# NIN-Like Protein Transcription factors regulate leghemoglobin genes in legume nodules

**Authors:** Suyu Jiang[1], Marie-Françoise Jardinaud[2], Jinpeng Gao[1], Yann Pecrix[2], Jiangqi Wen[3], Kirankumar Mysore[3], Ping Xu[4], Carmen Sanchez-Canizares[5], Yiting Ruan[1], Qiujiu Li[1], Meijun Zhu[1], Fuyu Li[1], Ertao Wang[1], Phillip S. Poole[5], Pascal Gamas[2], and Jeremy D. Murray*[1,6]

**Affiliations:**

[1]CAS-JIC Centre of Excellence for Plant and Microbial Science (CEPAMS), Centre for Excellence in Molecular Plant Sciences (CEMPS), Shanghai Institute of Plant Physiology and Ecology (SIPPE), Chinese Academy of Sciences, China.

[2]LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France.

[3]Noble Research Institute, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

[4]Shanghai Engineering Research Center of Plant Germplasm Resource, College of Life Sciences, Shanghai Normal University, Shanghai 200234, China

[5] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK.

[6]John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK

*Correspondence to: jeremy.murray@jic.ac.uk.

Leghemoglobins enable endosymbiotic $N_2$-fixation in legume nodules by channelling $O_2$ for bacterial respiration while maintaining a microoxic environment to protect $O_2$-sensitive nitrogenase. We show that the NIN-like Protein (NLP) transcription factors NLP2 and Nodule Inception (NIN) directly activate expression of leghemoglobins through a promoter motif resembling a 'double' version of the Nitrate Responsive Elements (NREs) targeted by other NLPs that has conserved orientation and position across legumes. CRISPR-KO of the NRE-like element resulted in strongly decreased expression of the associated leghemoglobin. Our findings indicate that the origins of the NLP-leghemoglobin module for $O_2$-buffering in nodules can be traced to an ancient pairing of NLPs with non-symbiotic hemoglobins that function in hypoxia.

**One Sentence Summary:** NLP regulation of leghemoglobins in N-fixing nodules descended from of an ancient NLP-hemoglobin regulatory module.

Most legume species can form an endosymbiosis with rhizobia bacteria, forming highly specialized root nodules containing cells in which rhizobial nitrogenase reduces dinitrogen to ammonia. Nitrogenase is rapidly and irreversibly deactivated by oxygen, while $N_2$-fixation is energy intensive, so the rate of $N_2$-fixation is a compromise between the $O_2$ required to support the energy demands of the symbionts and the poisoning of nitrogenase by $O_2$, the so-called "oxygen paradox" (1). To address this constraint, up to 40% of nodule protein is comprised of leghemoglobins (LgHbs), which contain an $O_2$-binding heme group that colors the nodule red. LgHbs feature an extremely fast $O_2$-association rate and a relatively slow $O_2$-dissociation rate, allowing them to buffer the free oxygen concentration at around 10 nM, and thereby maintain $O_2$ levels compatible with $N_2$-fixation. Notwithstanding the central role of LgHbs in $N_2$-fixation, their transcriptional regulation and the manner of their recruitment to nodulation have not been elucidated.

The founding member of the NIN-like Protein family, Nodule Inception (NIN), is required for infection by rhizobia as well as for nodule formation (*2*). In contrast, *Medicago truncatula* NLP1 and *Lotus japonicus* NLP4 are required for nitrate-suppression of nodulation (*3, 4*). *NIN* and *NLP2*, however, are the only NLPs expressed in medicago nodules (**Fig. 1A**; **fig. S1A**). The expression of *NLP2* and several *LgHbs* was strongly reduced in nodules induced by rhizobia mutants *bacA* and *fixJ* where $N_2$-fixation is defective due to lack of symbiosomes or nitrogenase activity, respectively, indicating a tight connection of *NLP2* with $N_2$-fixation (**Fig. 1B**).

To investigate NLP2's role, we evaluated nodulation on *NLP2*-knockout roots using an aeroponic system. The resultant nodules had normal zonation but were smaller and had reduced $N_2$-fixation potential (**Fig. 1C, D**). We then isolated two mutants, *nlp2-1* and *nlp2-2* (**fig. S1B**). The *nlp2-1* mutant had smaller nodules in aeroponics, matching the CRISPR knockout phenotype (**Fig. 1E)**, and had a greatly reduced plant biomass, but developed a normal number of nodules (**Fig. 1F)**. However, when grown in substrate the mutant developed normal sized nodules (**fig. S1C**). The increase in phenotype severity seen in aeroponics vs substrate conditions may reflect the much larger and more extensively nodulated root system that develops under aeroponics, which may exaggerate the N-fixation defect in *nlp2*. Nevertheless, the N-fixation potential of both *nlp2-1* and *nlp2-2* was also compromised, albeit to a lesser degree, when grown in substrate (**fig. S1D**), and could be partially complemented by expression of *NLP2* from its native promoter, but not by a mutant version of *NLP2* (**fig. S1E, F**). Furthermore, $^{15}N_2$ assimilation was compromised in *nlp2-1* plants nodulated in substrate, (**Fig. 2A)**, which corresponded to a decrease in N-content in nodules and shoot, while %N was unchanged (**Fig 2A; fig. S2D**). Evaluation of *nlp2* mutants under inhibitory nitrate regimes indicate that loss of NLP2 does not affect N-suppression of nodulation (**fig. S2A to C**). Together these results suggest a positive role for NLP2 in $N_2$-fixation.

In addition to reduced N-fixation, *nlp2* nodules appeared less pink (**fig. S2E**), which was reflected in lower LgHb expression and heme levels (**Fig. 2B, C; fig. S2F**). We then profiled gene expression of *nlp2-1* nodules under different N-regimes using substrate where the mutant forms normal sized nodules, rather than using aeroponics where the smaller nodules on the mutant would confound the analysis. Results revealed nitrate-dependent effects on gene expression (**Fig. 2D**; full data set **Auxiliary File 1**). Genes encoding LgHbs, heme biosynthesis enzymes, and nitrate uptake transporters were strongly downregulated in *nlp2-1* (**Fig. 2E**, **F**; **Supplemental Table S1**) while jasmonate-related genes were upregulated (**Supplemental Table S2**) as previously reported in root hairs of *M. truncatula nin* (*5*). Expression of *NLP2* from its own promoter partially rescued gene expression in *nlp2*, but a mutant version did not (**Fig. 2G**). *In silico* promoter analysis revealed that 28 of the 55 genes with the largest overall decrease in expression in *nlp2* nodules contained Nitrate Responsive Elements (NREs), which are required for transcriptional activation by NLPs (*6*), including known genes required for N-fixation like *SymCRK*, *SEN1*, and Nodule Cysteine-Rich (NCR) secreted peptides (*7*, *8*; **Supplemental Table S3**). In addition, a larger motif consisting of two partly overlapping NREs was detected in the *LgHb* genes. We designated these as 'double-NREs' (dNREs), in which the second half of the first semi-palindrome (NRE1) overlaps with the first half of the second element (NRE2) (**Fig. 3A**). The dNREs all have the same orientation and are positioned between -100 and -300 bp upstream of the TSS (**Fig. 3B**). dNREs were also detected at same position and orientation in promoters of *LgHbs* of soybean and common bean, indicating their conservation across papilionoid legumes (**fig. S3A**). Examination of dNREs in promoters of soybean and medicago *LgHbs* suggests that those strongly differing from the consensus motif have lower expression in nodules (**fig. S3B, C**), suggesting the direct involvement of dNREs in the regulation of *LgHb* expression. Our transcriptomic data verified that *NLP2* and *NIN* are the primary *NLPs* expressed in medicago nodules (**fig. S1A**, **fig. S4)**, and since NIN binding sites strongly resemble NREs (*9*, *10*; **fig. S5**) we speculated that both NLP2 and NIN may directly promote *LgHb* expression. Using an EMSA we found that truncated NLP2 and NIN proteins containing their RWP-RK DNA binding domains could bind to dNRE-containing DNA fragments but not to those with mutated dNREs (**Fig. 3C**, **D;** details in **fig. S6, S7**). To test its importance for *LgHb* expression we mutated the dNRE element in the *LgHb1* promoter using CRISPR/CAS9. This resulted in strongly decreased expression of *LgHb1*, but not *LgHb2*, whose promoter contained the gRNA target sequence but lacked the PAM (**Fig. 3E; fig. S8**). When expressed in *Nicotiana benthamiana* leaves, both NIN and NLP2 activated reporter activity when driven by a dNRE-containing fragment of the *LgHb7* promoter, but not by a dNRE-deleted version (**Fig. 3F**). Versions of NIN and NLP2 with mutations in their RWP-RK domains were unable to activate the reporter (**fig. S9A**). Co-infiltration of *NIN* and *NLP* constructs resulted in higher reporter gene activity than either construct alone (**fig. S9B**). These results support a role for NLP2 and NIN in activation of *LgHb* expression in nodules. Since *nin* loss-of-function mutants do not form nodules, we used a CRISPR/CAS9-based strategy to selectively knockout *NIN* in a tissue-specific manner by expressing *CAS9* in the nodule interzone/nitrogen fixation zone of WT and *nlp2-1* nodules. Successfully edited nodules showed decreased *LgHb* expression suggesting that NLP2 and NIN act redundantly to regulate the transcription of nodule *LgHbs* (**fig. S10**).

