SUPPORTING INFORMATION

Supplemental Figure S1 to S4

Phylogenetic trees for four protein families were constructed to find the number of proteins for enzymes that are likely to have the same catalytic function in flavonoid biosynthesis based on their sequence similarity (see Table S1). Only clades depicting the proteins of interest are highlighted in colour. In these regions, statistically significant nodes (>90% using 100 or 1000 bootstrap datasets) are shown in red and proteins are indicated as follows: Lotus (Lj) - green, Medicago (Ms/Mt) and other legumes - dark green, Arabidopsis (At) - black, monocot species (rice (Os), maize (Zm), Brachypodium (Bd)) - yellow and other species - blue. Relevant accession numbers are shown to the right of the gene name. To view the larger trees, zoom to 400X or greater. These trees are also available at the IT3F website (Bailey et al. 2008).

Supplemental Figure S1. Phylogenetic tree of CHI-like proteins. In Lotus, the proteins involved in isoflavonoid biosynthesis are LjCHI1, LjCHI3 and LjCHI4

Supplemental Figure S2. Phylogenetic tree of the Nmr family containing proteins that encode enzymes for IFR, LAR, PTR and PLR (pinoresinol-lariciresinol reductase).

Supplemental Figure S3. Phylogenetic tree of the epimerase family containing proteins that encode enzymes for DFR and VR.

Supplemental Figure S4. Phylogenetic tree of the 2-oxoglutarate and Fe(II)-dependent oxygenase family containing proteins that encode enzymes for F3H, ANS/LDOX, and FLS.

Supplemental Figure S5. Phylogenetic tree of subgroup 2 proteins from Arabidopsis, Lotus, rice and Brachypodium and a related but distinct set of proteins that were found to date (July 2011) to be present in 13 dicot species, including legume species.

Bootstrap values (>70%) are indicated and are shown in red if they are >90% and help to define a clade of proteins that are distinct from the rest of the proteins in the data set. Of particular note is a subclade of legume-specific proteins within the subgroup 2 sister clade that cluster with a significant bootstrap value of 99% reflecting the difference in their DNA binding domain protein sequence compared to the rest of the subgroup 2 sister clade, including two legume genes from Medicago and Lotus that fall outside the subclade.

Supplemental Figure S6. Phylogenetic tree of the WRKY family showing proteins from Lotus and Medicago.

Indicated are Lotus (Lj) and Medicago (Medtr) proteins with a greater than 2-fold increase (up - red) or decrease (down - blue) in expression after 3 h GSH treatment (Lotus) or 2h YE (Medicago; Naoumkina et al. 2007). Bootstrap values (>70% of 100 datasets) are indicated and are shown in red if they are statistically significant (>90%) and confirm orthologous proteins in Lotus and Medicago.

[*]genes induced in both YE-treated and cells treated with methyl jasmonate (MJ).

Supplemental Figure S7. Graph showing that SYBR-Gold nucleic acid gel stain detects DNA in a linear manner from 4 to 56 ng. This range coincided with the amount of PCR product quantified in the experiments presented in Fig.3 and Fig 4.

Supplemental File S1. Protein sequences with appropriate accession identifiers of all the Lotus TFs reported in this study and of all other R2R3MYB and WRKY proteins used in the phylogenetic analyses.

Supplemental Table S1. Summary of Lotus genes encoding enzymes in the phenylpropanoid pathway used in the microarray experiments. The pFAM protein families to which these genes belong are indicated.

Pathway	Enzyme	pFAM Clan ID (Description)	pFAM ID (Description)	Number Of Genes For Each Enzyme In Lotus
General phenylpropanoid:				
	phenylalanine ammonia lyase (PAL)		PF00221 (Phenylalanine and histidine ammonia-lyase)	14
	cinnamate 4-hydroxylase (C4H;CYP73A)		PF00067 (cytochrome P450)	5
	p-Coumaroyl CoA Ligase (4CL)	CL0378 (ANL)	PF00501 (AMP-binding enzyme)	19
Flavonoid:				
	chalcone synthase (CHS)	CL0046 (thiolase)	PF00195 and PF02797 (chalcone and stilbene synthase)	20
	chalcone reductase (CHR) also known as polyketide reductase (PKR)		PF00248 (Aldo/keto reductase)	17
	chalcone isomerases (CHI)		PF02431 (chalcone-flavanone isomerise)	type I: 1[a] type II: 3[a]
	flavanone 3-hydroxylase (F3H)	CL0029 (Cupin)	PF03171 (2-oxoglutarate and Fe(II)-dependent oxygenase)	2[a]
	dihdroflavonol reductase (DFR)	CL0063 (NADP Rossmann)	PF01370 (NAD dependent epimerase/dehydratase)	5[a]
	flavonol synthase (FLS)	CL0029	PF03171	3[a]
	leucoanthocyanidin dioxygenase (LDOX)	CL0029	PF03171	1[a]
	leucoanthocyanidin reductase (LAR)	CL0063	PF01370	1[a]
Isoflavonoid:				
	isoflavone synthase (IFS; CYP93C)		PF00067	3
	S-Adenosyl-L-Methionine: 2,7,4′-Trihydroxyisoflavanone 4′-O-Methyltransferase (HI4′OMT)	CL0063	PF00891 (O-methyltransferase)	6
	2-hydroxyisoflavanone dehydratase (HID)	CL0028 (AB_hydrolase)	PF07859 (Abhydrolase_3)	4
	isoflavone 2′-hydroxylase (I2′H; CYP81E)		PF00067	3
	isoflavone reductase (IFR)	CL0063	PF05368 (NmrA-like)	2[a]
	vestitone reductase (VR)	CL0063	PF01370	1[a]
	pterocarpan reducase (PTR)	CL0063	PF05368	5[a]

[a] numbers concluded from a phylogenetic tree of the protein family (Supplemental Fig. S1 to S4)

Supplemental Table S2. Genes encoding phenylpropanoid biosynthetic enzymes that increased in expression after three hours or six hours following elicitation, normalised with absolute fold change (FCA).

