRRKMSKDIELARMKKLAKEEKEKLRLDLE-----
Atha      298  ATVYPLKVREGEKATLEEPITFAEMYRRKMGRLLELARKKLAKEE-RD-----

Lses      345  ---KAKLEPKPMNEIFA-
Ljap      353  ---KAKLEPKPMNEIFAX
Gmax      359  LDQKAKLEAKPLKEILAX
Mtru      352  ---KAKIEAKPLNEILAX
Atha      346  ----HKEVDKPVQIFAX

```

Figure S54. Amino acid alignment for F3'H

Lses	1	MFQWMIIGFATITVIFFIHRVVKFAT-RP-SLPLPPGPKPWPIIGNLPHMGVPVPHHSLAA
Ljap	1	MFPWMIIGFATITFLIFFIHRVVKFST-RP-SLPLPPGPKPWPIIGNLPHMGVPVPHHSLAA
Gmax	1	MSPLI-VALATIAAAILIYRIIKFIT-RP-SLPLPPGPKPWPIVGNLPHMGVPVPHHSLAA
Mtru	1	MSLWF-TAIASTFLCILIYRFMKFAK-RSSSLPLPPGPKPWPIIGNMPLHGPAPHQSIATA
Atha	1	MATLF-LTILLATVLFILIRIFSHRRNRSHNNRLPPGPNPWPIIGNLPHMGTKPHRTLISA
Lses	59	LAQTHGPLMHLKLGFDVDDVVAASATVAEQFLKVHDANFSSRPPNAGAKYIAYNYQDLVFA
Ljap	59	LARAHGPLMHLKLGFDVDDVVAASAAVAEQFLKVHDANFSSRPPNAGAKYIAYNYQDLVFA
Gmax	58	LARIHGPLMHLRLGFDVDDVVAASASVAEQFLKTHDSNFSSRPPNAGAKYIAYNYQDLVFA
Mtru	59	LAKIHGPLMHLKLGFDVDDVVAASGSVAEQFLKVHDANFSSRPPNTGAKYIAYNYQDLVFA
Atha	60	MVTTYGPILHLRLGFDVDDVVAASKSVAEQFLKTHDANFASRPPNSGAKHMAINYQDLVFA
Lses	119	PYGARWRYLRKITSLLHLSFGKALDNFKHLRQEEVSRLTRNISK--SNSKVVNLGQLLNVC
Ljap	119	PYGARWRYLRKITNLHLSFGKALDNFKHLRQEEVSRLTRNISK--SNSKVVNLGQLLNVC
Gmax	118	PYGPRWRLLRKITSVHLFSKGAMNEFRHLRQEEVARLTCLNAS--SDTKAVNLGQLLNVC
Mtru	119	PYGPRWRMLRKISSVHLFSNKVMEEFKHLRQEEVARLTCLNASNYSDTKAVNLGQLLNVC
Atha	120	PYGHRWRLLRKISSVHLFSAKALEDKHFVRQEEVGTLTRELVLR--VGTKPVNLGQLLVNMC
Lses	177	TTNALS RVMIGRRVFNDGDDGGCDPRADEFKAMVVELMVLAGVFNIGDFIPSLEWLDLQGV
Ljap	177	TTNALS RVMIGRRVFNDGDDGGCDPRADEFKAMVVELMVLAGVFNIGDFIPSLEWLDLQGV
Gmax	176	TTNALARAMIGRRVFNDGNGGCDPRADEFKAMVVELMVLAGVFNIGDFIPSLEWLDLQGV
Mtru	179	TTNALARVMI GRRVFNDGNGGCDPKADEFKEMVLELMVLAGVFNISDFIPSLEWLDLQGV
Atha	178	VVNALGREMIGRRIFGA---DADHKADEFRSMVTMMALAGVFNIGDFVPSLDWLDLQGV
Lses	237	QAKMKNLHNRFD EFLTSIIEEHDT-SSKSENHKDLLSTLLSLKEV-ADDDGNKLN DIEIK
Ljap	237	QAKMKKLHNRFD EFLTSIIEEHNT-SSKSENHKDLLSTLLSLKDV-PDDDGKNLN DIEIK
Gmax	236	QAKMKKLHNRFD AFLTSIIEEHNSSSKNENHKNFLSILLSLKDV-RDDHGNHLTDTEIK
Mtru	239	QAKMKKLHNRFD AFLTNII DERENS NF KSEKHKDLLSTLLLLKEE-TDVGDNKLTYTEIK
Atha	235	AGKMKRLHNRFD AFLSSILKEHEM-NGDQKHDTMLSTLISLKGTDLDGDDGSLTDTEIK