We then investigated if the above regulation was a legume innovation or predated the symbiosis. Phylogenetic analysis showed that close homologs of NLP2 are found only in legumes, and group separately from their next closest homologs, most of which are expressed in seeds (**fig. S11, S12)**, while NIN was reported to be present in legumes and nodulating non-legumes (*11*). The absence of NLP2 in *Prosopis alba,* a Caesalpiniod legume, suggests that NLP2 arose before the divergence of Papilionoid legumes. LgHbs evolved from non-symbiotic hemoglobins (nsHbs) involved in energy maintenance under hypoxia and $O_2$ transport (*12, 13*), with class I nsHbs acting in the nsHB-Nitric Oxide cycle (*14-15*). Analysis of *nsHb* promoter sequences across angiosperms revealed that they frequently harbor NRE elements, suggesting they may be NLP targets (**fig. S13A, B**). Mining of data from a genome-wide ChIP analysis revealed binding of *L. japonicus* NIN to a class I *nsHB* (*16*). In addition, *nsHB1* was identified as a direct target of AtNLP7 in an Arabidopsis DAPseq study (*17*), while its expression was shown to be dependent on NLP7 or NLP8 in three transcriptomic studies (*18*, *19, 20*). Our own investigation of medicago *nsHB1* showed its expression is dependent on NLP1, and that it may be a direct target (**fig. S14**). These findings suggest that NLPs were recruited to nodulation as part of an ancient NLP-nsHb regulatory module. The dNRE element was detected in *Hb* promoters of a mimosoid legume, and thus appears to predate NLP2, but was absent in nodulating non-legumes (**fig. S13C;** these and other promoter sequences used are provided in the **Appendix** in the Supplemental materials). It therefore seems likely that regulation of hemoglobins by NLPs served as a basis for recruitment of NIN in nodulating non-legumes, then later NLP2, to enhance LgHb expression in legume nodules.

**References:**

1. M. K.,Vanderleyden *J. Biology and Fertility of Soils* **30**, 363–373 (2000).
2. L. Schauser *et al.*, *Nature*. **402**, 191-195 (1999).
3. H. Nishida *et al., Nature Commun.* **9,** 499 (2018).
4. J. S. Lin *et al.*, *Nature Plants* **4**, 942-952. (2018).
5. C. W. Liu *et al. Plant Physiol.* **179**, 1704-1722 (2019).
6. M. Konishi *et al., Nature Commun.* **4,** 1617 (2013).

7. F. Berrabah *et al., New Phytol.* **203**, 1305-1314 (2014).

8. H. Kouchi *et al., Plant and Cell Physiol.* **53**, 225-236 (2011).

9. H. Nishida *et al., Plant Cell* https://doi.org/10.1093/plcell/koab103 (2021).

10. T. Soyano *et al., Plant Cell Physiol.* **56**, 368-76 (2015).

11. M. Griesmann *et al., Science*, **361**(6398):eaat1743. doi: 10.1126/science.aat1743. (2018).

12. C. Vázquez-Limón *et al., Plant Sci.* **191**, 71-81 (2012).

13. H. Vigeolas *et al., Plant Physiol.* **155**, 1435-1444 (2011).

14. M. Perazzolli *et al.*, *Plant Cell* **16,** 2785–2794 (2004).

15. A. U. Igamberdiev *et al., J. Exp. Bot.* **55**, 2473-82 (2004).

16. T. Soyano *et al., PNAS*, **111**, 14607-14612 (2014).

17. R.C. O'Malley *et al., Cell* **165**, 1280-1292 (2016).

18. L. Zhao *et al. Sci. Rep.* **8**, 1487 (2018).

19. J.M. Alvarez *et al., Nat. Commun.* **11**, 1157 (2020).

20. D. Yan *et al. Nat. Commun.* **7**, 13179 (2016).

21. B. Roux *et al.*, *Plant J.* **77**, 817-37 (2014).

22. D. Barker *et al.*, *The Medicago Handbook* (2006).

23. S. Marillonnet *et al.*, *Curr. Protoc. Mol. Biol.* **130**, e115 (2020).

24. K. Xie *et al.*, *PNAS* **112**, 3570 (2015).

25. H.L. Xing *et al.*, *BMC Plant Biol.* **14**, 327 (2014).

26. M. Padidam *et al.*, *Biotechniques* **31**, 328 (2001).

27. H. Liu *et al.*, *Mol. Plant* **10**, 530 (2017).

28. M. M. Bradford *Anal. Biochem*. **72**, 248 (1976).

29. T. A. Larue *et al.*, *Anal Biochem.* **92**, 11 (1979).

30. K. J. Livak *et al.*, *Methods* **25**, 402 (2001).

31. E. M. Lodwig *et al.*, *Nature* **422**, 722 (2003).

32. V. A. Benedito *et al.*, *Plant J.* **55**, 504 (2008).

33. D. Winter, *et al.*, *PLoS One* **2**, e718 (2007).

34. R. C. O'Malley *et al.*, *Cell* **165**, 1280 (2016).

**Acknowledgments:** The authors acknowledge Xie Fang for providing purified NIN protein and *nlp1* mutant seeds, and Jian Feng for providing seeds of *nlp2-1*.We thank Dian Guan and Fran Robson for their contributions early in the project. We thank Giles Oldroyd (University of Cambridge, UK) and the ENSA program for providing plasmid backbones for Golden Gate cloning and *M. truncatula* root transformation.

**Funding:** This research was supported by the CAS Project for Young Scientists in Basic Research (YSBR-011), the Ministry of Science and Technology (2019FA0904703), Strategic Priority Research Program of the Chinese Academy of Sciences (XDB27040209), the National Key R&D Program of China (2016YFA0500500), the Chinese Academy of Science (153D31KYSB20160074), and the Biotechnology and Biological Sciences Research Council (BBSRC; grant nos. BB/L010305/1[David Phillips Fellowship]; BB/J004553/1) to JM, BB/N003608/1 to CSC, and by a grant to SJ from the China Postdoctoral Science Foundation (2018M642099). PG, MFJ and YP were supported by ANR grants EPISYM (ANR-15-CE20-0002) and Laboratoire d'Excellence (LABEX) TULIP (ANR-10-LABX-41).

**Author contributions:** The experiments were conducted by SJ, M-FJ, J-PG, YP, JW, FL, CSC, QL, YT, and MZ. KM and JW generated the *Tnt1* insertion mutants. Experiments were designed by SJ, M-FJ, J-PG, PX, PG and JDM. PX, PSP, EW, PG, and JDM wrote the manuscript.

**Competing interests:** Authors declare no competing interests.

**Data and materials availability:** All data is available in the main text or the supplementary materials.

**Supplementary Materials:**
Materials and Methods
Figures S1-S14
Tables S1-S5
References (*22-35*)

**Fig. 1. Expression of *NLP2* in *M. truncatula* nodules and *nlp2* nodulation phenotypes.**

**(A)** Nodule expression of *NLP2*. Histochemically stained nodule cross-section (15 dpi with *S. meliloti* Rm2011) expressing *pMtNLP2:GUS*. Bar, 100 µM. (**B**) Expression (qRT-PCR) of *NLP2*, *NIN, LgHb2, LgHb7,* and *LgHb8* relative to reference genes in nodules of WT plants inoculated with *S. meliloti* Rm2011 grown in attapulgite substrate with no nitrogen. n=3. **(C)** Representative DAPI-stained 70 µm nodule sections of *NLP2* knockout (*NLP2*-KO) and control (EV) roots grown using aeroponics with no nitrogen. Bar, 100 µM. Nodules were evaluated 11 dpi with *S. meliloti* Rm2011 carrying a *pNifH:GFP* reporter. **(D)** ARA of nodules of *NLP2* CRISPR-edited roots grown aeroponically without nitrogen. WT *M. truncatula* seedlings (A17) were transformed with a gRNA targeting *NLP2* and roots were genotyped. The mean number of nodules from edited and non-edited roots were compared (n=19) with a Kruskall non-parametric test, α=0.05. **(E)** Nodules of *nlp2-1* and WT nodules grown aeroponically without nitrogen; 18 dpi with *S. meliloti* Rm2011. (**F**) Biomass of inoculated, 18 dpi with *S. meliloti* Rm2011 (nod), and uninoculated plants. Right side: nodule number. Plants were grown aeroponically without nitrogen. n=10 for WT, 11 for *nlp2-1*. In (B, E and F), bars indicate ±SD. In (B and F), means were compared with Student's *t*-test α=0.05.