Pathway	Gene Name	Kazusa Identifier	FCA 0-3hrs	FCA 0-6hrs
General phenylpropanoid:				
	PAL2	chr1.CM0033.800	16.2	5.7
	PAL3	chr1.CM0033.830	10.2	4.7
	PAL4	chr1.CM0033.850	11.4	4.0
	PAL8	chr3.CM0574.220	24.7	5.3
	PAL14	LjSGA_036826.1	36.0	10.7
	CYP73A (C4H1)	chr5.LjT24B10.30	768.7	729.0
	CYP73A (C4H2)	LjSGA_003295.2	4376.0	3512.0
	CYP73A99[b](a) (C4H3)	LjSGA_082453.1	10.7	18.2
	CYP73A99[b] (b) (C4H4)	LjSGA_029001.1	27.7	33.6
	4CL1/16[c]	chr2.CM0018.1240/chr6.CM0057.160	27.9	12.0
	4CL12	LjSGA_063891.1	11.1	16.5
	4CL6	chr4.CM0061.190	7.7	6.3
Flavonoid:				
	CHS8[a]	chr2.CM0018.1080.nc	24.0	31.5
	CHS9[a]	chr2.CM0018.1140.nc	15.6	12.7
	CHS10[a]	chr2.CM0018.700.nc	21.5	18.4
	CHS12[a]	chr1.CM0284.240.nc	21.4	17.1
	CHS13[a]	chr1.CM0284.250.nc	22.3	16.8
	CHS14	chr2.CM0018.1100.nc	23.8	18.8
	CHS15	chr2.CM0018.1150.nc	50.0	29.0
	CHS16	chr2.CM0018.1250.nc	27.8	21.9
	CHS17	chr2.CM0018.680.nc	20.8	17.1
	CHS18	chr2.CM0018.710.nc	26.0	22.7
	CHS20	chr3.CM0590.740.nc	60.7	35.5
	CHS22	chr4.CM0044.260.nd	11.9	10.9
	CHS24	chr1.CM0591.370.nd	69.3	50.4
	CHS25	chr1.CM0591.390.nd	28.3	21.9
	CHS27	chr1.CM0591.360.nd	33.7	21.0
	CHS28	LjSGA_132696.1	37.2	27.3
	PKR1[a]	chr2.CM0191.390.nc	91.5	43.4
	PKR2[a]	chr2.CM0191.400.nc	188.6	154.3
	PKR3[a]	chr2.CM0191.410.nc	147.3	116.0
	PKR4[a]	chr2.CM0191.480.nc	101.4	40.6
	PKR5[ac]	CM1092.70.nc	91.8	113.3
	PKR6[ac]	CM1092.80.nc	101.2	118.8
	PKR7	LjSGA_019910.1	56.5	29.9
	CHI6	chr5.CM0034.130.nc	17.4	24.2
Isoflavonoid:				
	IFS1[a]	LjT24P23.50.nc	5.6	3.7
	IFS2[a]	chr4.CM0692.280.nc	53.3	20.3
	IFS4	LjSGA_087856.1	26.2	10.9
	HI 4′*O* MT	LjT24P23.70.nc	15.8	10.8
	HID1	LjSGA_020298.1	5.9	20.8
	HID2	LjB01D01.120	12.3	7.2
	CYP81E6 (I2′H)	chr4.CM0256.420	15.4	10.3
	CYP81E41 (I2′H-like)	chr2.CM0250.30	52.4	99.8
	IFR1[a]	chr2.CM0249.1070.nc	10,8	5,4
	IFR2[a]	chr2.CM0249.1080.nc	6,4	3,0
	VR1[a]	chr1.CM1255.320.nc	22.6	16.0
	VR2[a]	chr1.CM1255.310.nc	26.4	50.9
	PTR1[a]	chr3.CM0216.20.nd	17.3	20.9
	PTR2[a]	chr3.CM0216.40.nd	127.4	113.2
	PTR3[a]	chr3.CM0216.50.nd	13.6	5.9

[a] Gene names reported by Shimada et al. (2007)

[b] Presumed to be a single gene represented by 2 non-overlapping gene fragments

[c] Sequence identity too high to discriminate individual genes

Supplemental Table S3. Genes encoding phenylpropanoid biosynthetic enzymes that decreased in expression after three hours or six hours following elicitation, normalised with absolute fold change (FCA).