Lses	295	ALLLNMF TAGTDT SASTTEW AIAELIRNPKI LAQVQQELDTVVGRDRNVREDDIPLHPYL
Ljap	295	ALLLNMF TAGTDT SASTTEW AIAELIRSPRI LAQVQQELDTVVGRERNVREDDIPLHPYL
Gmax	295	ALLLNMF TAGTDT SSSTTEW AIAELIKNPQILAKIQQEELDTVVGRDRSVKEEDIAHLPYL
Mtru	298	ALLLNMF AAGTDT SSSTTEW AIAELIRNPRILAQVQQELDNVVGDRNVKEDDIPNLPYL
Atha	294	ALLLNMF TAGTDT SASTVDWAIAELIRHPDIMVKAQEELDIVVGDRPFVNESDIAQLPYL
Lses	355	QAVVKETFRLHPSTPLSLPRVASESCEVLGYHIPKGSTLLVNVWAIARDPKEWADPLEFK
Ljap	355	QAVVKETFRLHPSTPLSLPRVASESCEVLGYHIPKGSTLLVNVWAIARDPKEWADPLEFK
Gmax	355	QAVIKETFRLHPSTPLSVPRAAAESCEIFGYHIPKGATLLVNIWAIARDPKEWNDPLEFR
Mtru	358	QAVIKETFRLHPSTPLSPRIASESCEIFGYHIPKGSTLLVNVWAIARDPKEWVDPLEFK
Atha	354	QAVIKENFRLHPPTPLSLPHIASESCEINGYHIPKGSTLLTNIWAIARDPDQWS DPLAFK
Lses	415	PERFLQGGGDKVHVVDVKGNDFEVIFPGAGRRICAGMSLGLRMVQLLTATLAHSFNWELEN
Ljap	415	PERFLE--GDKVDVDVKGNDFKVIFPGAGRRICAGMSLGLRMVQLLTATLVHSFNWELEN
Gmax	415	PERFLLG-GEKADVDRGNDFEVIFPGAGRRICAGISLGLQMVQLLTAA LAHSFDWELED
Mtru	418	PERFLPG-GEKCDVDVKGNDFEVIFPGAGRRICPGMSLGLRMVQLLTATLAHSFDWELEN

Atha 414 PERFLPG-GEKSGVDVKGSDFELIPFGAGRRICAGLSLGLRTIQFLTATLVQGFWDWELAG

Lses 475 GLNHEKLNMDLAYGLTLQRAVPLSVYSRPRLSPHV-----

Ljap 473 GLNHEKLNMDLAYGLTLQRAVPLSVYSRPRLSPHVYAAS-HX

Gmax 474 CMNPEKLNMDLAYGLTLQRAVPLSVHPRPRLAPHVYSMS-SX

Mtru 477 GLNAGKMMNMDGYGLTLQRAVPLSVHPRPRLSPHVYSSC-FX

Atha 473 GVTPEKLNMEESYGLTLQRAVPLVHPRPRLAPNVYGLGSGX

Appendix 14

Supplementary Figures S55 Amino acid alignments for yellow-differentially expressed flavonoid genes in *Lotus sessilifolius*

Figure S55. Amino acid alignment for FLS

Lses	1	-EVLRVQTVATQSKD-A--SIPSMFVRPKTEQPGLTTRVGVELEVPIIDLSGSDEEKVLR
Ljap	1	MEVLRVQSVAAQSKD-A--SIPAMFVRSETEQPGITTTRVGVELEVPIIDLNGTDEVKVLS
Gmax	1	MEVLRVQTIASKSKD-A--AIPAMFVRAETEQPGITTVOGVNLEVPIIDFSDPDEGKVH
Mtru	1	MEVLRVQTIHQSKD-ATSTIPSMFLRSETESPGTTTVQGVKLEVPIIDFNNPDEGKIQN
Atha	1	MEVERVQDISSSSLLTE--AIPLEFIRSEKEQPAITTFRGPTPAIPVVDLSDPDEESVRR
Lses	57	EIVEASKEWGMFQVVNHEIPVEVIAKLQSVGKHFFELPQGEKEVCGKIDGSDSVEGYGTK
Ljap	58	EIVEASKEWGMFQVVNHEIPSEVIAKLQAVGKEFFELPQEEKEVYCKIEGSDSLEGYGTK
Gmax	58	EILEASRDWGMFQIVNHDIPSDVIRKLQSVGKMFFELPQEEKELIAKPAGSDSIEGYGTK