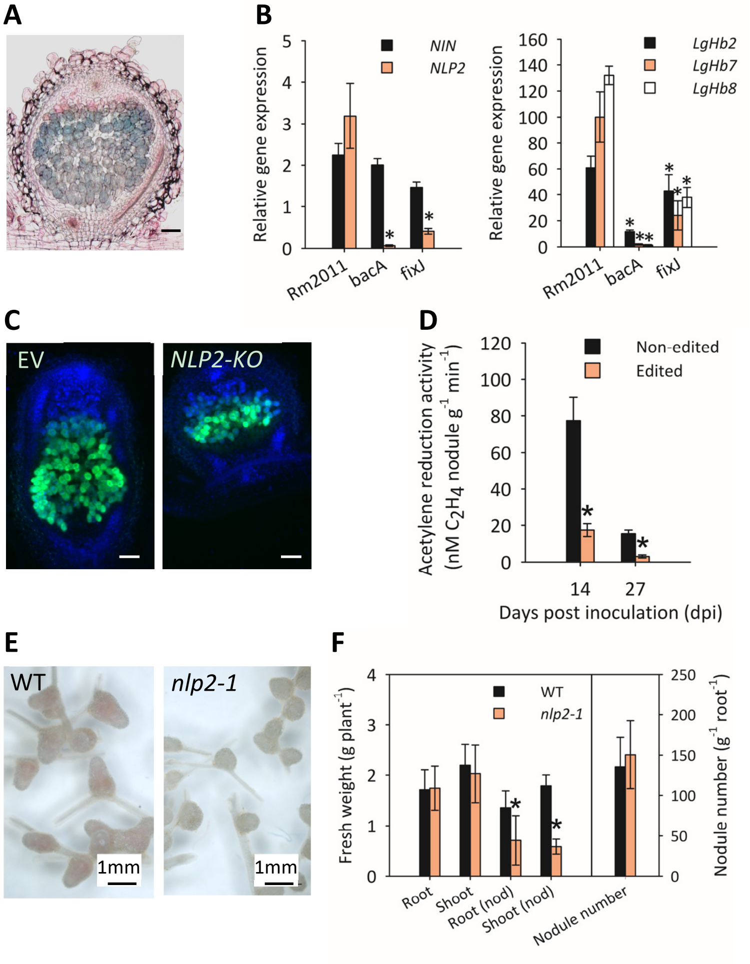
**Fig. 2. Loss of NLP2 reduces N-fixation, plant N-content, LgHb and heme levels.**

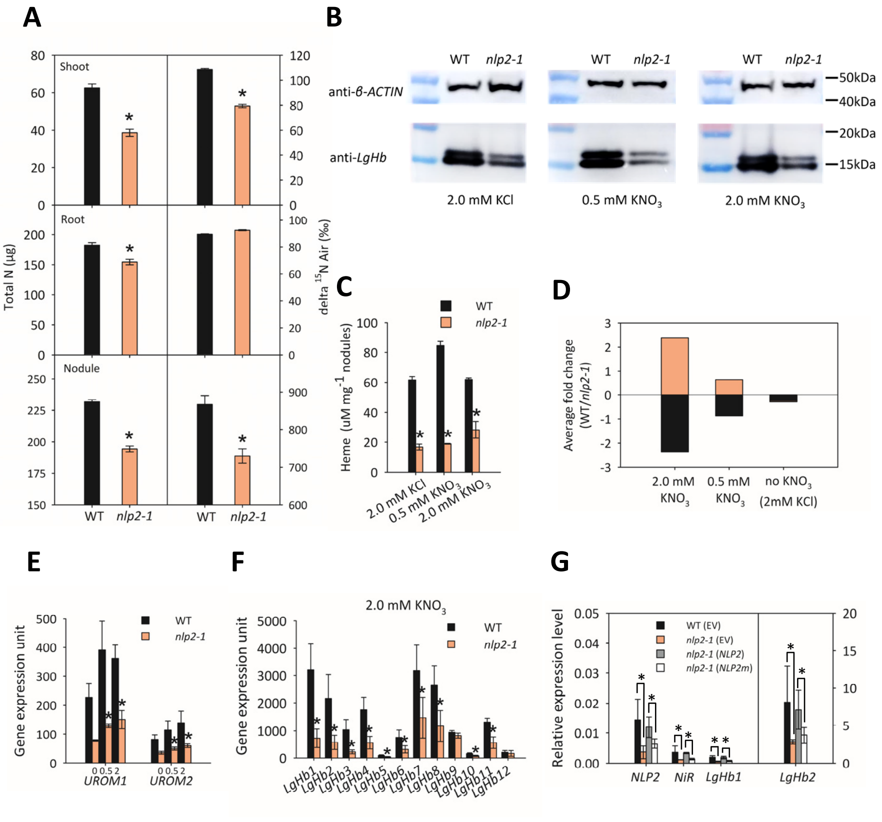
**(A)** Total N (Kjeldahl's) and $N_2$-fixation [d$^{15}$N (‰)] in *nlp2-1* and WT nodules. Plants were nodulated under 0.5 mM $KNO_3$. **(B)** LgHb content of WT and *nlp2-1* nodules. Western blot using α-LgHb antibodies to detect LgHb levels in nodules of WT and *nlp2-1* under different nitrogen regimes. The two bands reflect the presence of different sized LgHb isoforms. **(C)** Heme content of *nlp2-1* and WT nodules under different N-regimes. **(D)** Average fold changes for genes with changed expression in *nlp2-1* nodules at 2.0 mM $KNO_3$ (656 decreased, 915 increased) at different N-treatments. **(E)** Relative expression of *UROM1/2* in *nlp2-1* and WT nodules under 2.0 mM KCl (0), 0.5 mM $KNO_3$ (0.5), and 2.0 mM $KNO_3$ (2.0). **(F)** Expression of *LgHbs* in nodules in WT and *nlp2-1* under 2.0 mM $KNO_3$. **(G)** Complementation of gene expression in *nlp2-1* nodules under 2.0 mM $KNO_3$. Relative expression of *NLP2, LgHb1, LgHb2*, and *NiR* in nodules of the *nlp2-1* transformed with *pNLP2-NLP2, pNLP2-NLP2m* or EV (empty vector) using *A. rhizogenes* hairy-root transformation. *EF* was used as a reference. For (A), n=10, for (C, E, F and G), n=3. All plants were evaluated at 21 dpi with *S. meliloti* Rm2011. Means were compared with a Student's *t*-test, α=0.05. For (E and F) bars indicate 95% CI. For (A, C, and G) bars indicate ±SD.

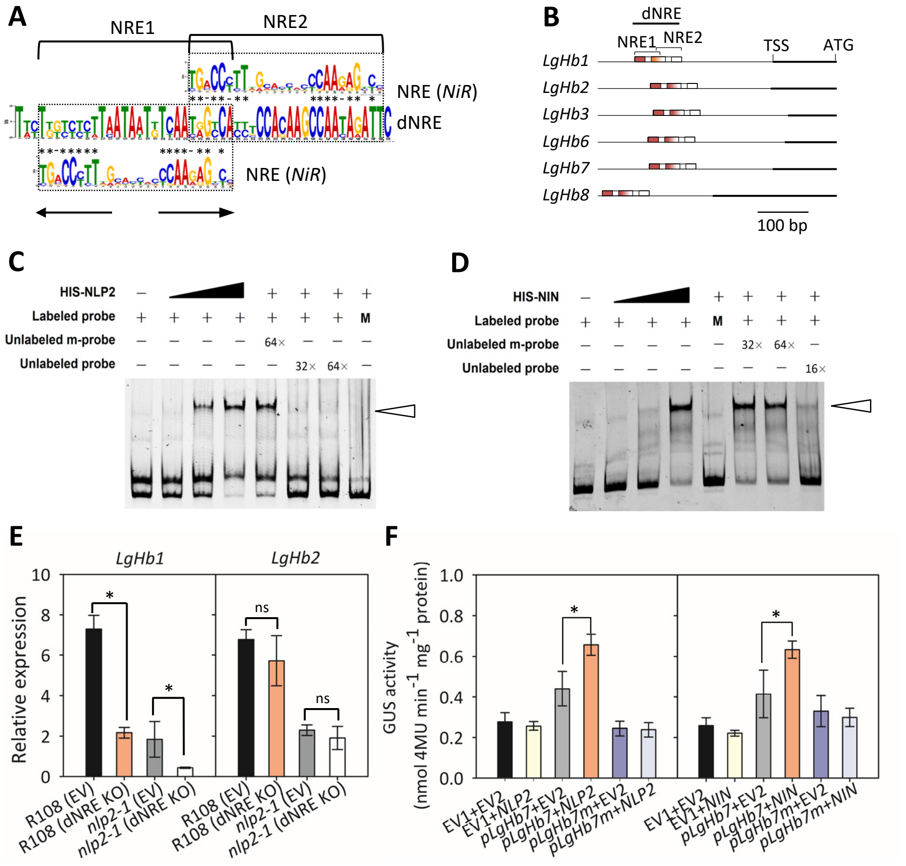
**Fig. 3. NLP2 directly activates *LgHb* expression via NRE-like promoter elements.**

(**A**) 'dNRE' motif detected in promoters of *M. truncatula nlp2-1* downregulated genes compared with NREs detected in *NiRs* across angiosperm plants (this study). The gapped semi-palindrome is indicated by arrows. The overlapping NREs are indicated by brackets. The matches of the NRE consensus to the dNRE are indicated by asterisks, and mismatches by hyphens. **(B)** The position of the dNRE sequence in promoters of *M. truncatula LgHbs* (NRE1 in orange, NRE2 in white).