Pathway	Gene Name	Kazusa ID	FCA 0-3hrs	FCA 0-6 hrs
General phenylpropanoid:				
	PAL9	chr5.LjT17D03.50.	40.7	8.7
	4CL14	LjSGA_094072.1	4.6	8.3
Flavonoid:				
	CHS1[a]	chr1.CM0593.380.nc	348.8	557.7
	CHI5	LjSGA_035547.1	14.2	40.7
	FLS2	chr1.LjB17A22.120.nc	219.7	446.1
Isoflavonoid(-like):				
	HID4	LjB01D01.90	21.2	16.9
	CYP81E40	chr2.CM0250.70	5.8	6.1
	VR-like	chr1.CM0017.190	10.7	18.3

[a] Shimada et al. (2007)

Supplemental Table S4. Sequences of primers used to confirm gene expression profiles**.**

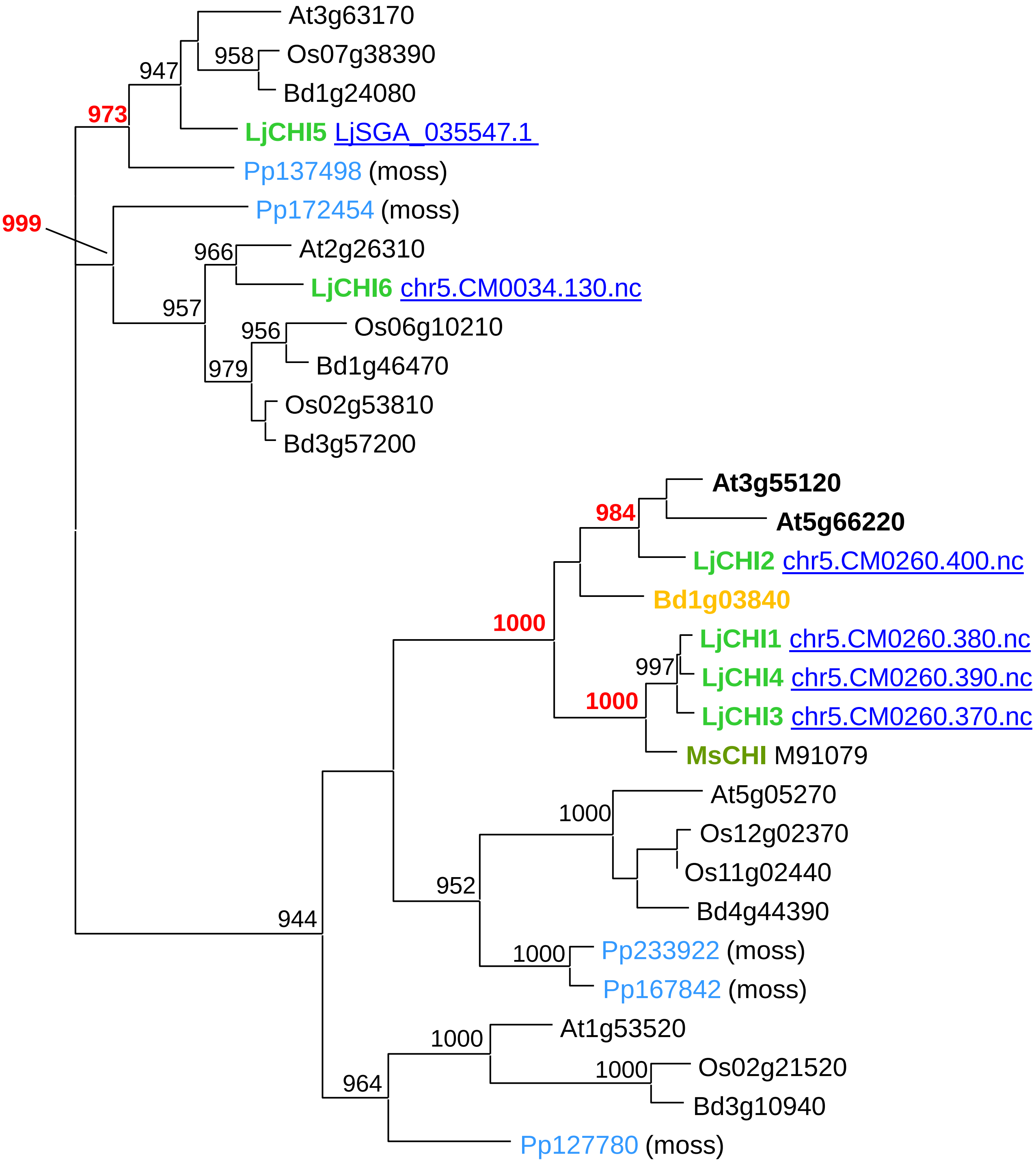
Gene	**Primer Sequence**
IFS1s	TTCTTCTTAAGATACAACACAAAC
IFS1as	TAGAGTAATGTAAATGGATATACG
IFS2s	TCATAAAAGTACAACACAAATAAC
IFS2as	GAAAAGTGAATATGATTATTAATAC
HI4′OMTs	GGTGTTGTGTGCTACAATTTC
HI4′OMTas	ATAGATTTAACTTGCACCTCTG
MYB13s	GCCTCCAAATCCAACAAAGATGCC
MYB13as	GAACCTGGATTTGTCTGACTATTGAG
MYB14s	CAGAACACCAACAACAACCAAGCTC
MYB14as	TACCCTCATTGATCATGCTTG
MYB15s	GATCTGATTCCAGTTCCAGCACAC
MYB15as	AAGATAAGATTAAAATTTGAGCCGTC
CHS1s	CTTGGAGGCTTGAACCGAG
CHS1as	ATATCACTATTCTACATTTATTGG
F3Hs	CAAACTCGACCCCTACTATTG
F3Has	GTTACTTCAAAGACCTCGTAGG
MYB3s	CAGAATCAAGAACCTGGTTCTTCTC
MYB3as	ACTAGCTATTTCTGAGAAATTGCCTC
MYB5s	GCCTCTTCCTGTGATCACAAATCAC
MYB5as	TTAATCAACATGCACTAGATATATATGC
MYB90s	TAAATGCTCAAGAAGCTGAAGATCAC
MYB90as	GGCTTTCTCTACCGTAGCAGGATC
MYB134s	CATCCGAAGGAAGCTAATAAAGATGG
MYB134as	ATCCAAAGGGAACACTACCAAGAAG
UBQs	CTACAACATTCAGAAGGAGTCCA
UBQas	GCACACACAGACAATCCTTCCAAC
LjMyb14s	CCATAACTCAATCATCAGACCC
LjMyb14as	AGAAACTCTCATCAATTTCAGG
LjHIDs	CTGTTTTGGAGTTCATACCCTG
LjHIDas	CAGATTCCTCTGTCTCTCAGTTC
LjVRs	TGGAAGTTGTGACTCTGATTCC
LjI2′Hs	GGGCTAACAACCCTGGAG
LjI2′Has	CATTGCCACACTTCTTCGTC
LjVRas	GACAATGAATGGGGAGCAGT
LjPTRs	TATCTCTACTGTGGGTCATGTACTGC
LjPTRas	GACCTTCACCTGGTATGTGCTC
LjActins	ATCATACCTTCTATAACGAGCTTC
LjActinas	GTGGCTGACACCATCACCAGAATC