Mtru	60	EIMEASTKWMFQIVNHDIPNEVIKKLQSVGKEFFELPQEEKEVIAKPIGSDSLEGYGTK
Atha	59	AVVKASEEWGLFQVVNHGIPTELIRRLQDVGRKFFELPSSEKESVAKPEDSKDIEGYGTK
Lses	117	LQKEVNGKKGWVDHLFHIIWPTSAINYRFWPKNPPSYREVNEEYGYKLRVSGDKLFKSLS
Ljap	118	LQKEVNGKKGWVDHLFHIIWPTSSINYRFWPKNPASYREVNEEYGYKLRVADKLFKSMS
Gmax	118	LQKEVNGKKGWVDHLFHIWPPSSINYSFWPQNPPSYREVNEEYCKHLRGVVDKLFKSMS
Mtru	120	LEKEVNGKKGWVDHLFHIWPPSSINYRFWPNNPASYREVNEEYGYKLREVDKLFKSLS
Atha	119	LQKDPEGKKAWVDHLFHRIWPPSCVNYRFWPKNPPEYREVNEEYAVHVKKLSETLLGILS
Lses	177	IGLGLEENELKEAAGGDDMIHLKINYYPPCPCPELVLGVPPTHDMFSMTILVPNEVQGL
Ljap	178	IGLGLEENELKEAAGGDDMIHLKINYYPPCPCPDVLGVPPHTDMSFVTILVPNEVQGL
Gmax	178	VGLGLEENELKEGANEDDMHYLLKINYYPPCPCPDVLGVPPHTDMSYLTILVPNEVQGL
Mtru	180	IGLELEEHELKKAAGGDELIHMLKINYYPPCPVDPDLVLGVPPHTDMSFLTILVPNEVQGL
Atha	179	DGLGLKRDALKEGLGGEMAEMYMMKINYYPPCPRPDIALGVPAHTDLSGITILVPNEVPGL
Lses	237	QAFRDGHWYDVKYVPNALVIHIGDQMEIVSNGKYKAVLHRTTVNKEETRMSWPVFIEPQG
Ljap	238	QAFRDGHWYDVKYVPNALVIHIGDQMEILSNGKYKAVLHRTTVNKEETRMSWPVFIEPQG
Gmax	238	QACRDGHWYDVKYVPNALVIHIGDQMEILSNGKYKAVFHRTTVNKEETRMSWPVFIEPKK
Mtru	240	QASRDGQWYDVKYVPNALVIHIGDQMEILSNGKYKAVLHRTTVNKEETRMSWPVFIEPKP
Atha	239	QVFKDDHWFDAEYIIPSAVIVHIGDQILRLSNGRYKNVLHRTTVNDEKETRMSWPVFIEPPR
Lses	297	HHEVGPHSKLVNHNDNPPKYKTKKYKDYAYCKLNKIPQ-X
Ljap	298	DHEVGPHSKLVNQDNPPKYKTKKYKDYAYCKLNKIPQ-X
Gmax	298	EQEVGPHPKLVNQDNPPKYKTKKYKDYAYCKLNKIPQ-X
Mtru	300	EHEIGPHPKLVNQENPPKYKTKKFEDLYCKLNKIPQ-X
Atha	299	EKIVGPIPELTGDDNPPKFKPFAFKDYSYRKLNLPLDX

Appendix 15

Supplementary Figures S56 – S57 Amino acid alignments for non-differentially expressed lignin genes in *Lotus sessilifolius*

Figure S56. Amino acid alignment for CAD

Lses	1	MGSLE--AERTTVGWAARDPSGILSPYTF	TLRNTGPDDVYIKVHYCGICHSDIHQVKNDLG
Ljap	1	MGSLE--AERTTVGWAARDPSGILSPYTF	TLRNTGPDDVYIKVHYCGVCHTDVHQVKNDLG
Gmax	1	MGSLE--AERTTVGLAARDPSGILSPYTYN	LRNTGPDDVYIKVHYCGICHSDLHQIKNDLG
Mtru	1	MGSIEVAERTTVGLAAKDPSGILT	PYTYTLRNTGPDDVYIKIHYCGVCHSDLHQIKNDLG
Atha	1	MGIME--AERKTTGWAARDPSGILSPYTYTLR	ETGPEVDNIRIICCGICHTDLHQTKNDLG
Lses	60	MSNYPMVPGHEVVGEVLEVGSDVTRFR	AGEIVGAGLLVGCKNCTACQSDIEQYCSKKIW
Ljap	60	MSNYPMVPGHEVVGEVLEVGSDVTRFT	VEIVGAGLLVGCKKCTACQSDIEQYCKKKIW