The bold line indicates the 5'UTR. TSS: transcription start site, ATG: start codon. **(C)** EMSA to test for interaction between a truncated NLP2 protein containing the RWP-RK domain with an N-terminal His-tag and the CY5-labeled *LgHb7* promoter fragment containing the dNRE. M: *LgHb7* promoter with a mutated dNRE. **(D)** EMSA to test for interaction between a truncated NIN protein containing the RWP-RK domain with an N-terminal HIS tag and the CY5-labeled *LgHb7* promoter fragment containing the dNRE. M: *LgHb7* promoter with deleted dNRE. (**E**) Relative expression (qRT-PCR) of *LgHb1 and LgHb2* in nodules of *M. truncatula* plants transformed with a CRISPR construct with two gRNAs targeting the dNRE of *LgHb1*. gRNA1 also matches the dNRE sequence in the *LgHb2* promoter but there is no corresponding PAM site. The vector was introduced by hairy-root transformation and gene expression was evaluated at 21 dpi with *S. meliloti* Rm2011. *EF* was used as a reference gene. **(F)** Trans-activation assay for NLP2 and the *LgHb7* promoter driving *GUS* in *Agrobacterium tumefaciens* co-infiltrated *Nicotiana benthamiana* leaves. EV1: empty vector no-promoter control, EV2: empty vector no-*NLP2* control. *pLgHb7*:*LgHb7* promoter fragment (-342 to -1 bp) containing the dNRE, *LgHb7m*: *LgHb7* promoter fragment with the dNRE deleted. *NIN* and *NLP2*: p35S promoter driving full-length *NLP2* or *NIN*. For (E) and (F) means are shown ±SD, and compared with Student's *t*-test, α=0.05. For (E), n=3. For (F) each replicate consisted of six leaves. For NLP2 experiments (left side) n=3-4, for NIN experiments (right side) n=6-8.


A
B
Relative gene expression
5
4
3
2
1
0
NIN
NLP2
Rm2011
bacA
fixJ
*
*
Relative gene expression
160
140
120
100
80
60
40
20
0
LgHb2
LgHb7
LgHb8
Rm2011
bacA
fixJ
C
EV
NLP2-KO
D
Acetylene reduction activity
(nM C2H4 nodule g-1 min-1)
120
100
80
60
40
20
0
Non-edited
Edited
14
27
Days post inoculation (dpi)
E
WT
nlp2-1
1mm
1mm
F
Fresh weight (g plant-1)
4
3
2
1
0
WT
nlp2-1
250
200
150
100
50
0
Nodule number (g-1 root-1)
Root
Shoot
Root (nod)
Shoot (nod)
Nodule number


A
Shoot
Root
Nodule
Total N (µg)
delta $^{15}$N Air (‰)
WT
nlp2-1
B
anti-β-ACTIN
anti-LgHb
50kDa
40kDa
20kDa
15kDa
2.0 mM KCl
0.5 mM KNO$_3$
2.0 mM KNO$_3$
C
Heme (uM mg$^{-1}$ nodules)
D
Average fold change (WT/nlp2-1)
2.0 mM KNO$_3$
0.5 mM KNO$_3$
no KNO$_3$ (2mM KCl)
E
Gene expression unit
UROM1
UROM2
F
2.0 mM KNO$_3$
LgHb1
LgHb2
LgHb3
LgHb4
LgHb5
LgHb6
LgHb7
LgHb8
LgHb9
LgHb10
LgHb11
LgHb12
G
Relative expression level
WT (EV)
nlp2-1 (EV)
nlp2-1 (NLP2)
nlp2-1 (NLP2m)
NLP2
NiR
LgHb1
LgHb2


A
NRE1
NRE2
NRE (NiR)
dNRE
NRE (NiR)
B
dNRE
NRE1
NRE2
TSS
ATG
LgHb1
LgHb2
LgHb3
LgHb6
LgHb7
LgHb8
100 bp
C
HIS-NLP2
Labeled probe
Unlabeled m-probe
Unlabeled probe
64×
32×
64×
M
D
HIS-NIN
Labeled probe
Unlabeled m-probe
Unlabeled probe
M
32×
64×
16×
E
LgHb1
LgHb2
Relative expression
10
8
6
4
2
0
*
*
ns
ns
R108 (EV)
R108 (dNRE KO)
nlp2-1 (EV)
nlp2-1 (dNRE KO)
F
GUS activity
(nmol 4MU min-1 mg-1 protein)
1.0
0.8
0.6
0.4
0.2
0.0
*
*
EV1+EV2
EV1+NLP2
pLgHb7+EV2
pLgHb7+NLP2
pLgHb7m+EV2
pLgHb7m+NLP2
EV1+EV2
EV1+NIN
pLgHb7+EV2
pLgHb7+NIN
pLgHb7m+EV2
pLgHb7m+NIN

Science
AAAS

# Supplementary Materials for

## NIN-Like Protein Transcription factors regulate leghemoglobin genes in legume nodules

**Authors:** Suyu Jiang[1], Marie-Françoise Jardinaud[2], Jinpeng Gao[1], Yann Pecrix[2], Jiangqi Wen[3], Kirankumar Mysore[3], Ping Xu[4], Carmen Sanchez-Canizares[5], Yiting Ruan[1], Qiujiu Li[1], Meijun Zhu[1], Fuyu Li[1], Ertao Wang[1], Phillip S. Poole[5], Pascal Gamas[2], and Jeremy D. Murray*[1,6]

**Affiliations:**

[1]CAS-JIC Centre of Excellence for Plant and Microbial Science (CEPAMS), Centre for Excellence in Molecular Plant Sciences (CEMPS), Shanghai Institute of Plant Physiology and Ecology (SIPPE), Chinese Academy of Sciences, China.

[2]LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France.

[3]Noble Research Institute, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

[4]Shanghai Engineering Research Center of Plant Germplasm Resource, College of Life Sciences, Shanghai Normal University, Shanghai 200234, China

[5] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK.

[6]John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK

*Correspondence to: jeremy.murray@jic.ac.uk.

**This PDF file includes:**

Materials and Methods
Figs. S1 to S14
Tables S1 to S5
References

## Materials and Methods

### Gene sequences

All accession numbers for genes mentioned in this paper are listed in **Table S4.**

### Plant materials

*M. truncatula* wild type R108-C3 and A17 were used in this study. Plants homozygous for a *Tnt1* transposon insertion in *NLP2* were isolated from lines NF7812 and NF8520 and were designated as *nlp2-1*, and *nlp2-2*, respectively. Genotyping primers are provided in **Table S5**. *M. truncatula NLP1 Tnt1* insertion mutants *nlp1-1* (NF12956) and *nlp1-2* (NF14709) were kindly provided by Fang Xie (CEMPS).

### Growth conditions

For plants grown in aeroponic conditions, seeds were first scarified and sterilized and then grown in caisson units where they were initially provided 5 mM $NH_4NO_3$, then grown in nitrate-free medium before being inoculated with rhizobia as previously described in The Medicago Handbook, Growing *Medicago truncatula*: choice of substrates and growth conditions (ISBN 0-9754303-1-9. http://www.noble.org/MedicagoHandbook/, 2007) (***22***). For plants grown in substrate, seeds were scarified and sterilized with NaClO, rinsed thoroughly, plated on deionized water-agar plates, and then grown inverted for 7 days in the dark at 4 °C. Seedlings were then transferred to cells in a tray (32 per tray) containing vermiculite: perlite substrate (1:1, v:v) and cultivated in a growth chamber (22 °C, 16 h/8 h day/night, ~70 % humidity) with weekly provision of Fahräeus Plant (FP) medium supplemented with the indicated concentrations of $KNO_3$. For the plants inoculated with the *fixJ* and *bacA* rhizobia mutants, the plants were grown on attapulgite with caisson medium without N, then inoculated with 200 µl (per plant) of a solution of *Sinorhizobium meliloti* Rm2011 $OD_{600nm} = 1$ and provided only $H_2O$.