Supplemental Table S5. Raw Intensity Signal for isoflavonoid-related genes with increased expression levels in all three independent transgenic lines constitutively expressing *LjMyb14* compared to wild-type MG20. Each gene was represented by two unique eArray probes and the Raw Intensity values are listed separately for each probe.

Pathway Or Enzyme Class	**Gene Name**	**Kazusa Identifier**	**eArray Probe**	**Raw Intensity Wild-type**	**Raw Intensity Line LjMYB14-1**	**Raw Intensity Line LjMYB14-2**	**Raw Intensity Line LjMYB14-3**
General phenylpropanoid:							
	PAL4	chr1.CM0033.830	CUST_510_PI399407686	263.08	620.09	565.57	577.76
			CUST_509_PI399407686	7727.23	19960.97	16912.80	21488.67
	PAL6	chr1.CM0033.960	CUST_123_PI399407686	14366.51	34648.23	30219.22	39488.58
			CUST_124_PI399407686	5330.25	14273.21	12195.87	13802.19
	CYP73A99 (b)[*] (C4H3)	LjSGA_082453.1	CUST_58_PI398500264	58.52	558.56	712.89	419.97
			CUST_57_PI398500264	312.48	2721.03	3049.95	3223.08
	CYP73A99(a)[*] (C4H4)	LjSGA_029001.1	CUST_451_PI398500264	51.53	374.37	487.24	216.79
			CUST_452_PI398500264	20.28	166.26	214.68	88.26
	CYP73A98 (C4H5)	LjSGA_031602.1	CUST_699_PI398500264	5584.75	18423.06	18351.91	15118.04
			CUST_698_PI398500264	6459.34	20151.57	19897.47	20806.50
	4CL1	chr2.CM0018.1240	CUST_396_PI399407686	517.74	1497.57	1495.99	1504.88
			CUST_395_PI399407686	560.24	1513.33	1502.94	1604.09
	4CL16	chr6.CM0057.160	CUST_419_PI399407686	490.13	1410.52	1330.17	1470.32
			CUST_420_PI399407686	529.60	1506.98	1420.77	1538.94
	4CL20	LjSGA_016012.1	CUST_420_PI399407686	529.60	1506.98	1420.77	1538.94
			CUST_325_PI399407686	624.48	1801.97	2008.03	1731.23
Isoflavonoid:							
	PKR5	CM1092.70	CUST_311_PI399407686	372.00	3136.56	3537.05	4667.22
			CUST_312_PI399407686	263.93	2267.34	2533.58	3308.96
	PKR6	CM1092.80	CUST_78_PI399407686	248.20	2179.81	2397.24	3153.81
			CUST_77_PI399407686	383.20	3245.40	3702.96	4741.98
	HI4′ OMT	LjT24P23.70	CUST_431_PI399407686	2032.47	5092.47	4669.24	6469.47
			CUST_432_PI399407686	1132.56	2777.39	2738.68	3069.04
	HID2	LjB01D01.120	CUST_137_PI399407686	482.32	1531.75	1686.63	1897.41
			CUST_138_PI399407686	580.57	1852.57	2102.53	2570.44
	CYP81E40 (I2′H)	chr2.CM0250.70	CUST_504_PI398500264	34.23	171.16	178.04	111.30
			CUST_503_PI398500264	74.37	322.53	327.65	240.17
	VR	chr1.CM1255.320	CUST_591_PI399407686	301.28	1836.06	2103.42	1887.81
			CUST_592_PI399407686	209.34	1447.14	1730.80	1429.27
	PTR3 (AB265591)	chr3.CM0216.50	CUST_85_PI399407686	477.19	1595.92	1512.80	1629.25
			CUST_86_PI399407686	237.44	765.71	723.39	788.19
UDP-glycosyltransferases:							
	UGT71B-l.ike	LjSGA_025093.1	CUST_148_PI399390334	16.13	147.13	151.33	120.30
			CUST_147_PI399390334	29.14	184.65	181.19	160.29
	UGT71B-l.ike	LjSGA_066183.1	CUST_18_PI399390334	197.78	214.80	207.29	33.86
			CUST_19_PI399390334	7.19	12.41	18.54	5.66
	UGT71B-l.ike	LjSGA_084705.1	CUST_3_PI399390334	25.49	161.64	126.85	111.23
			CUST_4_PI399390334	18.63	83.12	77.84	79.17
	UGT71B-l.ike	LjSGA_128268.1	CUST_242_PI399390334	43.61	282.24	245.24	249.62
			CUST_243_PI399390334	8.41	97.77	97.42	82.48
	UGT72B4	chr1.CM0579.30	CUST_265_PI399390334	2083.28	5607.01	7132.80	6070.81
			CUST_264_PI399390334	1927.91	5386.74	6789.96	5939.55
	UGT72B5	chr3.LjT41F16	CUST_170_PI399390334	3.83	46.65	43.03	46.53
			CUST_169_PI399390334	14.81	79.86	86.24	66.00
	UGT73B-like	chr4.CM0739.130	CUST_315_PI399390334	9.57	106.13	43.16	37.80
			CUST_314_PI399390334	7.37	80.53	28.31	26.07
	UGT73B-like	chr6.LjT16B20.10	CUST_300_PI399390334	9836.20	20043.77	21695.49	26850.51
			CUST_301_PI399390334	12831.02	26130.18	28163.07	34182.27
	UGT79B13	chr3.LjT14J20.160	CUST_232_PI399390334	109.30	341.03	393.45	451.46
			CUST_231_PI399390334	185.65	655.21	639.63	890.66
	UGT87E1	chr3.LjT09A09.120	CUST_196_PI399390334	1.71	17.82	11.17	12.50
			CUST_195_PI399390334	6.95	23.72	20.42	24.73
	UGT90A-like	chr1.LjT14B18.140	CUST_274_PI399390334	66.91	1906.74	1817.84	2108.00
			CUST_275_PI399390334	3.59	86.07	115.76	43.46
	UGT92A-like	chr1.LjT14B18.160	CUST_174_PI399390334	17.59	460.78	468.90	458.60
			CUST_173_PI399390334	8.09	227.78	221.33	169.27
	UGT92A-like	LjT30N24.300	CUST_303_PI399390334	2.23	28.43	73.65	33.08
			CUST_302_PI399390334	14.92	267.35	535.96	775.21