Gmax	60	MSNYPMVPGHEVVGEVLEVGSDVSRFR	VGELVGVGLLVGCCCKNCQPCQQDIENYCSKKIW
Mtru	61	MSNYPMVPGHEVVGEVLEVGSNVTR	FKVGEIVGVGLLVGCCCKSCRACDSEIEQYCNKKIW
Atha	60	MSNYPMVPGHEVVGEVLEVGSDVSKFT	VDIVGVCLVGGCGGSPCERDLEQYCPKKIW
Lses	120	NYNDVYVDGKPTQGGFAETII	IVEQKFVVKIPEGMAPEQVAPLLCAGVTVYSPLAHFGLKE
Ljap	120	NYNDVYVDGKPTQGGFAETII	IVEQKFVVKIPEGMAPEQVAPLLCAAVTVYSPLSHFGLKE
Gmax	120	SYNDVYVDGKPTQGGFAETMV	VEQKFVVKIPEGIAPEQVAPLLCAGVTVYSPLVHFGLKE
Mtru	121	SYNDVYT	DGKIITQGGFAESTVVEQKFVVKIPEGLAPEQVAPLLCAGVTVYSPLSHFGLKT
Atha	120	SYNDVYINGQPTQGGFAKATV	VHQQKFVVKIPEGMAVEQAAPLLCAGVTVYSPLSHFGLKQ
Lses	180	SGLRGGILGLGGVGHMGVKIAKAL	GHHVTVISSDRKKKEAIEDLGVDAYLVSSDTTSMQ
Ljap	180	SGLRGGILGLGGVGHMGVIAKAM	GHHVTVISSDRKKKEAIEDLGADAYLVSSDTTSMQ
Gmax	180	SGLRGGILGLGGVGHMGVKIAKAL	GHHVTVISSDKKKQEALEHLGADQYLVSSDVTAMQ
Mtru	181	PGLRGGILGLGGVGHMGVIAKAF	GHHVTVISSDKKKKEALEDLGADSYLVSSDVTVMQ
Atha	180	PGLRGGILGLGGVGHMGVKIAKAM	GHHVTVISSSNKKREALQDLGADDYVIGSDQAKMS
Lses	240	GAADSLDYIIDTVPVGH	PLEPYLSLLKLDGKLILMGVINTPLQFITPMVMLGRRSITGSF
Ljap	240	EAADSLDYIIDTVPVGH	PLEPYLSLLKLDGKLILMGVINTPLQFITPMVMLGRRSITGSF
Gmax	240	EAADSLDYIIDTVPVGH	PLEPYLSLLKLDGKLILMGVINTPLQFVSPMVMLGRRSITGSF
Mtru	241	EAADSLDYIIDTVPVGH	PLEPYLSLLKIDGKLILMGVINTPLQFVTPMVMLGRKSITGSF
Atha	240	ELADSLDYVIDTVPVHH	HALEPYLSLLKLDGKLILMGVINNPLQFITPLIMLGEKVITGSF
Lses	300	IGSIKETEEMLEFWKEKGLSSMIE	VVNMDYINKAFERLEKNDVRYRFVVDVKGS-----
Ljap	300	IGSIKETEEMLEFWKEKGLTSMIE	IVVNMDYINKAFERLEKNDVRYRFVVDVKGSKLIDQX
Gmax	300	IGSMKETEEMLEFWKEKGLSSMIE	VVNMDYINKAFERLEKNDVRYRFVVDVKGSKLVD-X
Mtru	301	VGSVKETEEMLEFWKEKGLSSMIE	IVTMDYINKAFERLEKNDVRYRFVVDVKGSKFED-X
Atha	300	IGSMKETEEMLEFC	KEKGLSSIIEVVKMDYVNTAFERLEKNDVRYRFVVDVEGSNLDA-X

Figure S57. Amino acid alignment for CCR-3

Lses	1	-----KVCVTGAGGFVASWLVLKLLSKGYTVHGTVRQP--GSPKYEHLLKLEKA
Ljap	1	-----KVCVTGAGGFVASWLVLKLLSKGYTVHGTVRQP--GSPKYEHLLKLEKA
Gmax	1	MAA-----KVCVTGAGGFVASWLVLKLLSKGYIVHGTVRDPEPATQKYEHLLKLHGA
Mtru	1	MEASGGVNNNKVCVTGAGGFVASWLVLKLLSKGYFVHGTVREP--GSPKYEHLLKLEKA
Atha	1	MAD----VHKGKVCVTGAGGFVLSWVVDLLLSKDYFVHGTVRDP--DNEKYAHLKKLEKA
Lses	48	SENLTFLFKADLLNYESVHSAILGCTAVFHVASFPVPSTISSNPQVEVIEPAVKGTANVLEA