### Nodulation assays

*M. truncatula* seedlings were transferred to soil, watered with distilled water, and grown for one week. One seedling was then inoculated with *S. meliloti* Rm2011 (1 mL, $OD_{600nm}=0.8$). One day after inoculation they were watered with Modified Fahräeus Plant (MFP) with different N-regimes. They were initially provided with 2 L medium per tray, and then were provided 1 L per tray every week afterwards. FP contains 0.7 mM $KH_2PO_4$, 0.5 mM $MgSO_4$, 1 mM $CaCl_2$, 50 µM Fe-EDTA, 0.8 mM $Na_2HPO_4$, 0.1 µg $L^{-1}$ $H_3BO_3$, 0.1 µg $L^{-1}$ $MnSO_4 \cdot H_2O$, 0.1 µg $L^{-1}$ $ZnSO_4 \cdot 7H_2O$, 0.1 µg $L^{-1}$ $CuSO_4 \cdot 5H_2O$, and 0.1 µg $L^{-1}$ $Na_2MoO_4$. MFP is FP with different levels of $KNO_3$ added. Four different N-regimes were used: 2 mM KCl (no nitrate control), 0.5 mM $KNO_3$, 2.0 mM $KNO_3$ and 5.0 mM $KNO_3$. The nutrient solutions were adjusted to a pH of 6.5 ± 0.05 using $H_2SO_4$ or NaOH. Nodules were harvested 21 days post inoculation (dpi).

### Acetylene reduction assays

For determination of acetylene-reduction activity (ARA), plants were either assayed 21 dpi with *S. meliloti* (Rm2011) with weekly provision of either $KNO_3$ or KCl as described above or were nodulated under 1 mM $KNO_3$. Additional treatments with 500 mLof either 1mM $KNO_3$ or 5 mM $KNO_3$ per tray were carried out and ARA was evaluated 48-h later. For the ARA assay, ≥4 biological replicates with each comprised of 3 nodulated roots were analyzed for each genotype. The nodulated roots were introduced into 20 mL glass vials sealed with rubber stoppers. After injection of 1 mL acetylene, they were subsequently incubated for 1.5~2.5 h at room temperature. Samples of 200 µL of gas from each bottle were used to measure ethylene production using a gas chromatograph (7820A, Agilent Technology, Shanghai, China) and the roots were weighed. The nodules were then excised and weighed, and the total nitrogenase activity was calculated as nanomoles of ethylene per min per gram nodule.

### Complementation of *nlp2-1*

Plasmids were constructed using either Gateway or Golden Gate techniques in combination with Gibson cloning. For the *nlp2-1* complementation analysis, fragments of the *NLP2* promoter (2874 bp, *pNLP2*) and the *NLP2* CDS were amplified using the polymerase chain reaction (PCR) from genomic DNA and cDNA of *M. truncatula*, and the products were cloned into pL0V-PU-41295 and pL0V-SC-41308, respectively, using Gibson cloning (NovoRec plus One step PCR Cloning Kit, Novoprotein, Shanghai, China). The resultant L0 vectors containing the *pNLP2* and *NLP2* CDS were recombined with a L0 containing the T-Adh terminator to create pL1V-F2-47742 (*pNLP2-NLP2*). pL1V-F2-47742 was recombined with pL1V-R1-47802 containing the *dsRed* CDS under ~1.1kb of the *Lotus japonicus Ubiquitin* promoter. The NLP2m vector with a mutation in the RWP-RK domain was constructed in the same way after PCR mutagenesis of the L0. The resultant L2 constructs (pNLP2-NLP2-Adh-F2-47742-pLjUBQ-dsRed-NOS-R1-47802-ELE2) and the mutated version were introduced into the *nlp2-1* mutant background by *Agrobacterium rhizogenes* (Arqua-1) hairy-root transformation and an empty vector was used as a control.

### CRISPR/Cas9 knockouts

For CRISPR/Cas9-mediated gene editing of *NLP2*, five guide RNAs (gRNA) were designed with CRISPR webtool (version 4.8, http://crispor.tefor.net/, 20bp-NGG-Sp Cas9, SpCas9-HF1, eSpCas9 1.1) using the INRA A17r5.0 r1.6 *M. truncatula* genome release (NLP2_gRNA_1 GAGGAAGGACAAGGTATATC, NLP2_gRNA_2 GTTTCAGAAGCTGAACTAAG, NLP2_gRNA_3 GAAACGAACCTAAAAAATAG, NLP2_gRNA_4 GAAAGCAGGCTGAGAAAAAG, NLP2_gRNA_5 GGATGCTGCAAAAAAAATAG). Plasmids required for all the following assemblies were provided by the ENSA project (Engineering Nitrogen Symbiosis for Africa) (https://www.ensa.ac.uk) and the Golden Gate cloning strategy was systematically applied (***23***). Synthetic polycistronic genes containing two gRNAs (NLP2_gRNA_1 and NLP2_gRNA_2) and three gRNAs (NLP2_gRNA_3, NLP2_gRNA_4, NLP2_gRNA_5) interspersed by tRNAs were pre-assembled by PCR using pGTR plasmid as template and Golden Gate cloning and placed under the control of *MtU6.1* (MtrunA17Chr3g0136831) and *MtU6.6* (MtrunA17Chr7g0251721) promoters respectively (***24***). The gRNAs expression modules and a *Cas9* cassette (pLjUbi:Cas9-SV40:T35S (EC15251), (EC41414), (EC15747), with SV40 NLS C-fused by PCR to the AtCAS9 sequence) were then integrated into a binary vector (EC50506) in between a kanamycin resistance cassette (p35S:KanR:TNos (EC15529)), and a *dsRed* marker cassette (pAtUbi10:DsRed:TOcs (EC15062, EC15073 and EC41432)). The engineered binary vector was introduced into *A. rhizogenes strain* (Arqua-1) that was then used to transform *M. truncatula* A17 roots. Out of 227 transformed root systems inoculated with *S. meliloti* Rm2011 and genotyped using genomic PCR and amplicon sequencing, 64 were found to be homozygous for *NLP2* gene editions.

For CRISPR/Cas9-mediated gene editing of *NIN* a tissue-specific knockout strategy was used in which *Cas9* was expressed from the promoter of *Nodule-Cysteine Rich 158* (*NCR158*), which is specifically expressed in the interzone and $N_2$-fixation zone of nodules (**fig. S10A**). The CRISPR constructs were developed by modifying the previously described pKSE401 and pCDC-T1T2 vectors (***25***). The pKSE401 neomycin resistance gene was replaced by the *dsRed* CDS using the *Nco* I and *Rsr* II restriction sites. In addition, the *CaMV 35S* promoter upstream of *Cas9* was removed *by Bst*E II and *Xba* I double digestion and replaced by a 2112 bp fragment of the *NCR158* (Medtr7g027180) promoter using Gibson Assembly (NovoRec plus One step PCR Cloning Kit, Novoprotein, Shanghai, China). This vector was renamed as pKNF401.The gRNA vector pG-DT1T2.was synthesized containing the gRNA-Scaffold, *AtU6-26t*, transcriptional blocker sequence (***26***), and *AtU6-26pro* (synthesized by GenScript). We used CRISPR-P 2.0 for guide RNA design in this study (***27***). Two specific primers were synthesized, gRNA-T1-F and gRNA-T2-R, pG-DT1T2 were used as a template for amplification of the guide RNA cassette. The PCR product was recombined with the vector in a one-tube Golden Gate reaction using *Bsa* I (NEB) and T4 DNA ligase (NEB) (4 min at 37°C, 3 min at 16°C, 40 cycles), to make the CRISPR target vector pKGSD401-pNCR158:Cas9. Similar cloning strategy was used for constructing the binary vector for dNRE editing except *Cas9* expression was driven by the *CaMV 35S* promoter using the pKGSD401-p35S:Cas9 vector. The guide RNA sequences are listed in **Table S5**.

### Hairy-root transformations

The recombinant plasmids were introduced into *A. rhizogenes* Arqua-1. A single *A. rhizogene*s colony was picked and cultured overnight at 28°C with shaking. *M. truncatula* seeds were sterilized and germinated as described above. A 50 mL volume of the overnight culture was pelleted in a benchtop microfuge at 4,000×rpm speed for 5 min, supernatant was removed, and the pellet was resuspended in 2-5 mL autoclaved ddH$_2$O. About 3 mm of the root tip was removed from each seedling on a sterile flow bench, and then the end of wounded root tip was dipped in the *A. rhizogenes* Arqua-1 culture and seedlings were placed on FP plates. After a week of vertical growth on the plates in a growth chamber (22°C, 16-h light/8-h dark photoperiod), any extended roots were removed, and the seedlings were transferred to FP medium containing 0.5 mM KNO$_3$ and grown at 22°C under 16-h light/8-h dark photoperiod for 3 weeks. The dsRed positive roots were identified under a fluorescent microscope, and non-transgenic roots were removed. The plants were then transferred to pots containing a vermiculite: perlite substrate (1:1, v:v) and kept in a growth chamber at 22 °C under 16-h light/8-h dark photoperiod for 7-10 days before being inoculated.