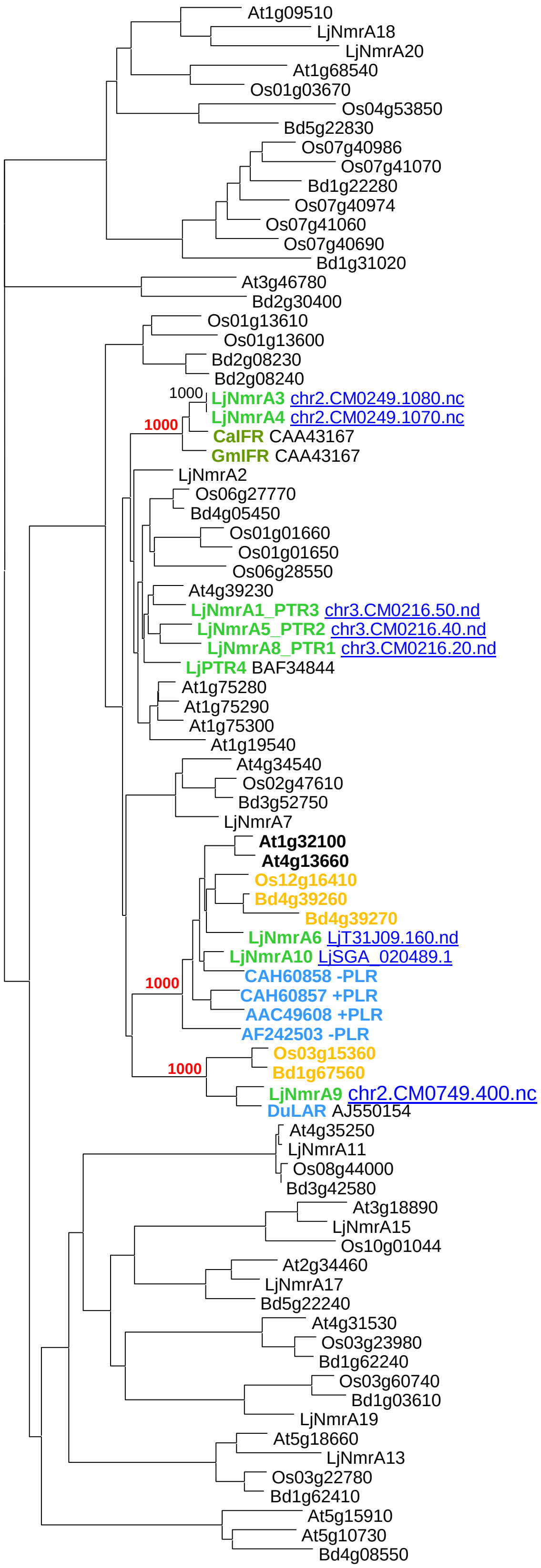
[*] Presumed to be a single gene represented by two non-overlapping gene fragments