Ljap	1	-----NVFHFV-----
Gmax	54	SENLTFLFKADLLNYESLSAISGCTAVFHLACPVPSISVPNPQVEMIEPAVKGTTNVLEA
Mtru	59	SENLTFLFKADILDYESVYSAIVGCSAVFHVASFPVPSTVVPNPQVEVIEPAVKGTANVLEA
Atha	55	GDKIKLFLFKADLLDYGSLQSAIAGCSGVFHVACPVPPASVPNPQVEVIEPAVDGTLNVLKA
Lses	108	SLQAKVERVVFVSSEAAICMNPNLPKDKLIDESYWSDKDYCKKTQNWYCYSKTEAEEQAL
Ljap	7	-----
Gmax	114	SLEAKVQRIIVFVSSLAATISNSPNLPKDKVIDESYWSDKDYCKTTQNWYCFKTEAEEQAL
Mtru	119	CLKANVERVVFVSSVAAVAINPNLPKDKAIDESCWSDKDYCKNTKNWYCYAKTEAEEQAL
Atha	115	CIEANVKRVVYVSSVAAAFMNPMSKQVLDEACWSDDQYCKKTENWYCLAKTRAESEAF
Lses	168	DFAKRTGLSVVSICPTTVLGPILQS--TVNASSLILLKLLKAEGCDSMENKLNWIVDVRDL
Ljap	7	-----GLVFELDCLKLKVLYSMENRLSWIVDVGD
Gmax	174	DFAKRTGLDVVSICPSLVLGPILQSTTVNASSLALLKLLK--GVNSMENKIRWIVDVRDV
Mtru	179	HFAKRTGLNVVTICPTLVLPILQS--TTNASSLVLLKLLK--EGCDSVENKLRWIVDVRDV
Atha	175	EFAKRTGLHLVSVCPPTLVLPILQONTVNASSLVLLKLLK--EGFETRDNQERHLVDVRDV
Lses	227	VDAVLLAYEKLEAEGRYICTSHPIKKRDLVEKLKSIYPNYKYPTKFTFVDGYRRLSSEKL
Ljap	36	VDAVLLAYEKLEAEGRYICTSHPIKKRDLVEKLKSIYPYKYPTKFTFVDSYRRLTSEKL
Gmax	232	ADAILLAYEKLEAEGRYICHSHTIKTRDMEKLKSIYPNYKYPKYTEVDDYISFSSEKL
Mtru	237	VNAILLAYENHEADGRYICTSHAIVTRDLVERLKGIIYPNYKYPTNYIEMDDYKMLSSEKL
Atha	234	AQALLLVYEKAEAEGRYICTSHTVKEEIVVEKLKSFYPHYNYPKKYIDAEDRVKVSSEKL
Lses	287	QRLGWKSRPLEETLIDSVESYREAGLLQSEX
Ljap	96	QRLGWKSRPLEETLIDSVESYREAGLLQSE-
Gmax	292	QRLGWKYRSLEETLVDSVESYREAGHLQSEX
Mtru	297	QSLGWKLRPLEETLIDSVESYKEAGLLQSQX
Atha	294	QKLGWTYRPLEETLVDSVESYRKAKLVDD--X

Appendix 16

Supplementary Figures S58 – S60 Amino acid alignments for yellow-differentially expressed lignin genes in *Lotus sessilifolius*

Figure S58. Amino acid alignment for CAD

Lses	1	RRKRMS--NNAGKLV	CVVTGGSGYIASWIVKFLLEHGYTVRATVRDPSNP	KKVEHLLKLDGA			
Ljap	1	----MS--NGAGKLV	CVTGGSGYIASWIVKFLLEHGYTVRATVRDPNNPNK	VEHLLKLDGA			
Gmax	1	----MSNNNAGKV	VCVTGASGFIAWIVKFLLRGYTVRATVRYP	SNLKKVDHLVKLEGA			
Mtru	1	----MS--GE	GKVVCVTGASGFIAWIVKFLLRGYTVRATVRDPSNP	KKVDHLLKLDGA			
Lses	60	KERLHLFKADLLEEGSFDSVVDGCDGVFHTASPVGLIVKDPQAE	MDPAVKGT	LNVLKSC			
Ljap	56	KERLHLFKADLLEEGSFDSAT	-----AELIDPAVKGT	LNVLKSC			
Gmax	57	KERLQLFKADLLEEGSFDSVVEGCHGVFHTASPVRFVNDPQAE	LIDPAVKGT	LNVLKSC			