### Western blot analysis

Fresh nodules (150-200 mg) were homogenized in liquid nitrogen with a cold mortar and pestle. 1 mL of extraction buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM proteinase inhibitor, 10 % glycerol and Triton X-100, pH=7.5) was added to each sample which were then vortexed, incubated on ice for 20 min, and then centrifuged at 12,000g for 20 min at 4°C. The supernatant was then collected in fresh tube and placed on ice. After total protein isolation, protein concentrations were quantified using the Bradford assay (***27***). For western blot analysis, protein samples were separated on 4-12 % gradient SDS-PAGE gel, then transferred to polyvinylidene difluoride membranes. For LgHb detection, an α-LgHb antibody made by ABclonal Co., Ltd. that was designed against a peptide sequence common to all *M. truncatula* LgHbs (Cat No. A20191, Wuhan, China) was used as a primary antibody and HRP-conjugated rabbit α-goat as a secondary antibody (ABclonal Co., Ltd.). Signals were acquired with Western ECL Substrate (Thermo Fisher, USA) under a ChemiScope western blot processor (Amersham Imager 600, GE healthcare, USA).

### Quantification of total heme

For heme extraction, nodules (~100 mg) were ground into a fine powder in liquid nitrogen and resuspended in 1 mL deionized water. After centrifugation (12,000g, 10 min), supernatants were incubated following the manufacturer's protocol (Heme Assay Kit, MAK316-1KT, Sigma-Aldrich, USA).

### Leghemoglobin (LgHb) concentration

For the measurement of LgHb concentration in nodules, 100 mg of fresh nodules were ground and homogenized in 0.3 mL precooled PBS buffer ($Na_2HPO_4$-$NaH_2PO_4$ buffer at 5°C, pH=6.8). The resulting slurry was then centrifuged at 12,000g for 15 min prior to assaying the supernatant by spectrophotometry at wavelengths of 540, 520 and 560 nm. The LgHb concentration was calculated as described previously (***28***). Bovine Hb was used as a protein standard in this experiment.

### RNA extraction and real-time PCR analysis

Total RNA was extracted from nodules using an Ultrapure RNA kit (Tiangen, Beijing, China). First strand cDNA was synthesized and carried out using an EasyScript One-Step gDNA Removal and cDNA Synthesis SuperMix Kit (Tiangen, Beijing, China). The UltraSYBR Green Mixture qPCR kit (GenStar, Beijing, China; Yeasen, Shanghai, China) was used in the qPCR reaction. The real-time PCR was carried out using StepOne Plus (Applied Biosystems, USA). Either *M. truncatula Elongation Factor* (*EF*), or Helicase, UBIQ, Mt0085_00020, and Mt0004_00216 were used as reference genes. The sequences of primers are listed in **Table S5**. Standard errors and statistical significance based on three biological replicates were calculated using the 2-ΔΔCt method (***29***).

### Electrophoretic Mobility-Shift Assays (EMSA)

To create constructs for EMSAs, the partial-length CDS of *NLP2* (encoding C-terminal residues 560–913) and *NIN* (encoding C-terminal residues 566-993) were cloned into pDONR207 and then recombined into pHGWA and pBAD-DEST49 to create pHGWA-NLP2 and pDEST49-NIN, then introduced into *E. coli* BL21. The *pLgHb7* and *pLgHb1* probes were amplified and then labeled with Cy5 by PCR amplification. Primers used for probe synthesis are listed in **Table S5**. NIN and NLP2 proteins were expressed in *E. coli* transformed with pDEST49-NIN and pHGWA-NLP2 and then purified as described as below. The purified proteins were then incubated with the Cy5-labelled probe at 37 °C for 30 min in EMSA buffer (20 mM Tris, pH=7.9, 5 % glycerol, 40 ng $\mu L^{-1}$ BSA, 50 mM $MgCl_2$ and 5 mM dithiothreitol). After incubation, the reaction mixture was electrophoresed on a 6 % native polyacrylamide gel with running buffer (25 mM Tris, 192 mM glycine, 0.1 % SDS) and then imaged with a Starion FLA-9000 imager (Fujifilm).

For the EMSA using the pLgHb7 promoter fragment the mutated variant pLgHb7m had three two bp mutations as shown below: the conserved dNRE indicated in bold, deleted bases are indicated with hyphens the corresponding bases in WT are shown in red, the consensus dNRE is shown for reference.

| | |
|---|---|
| pLgHb7 | TA**TTGTCTC**T**TTAA**TAATGT**CAACAGCCATTTC**CACAAG**CCAATAGATT**CTT |
| pLgHb7m | TA--**GTCTCTTTAA**TAATGT**C--CAGCCATTTC**CACAAG**CC--TAGATT**CTT |
| consensus dNRE | TK**KYYYY-WTMA**NNNNNNN**Y**AA**YRGYCAYYBC**NNNNNN**CC**AA**YARATT** |

For the *nsHB1* EMSA the following version of the *M. truncatula nsHB1* promoter sequence containing two copies (2X) of the NRE was used:

*original WT nsHB1 promoter sequence:*
GAATGAATGAAAGTATGAATGACTCCCCATTGGTATGAAGGGTCAATTTCATCTTCAATTAA

*pnsHB1 2×NRE probe used for EMSA- 126bp:*
cggtacctcgcgaatgcaGAATGAATGAAAGTATGAATGACTCCCCATTGGTATGAAGGGTCAAATGACTCCCCATTGGTATGAAGGGTCAATTTCATCTTCAATTAAatcccgggcccgtcgact

The NREs are indicated in red and green. Adaptor primers are shown in lowercase.

## Protein Purification

The constructed recombinant plasmid was transformed into *E. coli* BL21 competent cells, and plated on LB solid medium containing appropriate antibiotics, and the positive transformants were identified. Single colonies were inoculated into LB liquid medium with antibiotic and incubated at 37°C overnight. The next day, the cultures were inoculated into 10 mL antibiotic-containing LB medium with 1:100 dilution and incubated at 37°C for about 1-2 h until $OD_{600nm}$ ~ 0.4 to 0.5, and then induced with 0.8mM IPTG. A culture without IPTG was used as a control. The bacteria were incubated at 28°C with shaking for 6 h and then were kept at 4°C for 10-30 min. Then the bacteria were centrifuged at 4,000 rpm for 5 min at 4°C to form a pellet. Then 3~6 mL of pre-cooled HIS Lysis Buffer (with 50 mM $NaH_2PO_4$ and 300 mM NaCl, pH=8.0) was used to resuspend the bacteria and the liquid was transferred into a 5 mL centrifuge tube. The tubes were then placed on ice and sonicated (Noise Isolating Chamber, Ningbo Scientz Biotechnology Co. Ltd) until the bacteria solution clarified. The supernatant was then centrifuged at 12,000g for 5 min on a microcentrifuge (Eppendorf) at 4°C and the supernatant was then transferred to a new tube after. Then 5 mL of HIS Lysis Buffer was used to wash the protein purification column, and 20-50 µL of Ni-NTA Agarose beads was added. The column containing beads was washed three times using HIS Lysis Buffer. The lysate supernatant was added to the column and was allowed to flow through for 1-2 h at 4°C. The column was then washed twice using HIS Lysis Buffer and once using a wash buffer containing 50 mM imidazole. In the end, 1-2 mL HIS Elution Buffer (with 50 mM $NaH_2PO_4$, 300 mM NaCl and 50 mM imidazole, pH=8.0) was applied to the column and collected. 20 µL samples in every step were collected and electrophoresed and identified by Coomassie brilliant blue staining. The imidazole in the eluate was removed by ultrafiltration to obtain the purified protein.