Figure S1



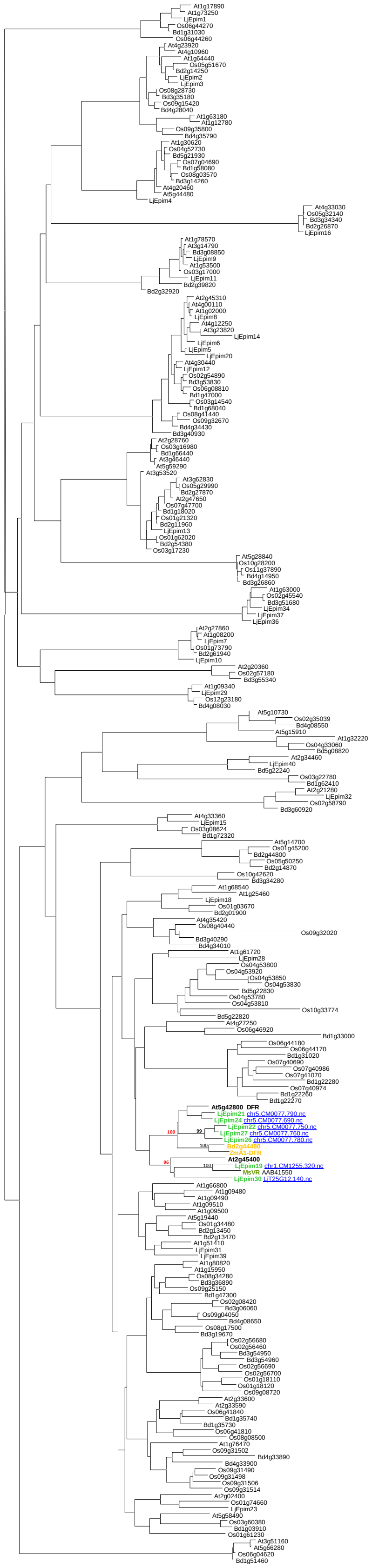
Supplemental Figure S1. Phylogenetic tree of CHI-like proteins. In Lotus, the proteins involved in isoflavonoid biosynthesis are LjCHI1, LjCHI3 and LjCHI4

Figure S2



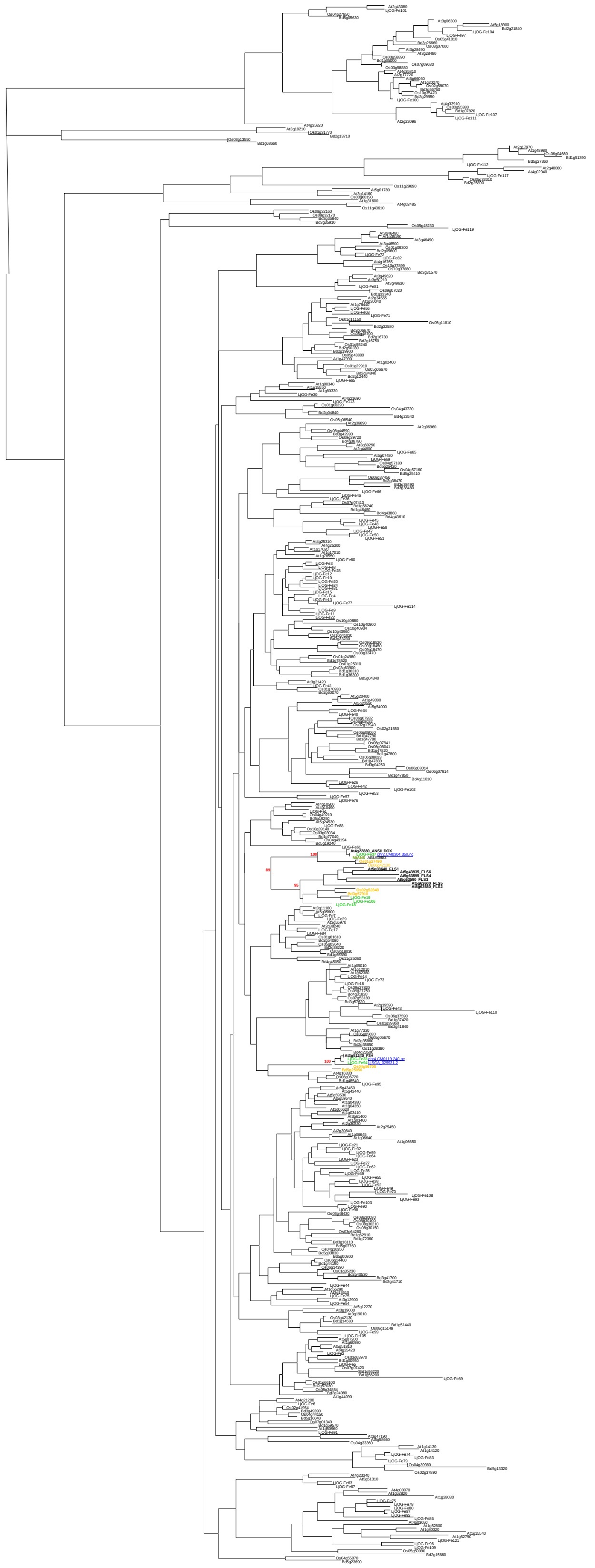
Supplemental Figure S2. Phylogenetic tree of the Nmr family containing proteins that encode enzymes for IFR, LAR, PTR and PLR (pinoresinol-lariciresinol reductase).