Mtru	55	KERLQLFKADLLEEGSFDSVVEGCDGVFHTASPVRFVNDPQ	VELIDPALKGT	LNVLKSC			
Lses		-----	-----	-----			
Ljap	95	AKSPSVRRVVL	TSSVGAQYNERPKTPEVLI	DETWFSDPDFCRKSKVCASF	SXPLFRNXS		
Gmax	117	AKSPSVKRVVL	TSSISAVAFNRRPKTPQVVVDETWFSDPDVCRELELWY	TL	SKTLAEDAA		
Mtru	115	AKSTSVKRVVL	TSSNAAVSFNTRPKNPEVVVDETWFSPDFCRESKLWYV	LS	SKTLAEAAA		
Lses		-----	-----	-----			
Ljap	155	-----	-----LE-----	-----	-----		
Gmax	177	WKFVNENS	IDMISINPTMVAGPLLQPEINESVEPI	LN	LINGKPF	PNKSF	GWVDVKDVANA
Mtru	175	WKFVNENN	IDMVSINPTMVAGPLLQPEVNESVEPI	LN	LINGIP	FPNKAIGWVNVKDVANA	
Lses		-----	-----	-----	-----		
Ljap	157	-----AHADGRC-----	-----	-----	-----		
Gmax	237	HI	LAYEIASASGRYCL	VERVIHYSELATILRGLYPTLQIPDKCEVDEPYIPT	YQISTEKA		
Mtru	235	HI	HAYEIASASGRCLL	AERVVHYSELAMILRDL	LYPTLPISDKQSTSSSKFVIL	L	AYTTL
Lses		-----	-----	-----	-----		
Ljap	164	-----	-----	-----	DSD		
Gmax	297	KKDLGIEFT	PLEVSLRETVESFREKKIVNF	NPX			
Mtru	294	FHQIGAEPT	CIQGMQLNPLNPLIYSYYFG	FLPCX			

Figure S59. Amino acid alignment for CCoA-OMT

Lses	1	MATN-----EDQKQTEAGRHQEVGHKSLLQSDALYQYILETSVYPREHEAMKE
Ljap	1	MATN-----EDQKQTEAGRHQEVGHKSLLQSDALYQYILETSVYPREHEAMKE
Gmax	1	MAEQ-----NQNTTEAGRHQEVGHKSLLQSDALYQYILETSVYPREPESMKE
Mtru	1	MATN-----EDQNQTESGRHQEVGHKSLLQSDALYQYILETSVFPREHEAMKE
Atha	1	MATTTTEATKTSSTNGEDQKQSQNLRHQEVGHKSLLQSDDL YQYILETSVYPREPESMKE
Lses	49	LRDLTAKHPWNIMTTSADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALAIPDDGKIL
Ljap	49	LRDLTAKHPWNIMTTSADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALAIPDDGKIL
Gmax	49	LRELTAKHPWNIMTTSADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALALPEDGKIL
Mtru	49	LREVTAKHPWNIMTTSADEGQFLSMLLKLINAKNTMEIGVYTGYSLLATALAIPEDGKIL
Atha	61	LREVTAKHPWNIMTTSADEGQFLNMLIKLVNAKNTMEIGVYTGYSLLATALALPEDGKIL
Lses	109	AMDVNKENYELGLPVIKKAGVAHKIDFREGPALPVLDEMVKDEKNHASYDFIFVDADKDN
Ljap	109	AMDVNKENYELGLPVIKKAGVAHKIDFREGPALPVLDEMVKDEKNHGSYDFIFVDADKDN
Gmax	109	AMDINRENYELGLPVIKKAGVDHKIEFREGPALPVLDEMIKDEKNHGSYDFIFVDADKDN
Mtru	109	AMDINKENYELGLPVIKKAGVDHKIDFREGPALPVLDEMIKDEKNHGSYDFIFVDADKDN
Atha	121	AMDVNRENYELGLPIIEKAGVAHKIDFREGPALPVLDEIVADEKNHGT YDFIFVDADKDN