## Heterologous expression assays

For promoter-GUS analyses, the *LgHb7* promoter sequence was amplified by PCR and cloned into pENTR and then transferred to pKGWFS7 by the LR reaction; the vector PUB-GFP-HA was cut open with *Kpn* 1 and *Xba* 1, and then the *NLP2* and *NIN* CDS were amplified from cDNA and used to replace the GFP-HA in PUB-GFP-HA, using Gibson cloning (NovoRec plus One step PCR Cloning Kit, Novoprotein, Shanghai, China). For the mutant variants, PCR mutagenesis was used to mutate the RWP-RK domain as follows: *NLP2* WT 1861-1878 bp (AGATGGCCATCGCGCAAG: RWPSRK) to *NLP2m* (AGAGCGGCATCGCGCAAG: RAASRK), *NIN* WT 1966-1984bp (AGATGGCCTTCAAGAAAA: RWPSRK) to *NINm* (AGAGCGGCTTCAAGAAAA: RAASRK). The new constructs pKGWFS7-pLgHb7, PUB-NLP2 and PUB-NLP2m were used for transient transformation of *Nicotiana tabacum* L.. *N.*

*tabacum* leaves co-infiltrated with the constructs were stained with a solution containing 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 0.5 mg/mL 5-bromo-4-chloro-3-indolyl-$\beta$-D-glucuronic acid (X-Gluc), 0.1 % Triton X-100, and 0.1 M sodium phosphate buffer (pH=7.0). The leaves were incubated at 37 °C in the dark for 12-24 h. To stop the reaction and clear the roots, the samples were washed with 75 % ethanol three times. GUS-staining was then photographed. For GUS quantification, *N. tabacum* leaves were infiltrated with the constructs and leaf discs were macerated with 50 mM disodium hydrogen phosphate, 10 mM $Na_2EDTA$, $\beta$-mercaptoethanol, 0.1 % Sarcosyl and 0.1 % TritonX-100) using a JXFSTPRP-48 tissue lyser (Jingxin, Shanghai). The resulting slurry was then centrifuged at 12,000 rpm for 10 min, the supernatant was then mixed with 20 mM 4-MUG and incubated at 37 °C in the dark for 60 min, and then stopped by adding 0.2 M sodium carbonate and then assayed spectrophotometry by microplate reader at 355 nm excitation and 460 nm emission filter. Protein concentrations of the samples were determined using a standard curve based on a concentration gradient made using 4-MUG and BSA.

The design of the dNRE-deleted version of the *LgHb7* promoter is shown below, the bases deleted are indicated by hyphens and the dNRE is indicated in red.

```
pLgHb7     AAATGATTATTGTCTCTTTAATAATGTCAACAGCCATTTCCACAAGCCAATAGATTCTTTAAAAAAAAAATCATAAAAAATC
pLgHb7 mut AAATGATTATTG-----------------------------------------------------------CATAAAAAATC
```

## $^{15}N_2$ fixation assays

The procedure used followed that described by (***30***). Briefly, *M. truncatula* plants 21 dpi with *S. meliloti* Rm2011 were removed from the soil substrate and the intact plants were incubated in 20 % $^{15}N_2$ gas (29 mLof air was removed and replaced with 29 mL $^{15}N_2$), and then incubated for 15 min. The incorporation of $^{15}N_2$ was determined separately by separating roots, shoots, and nodules. Tissue samples were dried overnight at 700 °C, and dry weights were determined. Then, around 100 mg of dried root or shoot was then crushed and sent for combustion and mass spec analysis. Total nitrogen and $^{15}N$ enrichment were measured in a continuous flow isotope ratio mass spectrometer (James Hutton Institute, UK).

## Transcriptomics data

The Affymetrix GeneChip Medicago Transcriptome Array was used in this experiment (***31, 33***). Total RNA was quantified using a NanoDrop ND-2000 (Thermo Scientific) and the RNA integrity was assessed using Agilent Bioanalyzer 2100 (Agilent Technologies). The sample labeling, microarray hybridization and washing were performed based on the manufacturer's standard protocols. Briefly, total RNAs were transcribed to double strand cDNAs and then synthesized cRNAs. Next, 2nd cycle cDNAs were synthesized from cRNAs. Followed fragmentation and biotin labeling, the 2nd cycle cDNAs were hybridized onto the microarray. After washing and staining, the arrays were scanned by the Affymetrix Scanner 3000 (Affymetrix). Affymetrix GeneChip Command Console (version 4.0, Affymetrix) software was used to extract raw data, carry out RMA normalization and to calculate P-values, and FDR corrected P-values. CVs were calculated using Excel.

### Promoter Analysis

The MEME suite was used to analyze 1kb regions upstream of the promoters of the 55 genes most strongly downregulated in *nlp2-1* nodules across conditions. The identified motifs were analyzed with the Tomtom using the Arabidopsis DAP motifs from O'Malley et al. (***34***).

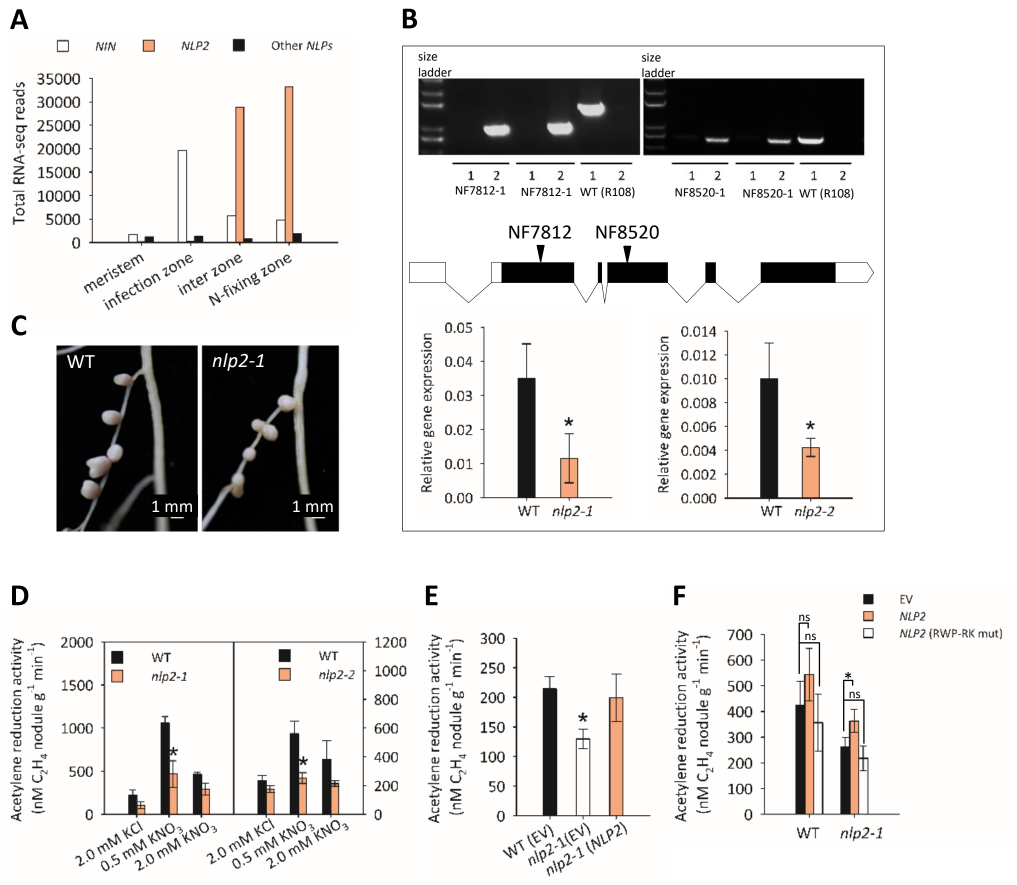


**figure S1. Expression of *NLPs* in *M. truncatula* and isolation of *NLP2 Tnt1* insertion mutants and their nodulation phenotypes.**

(**A**) Expression of *NIN*, *NLP2*, and the cumulative expression of all other *NLPs* (*NLP1*, *3*, *4*, *5*) across nodule zones. Data from laser capture dissection **(*21*)**. (**B**) Identification and characterization of *NLP2* transposon insertion mutants. Detection of two homozygous individuals with a *Tnt1* insertion in *NLP2* at +483 bp (*nlp2-1;* line NF7812) and +1188 bp (*nlp2-2*, NF8520). *EF* was used as a reference gene. (1) *NLP2* gene-specific primers (2) *Tnt1*-specific primer with a *NLP2*-specific primer. The positions of the insertions are indicated by arrowheads, solid black and white indicates exons and UTRs, respectively. Thin lines indicate introns. Relative expression of *NLP2* in *nlp2-1*, *nlp2-2*, and WT using qRT-PCR. n=3. *EF* was used as a reference gene. (**C**) Nodulated roots of WT and *nlp2-1* mutants, 21 dpi with *S. meliloti* Rm2011 (under 0.5 mM $KNO_3$). **(D)** ARA of plants nodulated with *S. meliloti* Rm2011 under different levels of $KNO_3$ or KCl and nodules assayed at 21 dpi. n=3. **(E)** Complementation of ARA in *nlp2-1*. *nlp2-1* seedlings were transformed with *pNLP2-NLP2* or empty vector (EV) controls and evaluated 21 dpi with *S. meliloti* Rm2011 nodulated under 0.5 mM $KNO_3$. n=3, each using nodules from 2-3 composite plants. (**F**) ARA of composite *nlp2* mutant plants transformed with either empty vector (EV) control or with *NLP2* or *NLP2* (RWP-RK mut) with a mutated RWP-RK domain, driven by the native *NLP2* promoter. n=3. For (B, D-F) means (±SD) were compared using Student's one-sided *t*-test, $\alpha$=0.05.