Figure S3



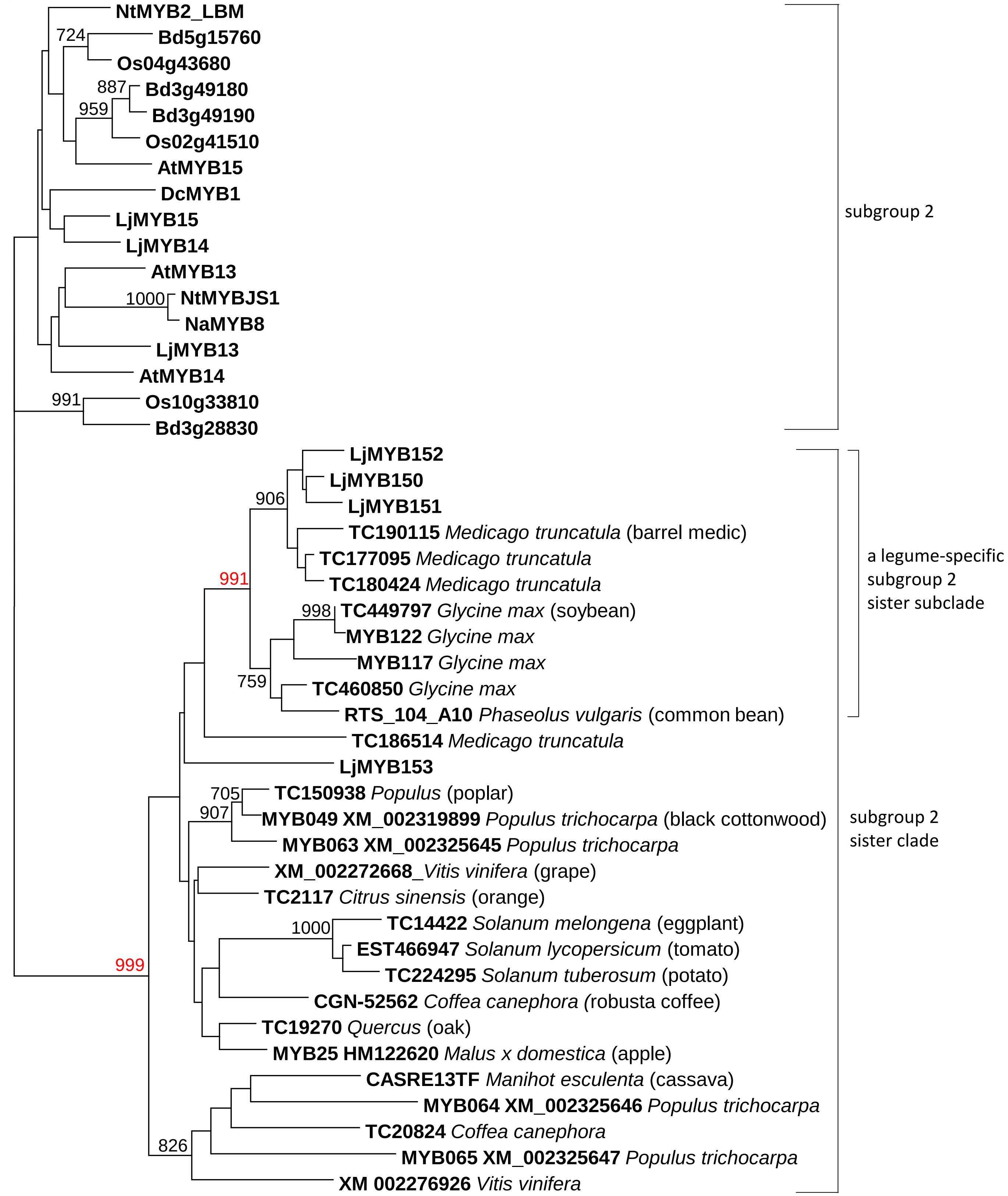
Supplemental Figure S3. Phylogenetic tree of the epimerase family containing proteins that encode enzymes for DFR and VR.