Lses	169	YLN YHKRLIELVKVGGVIGYDNTLWNGSVVAPPDAPLRKYVRYRDFVLELNKALAADPR
Ljap	169	YLN YHKRLIELVKVGGVIGYDNTLWNGSVVAPPDAPLRKYVRYRDFVLELNKALAVDPR
Gmax	169	YLN YHKRLIELVKVGGVIGYDNTLWNGSVVAPPDAPLRKYVRYRDFVLELNKALAVDPR
Mtru	169	YLN YHKRLIELVKVGGVIGYDNTLWNGSVVAPPDAPLRKYVRYRDFVLELNKALAVDPR
Atha	181	YI NYHKRLIELVKIGGVIGYDNTLWNGSVVAPPDAPMRKYVRYRDFVLELNKALAADPR
Lses	229	IEICMLPVG DGITICRRIKX
Ljap	229	IEICMLPVG DGITICRRIKX
Gmax	229	IEICMLPVG DGITICRRIKX
Mtru	229	IEICMLPVG DGITICRRIKX
Atha	241	IEICMLPVG DGITICRRISX

Figure S60. Amino acid alignment for OMT

Lses	1	MGSTGETQITPTHV	DEEANLFAMQLASASVLPVLKSAI	ELDLLEIIAKAGTGAQISPA		
Ljap	1	MGSTGETQITPTHV	SDEEANLFAMQLASASVLPVLKSAI	ELDLLEIIAKAGTGAQLSPA		
Gmax	1	MGSTGETQITPTHV	SDEEANLFAMQLASASVLPVLKSALELDLLEIIAKAGPGVHLSPS			
Mtru	1	MGSTGETQITPTH	SDEEANLFAMQLASASVLPVLKSALELDLLEIIAKAGPGAQISPI			
Atha	1	MGSTAETQITP	VQVITDEEAALFAMQLASASVLPMA	LKSALELDLLEIMAKNG--SPMSPT		
Lses	61	EIASQLPTTNPDAPT	VLDRI	LRLACYNILTC	SVQTQ	QDGKVQRLYGLAPVAKYLVKNED
Ljap	61	EIASQLPTTNPDAPT	VLDRI	LRLACYNILTC	SVRTE	QDGKVQRLYGLAPVAKYLVKNED
Gmax	61	DIASRLPT	HNPDAPV	MLDRILRLACYNIL	SFSIR	TLPHGKVERLYGLAPVAKYLVKNED
Mtru	61	EIASQLPTTNPE	APV	MLDRILRLACYNILTC	SVRTQ	QDGKVQRLYGLATVAKYLVKNED
Atha	59	EIASKLPT	KNPE	APV	MLDRILRL	ITSYSVLTCSNRKLSGDGVERTYGLGPVCKYLTKNED
Lses	121	GVSISALNLMNQDKVLMESWYHLKDAVLEGGIPFNKAYGMTAFEYHGTDP	RFNKVF	NKGM		
Ljap	121	GVSISALNLMNQDKVLMESWYHLKDAVLEGGIPFNKAYGMTAFEYHGTDP	RFNKVF	NKGM		
Gmax	121	GVSIAALNLMNQDK	ILMESWY	YLLKDAVLEGGIPFNKAYGMTAFEYHGTDP	RFNKVF	NKGM
Mtru	121	GVSISALNLMNQDKVLMESWYHLKDAVLDGGIPFNKAYGMTAFEYHGTDP	RFNKVF	NKGM		
Atha	119	GVSIAALCLMNQDKVLMESWYHLKDAITLDGGIPFNKAYGMSAFEYHGTDP	RFNKVF	NKGM		
Lses	181	SDHSTITMKKILETYTGFEGLKSLVDVGGGTGAVINMIVSKYPTIQ	GINFDLPHVIEDAP			
Ljap	181	SDHSTITMKKILETYTGFEGLKSLVDVGGGTGAVINMIVSKYPTIK	GINFDLPHVIEDAP			
Gmax	181	ADHSTITMKKILETYTGFEGLKSLVDVGGGTGAVIN	MIVSKYPTIKGINFDLPHVIEDAP			
Mtru	181	SDHSTITMKKILETYTGFEGLKSLVDVGGGTGAVINTIVSKYPTIK	GINFDLPHVIEDAP			
Atha	179	SNHSTITMKKILETYKGFEGLTSLVDVGGGIGATL	KMIVSKYPNLKGINFDLPHVIEDAP			
Lses	241	SYPGVEHVGGDMFVSVPKADAVFMKWICH	WSD	EHCLKFLKNCYEALPDNGKVIVAECIL		
Ljap	241	SYPGVEHVGGDMFVSVPKADAVFMKWICH	WSD	EHCLKFLKNCYEALPDNGKVIVAECIL		
Gmax	241	SYPGVEHVGGDMFVSVPKADA	IFMKWICH	WSD	EHCLKFLKNCYEALPDNGKVIVAECIL	
Mtru	241	SYPGVEHVGGDMFVS	IPKADAVFMKWICH	WSD	EHCLKFLKNCYEALPDNGKVIVAECIL	
Atha	239	SHPG	IEHVGGDMFVSVPKGD	ITFMKWICH	WSD	EHCLKFLKNCYESLPEDGKVILAEACIL
Lses	301	PVAPDSSLATKGVVHIDVIMLAHNPPGKERTEKEFEALAKGAGFQG	FKVMCCAFNS	SYVME		
Ljap	301	PVAPDSSLATKGVVHIDVIMLAHNPPGKERTEKEFEALAKGAGFQGFQ	VLCCAFNS	SYVME		
Gmax	301	PVAPDSSLATKGVVHIDVIMLAHNPPGKERTEKEFEALAKGS	GFQGFQ	VLCCAFNTYVME		
Mtru	301	PVAPDSSLATKGVVHIDA	ITMLAHNPPGKERTQ	KEFEDLAKGAGFQGFQ	FKVHCNAFN	TYIME
Atha	299	PETPDSSLSTKQ	VVHVDCITMLAHNPPGKERTEKEFEALAKAS	GFKG	IKVVCD	AFGVNLTIE
Lses	361	FLKKVX				
Ljap	361	FLKKVX				
Gmax	361	FLKKVX				
Mtru	361	FLKKVX				
Atha	359	LLKKIX				

Appendix 17

Figure S61. Expanded phylogeny of *Lotus* spp. exhibiting PACC, including those studied in this thesis. Figure and caption are reproduced from Figure 2 in Ojeda et al. 2013:

Molecular tree based on one nuclear (ITS) and three plastid regions (*CYB6*, *trnH-psbA* and *matK*). The tree was randomly selected from a maximum parsimony (MP) analysis from Ojeda et al., 2012a and Ojeda et al., 2012b. Character mapping of the trait flower color change after anthesis in *Lotus* sections *Pedrosia* and the “rhyncholotus group”. Red branches show clades where this trait has evolved and the numbers on the tree the times this trait evolved within this groups (1–3). Arrows indicate the numbers of reversals, one of which occurred in three species of the “rhyncholotus group”. The species from the outgroup belong to the *Lotus* section are not endemic of the Macaronesian region. Species with (*) are distributed in mainland Africa and/or in Europe (a) represents flower color at anthesis (pre-change) and (b) indicates flower color after change (post-change). (A) *L. jacobaeus*, (B) *L. purpureus*, (C) *L. eriosolen*, (D) *L. sessilifolius* subsp. *sessilifolius*, (E) *L. pyranthus*, (F) *L. emeroides*, (G) *L. argyrodes*, (H) *L. glaucus*, (I) *L. campylocladus*, and (J) *L. lancerottensis*.