Figure S4



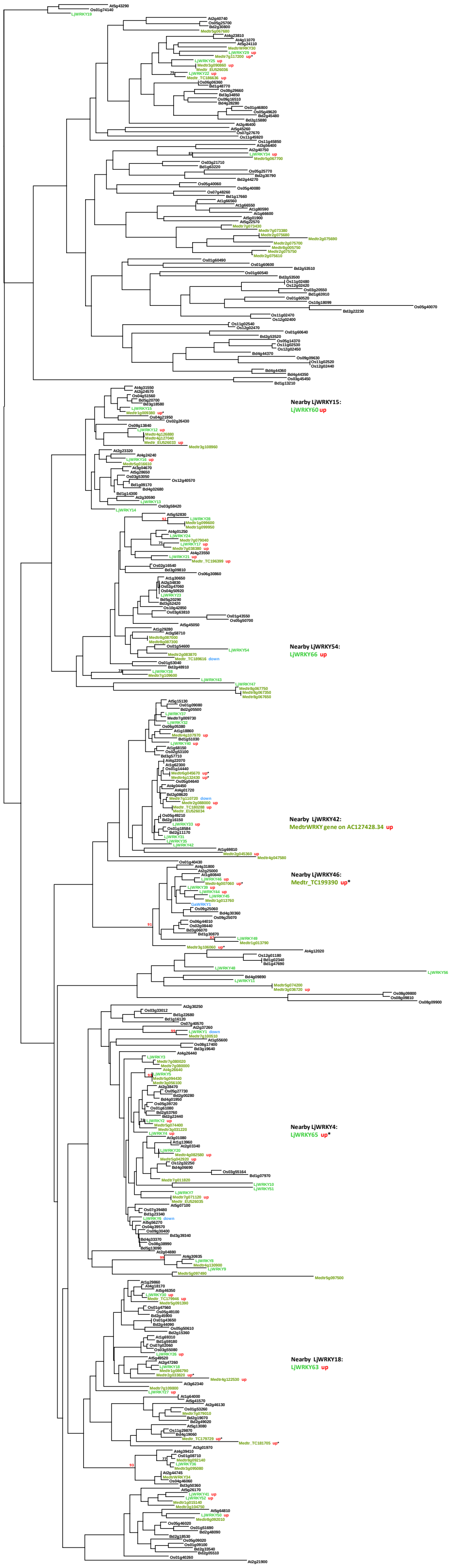
Supplemental Figure S4. Phylogenetic tree of the 2-oxoglutarate and Fe(II)-dependent oxygenase family containing proteins that encode enzymes for F3H, ANS/LDOX, and FLS.

Figure S5



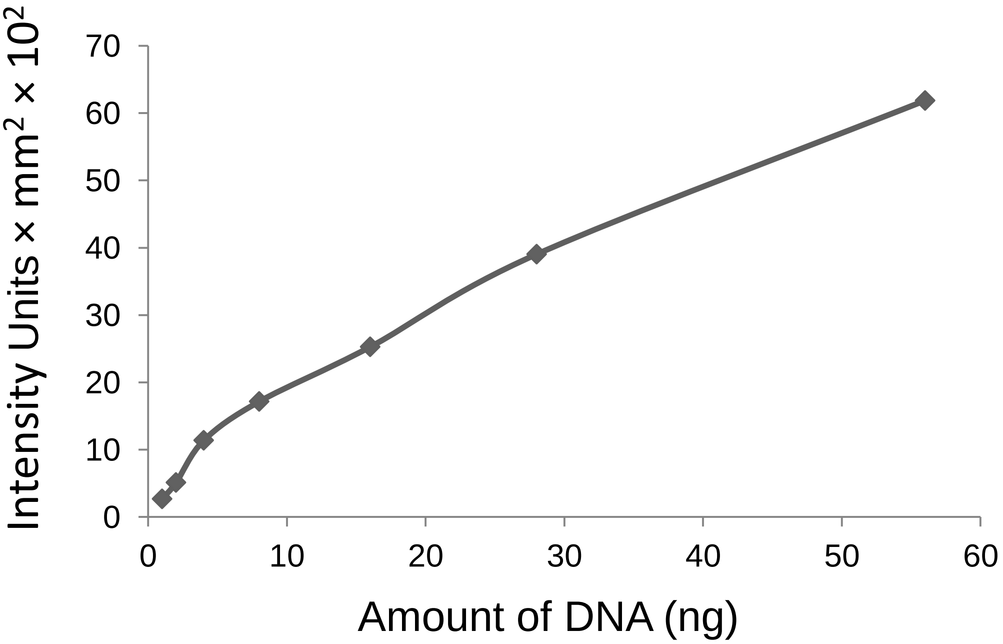
Supplemental Figure S5. Phylogenetic tree of subgroup 2 proteins from Arabidopsis, Lotus, rice and Brachypodium and a related but distinct set of proteins that were found to date (July 2011) to be present in 13 dicot species, including legume species. Bootstrap values (>70%) are indicated and are shown in red if they are >90% and help to define a clade of proteins that are distinct from the rest of the proteins in the data set. Of particular note is a subclade of legume-specific proteins within the subgroup 2 sister clade that cluster with a significant bootstrap value of 99% reflecting the difference in their DNA binding domain protein sequence compared to the rest of the subgroup 2 sister clade, including two legume genes from Medicago and Lotus that fall outside the subclade.

Figure S6



Supplemental Figure S6. Phylogenetic tree of the WRKY family showing proteins from Lotus and Medicago. Indicated are Lotus (Lj) and Medicago (Medtr) proteins with a greater than 2-fold increase (up - red) or decrease (down - blue) in expression after 3 h GSH treatment (Lotus) or 2h YE (Medicago; Naoumkina et al. 2007). Bootstrap values (>70% of 100 datasets) are indicated and are shown in red if they are statistically significant (>90%) and confirm orthologous proteins in Lotus and Medicago.

[*] genes induced in both YE-treated and cells treated with methyl jasmonate (MJ).



Supplemental Figure S7. Graph showing that SYBR-Gold nucleic acid gel stain detects DNA in a linear manner from 4 to 56 ng. This range coincided with the amount of PCR product quantified in the experiments presented in Fig.3 and Fig 4.