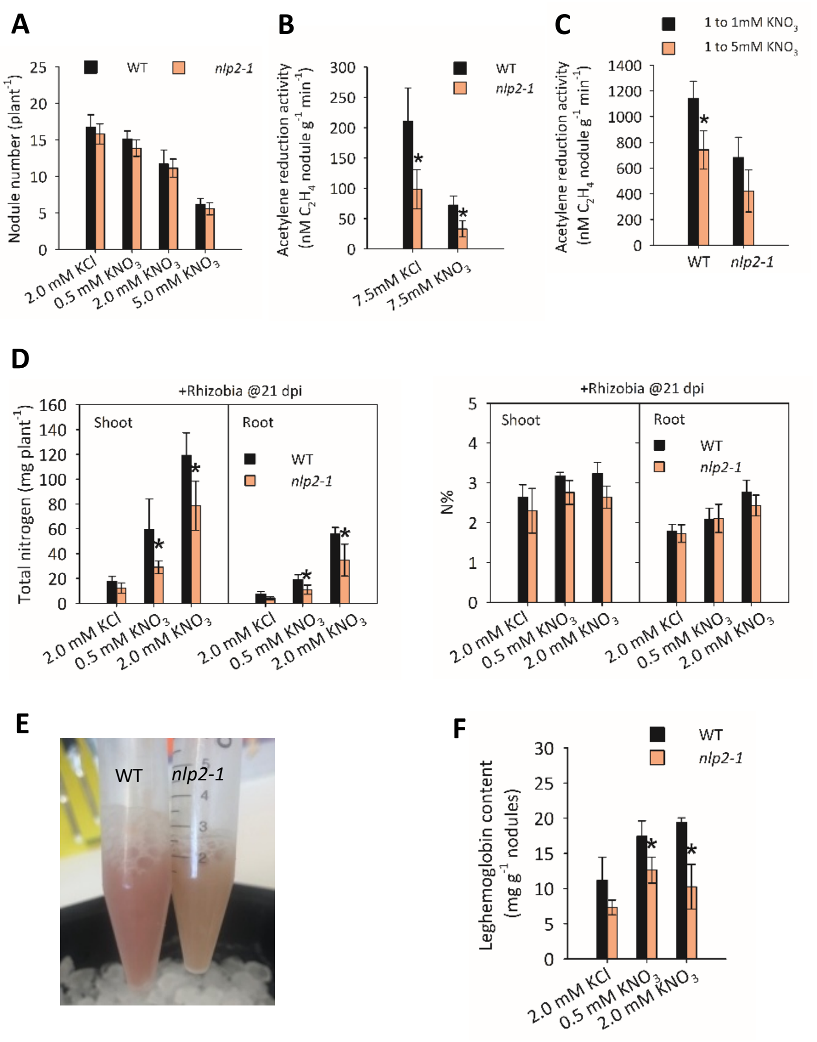
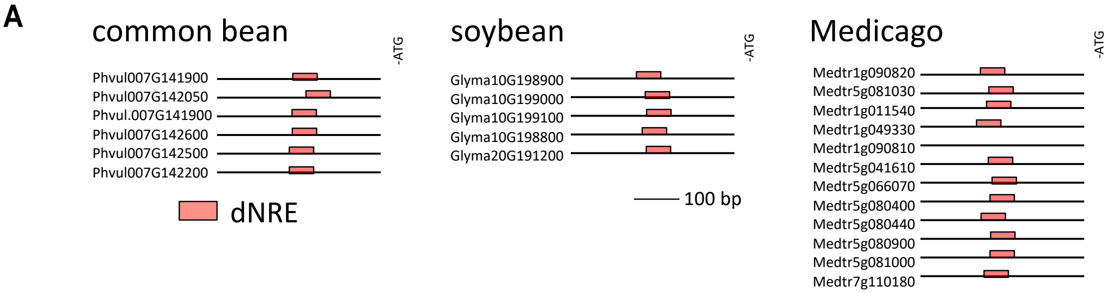


**figure S2**. **$N_2$-fixation and *LgHb* content is reduced in *nlp2-1* nodules.**

**(A)** Nodule numbers of WT and *nlp2-1* mutants under different weekly N-regimes at 21 dpi. **(B)** ARA (nM $C_2H_4$ nodule$^{-1}$ min$^{-1}$) of *nlp2-1* and WT nodules with weekly provision of 7.5 mM KCl or 7.5 mM $KNO_3$. (**C**) ARA of *nlp2-1* and WT nodules two days after addition of $KNO_3$. Plants were nodulated with a weekly provision of 1.0 mM $KNO_3$ and treated with either 1.0 or 5.0 mM $KNO_3$ at 21 dpi. **(D)** Total N (Kjeldahl's) and N% in *nlp2-1* and WT nodules. Plants were nodulated under 0.5 mM $KNO_3$ and evaluated at 21 dpi. (**E**) Ground nodule extracts of *nlp2-1* and WT at the same protein concentration. (**F**) nodule leghemoglobin content. For (A, B and F) , n= 3, for (C) n=4, for (D) n=3-4. Means are shown ±SD is indicated. Means were compared with Student's two-sided *t*-test, α=0.05. All plants were nodulated with *S. meliloti* Rm2011.



**B**

| Gene model | Gene | Promoter sequence | Expression log2(FKPM) | mismatches/ indels |
|---|---|---|---|---|
| Medtr5g066070 | LgHb9 | TTTTGTCTC-TTAATAACTACAATGGTCACCTCCACAAGCCATTATATTCTT | 19.7 | 2 |
| Medtr5g081030 | LgHb2 | TCTGGTCTC-TTCATAATTTCAATGGTCATTTCCACAAGCCAATAGATTCTT | 19.6 | 0 |
| Medtr5g081000 | LgHb8 | TATTGTCTCTATAATATTGTCAATAGCCATTTCCACAAGCCAATAGATTCAT | 19.6 | 0 (1 ins) |
| Medtr5g080440 | LgHb1 | TCTTGTCTC-TTAATAATTTTAATGGTCACCGCCACAAGCCAATAGATTCTT | 18.7 | 0 |
| Medtr1g011540 | LgHb7 | TATTGTCTCTTTAATAATGTCAACAGCCATTTCCACAAGCCAATAGATTCTT | 18.3 | 0 (1 ins) |
| Medtr5g041610 | LgHb6 | TCTTGTCTC-TTAATAATGTCAATAGCCACCTCCACAAGCCAACAAATTCTT | 17.8 | 0 |
| Medtr1g090820 | LgHb10 | AATTGTCTC-TTAATATTGCCAATGGCCACCTCTAAATTAAAGACCAATAGA | 17.7 | 6 |
| Medtr1g049330 | LgHb4 | GATTGTCTC-TTAATAATACCAATAGCCATCTCATAATTAAAGGCCAATAAA | 17.7 | 6 |
| Medtr5g080400 | LgHb3 | TCTTGTCTC-TTAATAATTTCAATGGTCACCCCCAGAAGCCAATAGATTCTT | 15.3 | 0 |
| Medtr1g090810 | LgHb11 | TATTATCTCTT----AATAAAGGTTGCCTCCTCTCAAAGGCCAAGAAATACT | 13.5 | 11 (3 del) |
| Medtr7g110180 | LgHb12 | TCACGACTC-TTTATAATAGTAATGGACGTTTCCACAAGCCAACAAATTTAA | 11.5 | 5 |
| Medtr5g080900 | LgHb5 | TATTGTTTCTTTGATAATGTCAATGACCATTTTCACGATTCTTTTAAATATA | 6.2 | 8 (1 ins) |
| Consensus dNRE | | **KKYYYY-WTMA** **YAAYRGYCAYYBC** **CCAAYARATT** | | /33 |

**C**

| Gene model | V2.0 Promoter sequence | Expression log2(FKPM | mismatches/ indels |
|---|---|---|---|
| Glyma10G199100 | TGTTTCTTCTTCATCATGCTGATTGACACCCTCCACAAGCCAAGAGAAACAC | 18.2 | 5 |
| Glyma10G199000 | CATTGGCTCTTCGTCATGCCGATTGACACCCTCCACAAGCCAAGAGAAACTT | 15.9 | 9 |
| Glyma20G191200 | ATTTGGCTCTTCATCATGCCGATTGACACCCTCCACAAGCCAAGAGAAACTT | 15.1 | 8 |
| Glyma10G198800 | GATTGTCTCTTCACCATACCAATTGATCACCCTCCTCCAACAAGCCAAGAGA | 15.0 | 10 |
| Glyma11G121800 | AAAATGGACTACGAAAACTTGAGTAACAAGGTAAGCCACACAAATGGGAATG | 6.2 | 18 |
| Glyma11G121700 | GTTTGTTTCTTGA-AAGCTCAATCAAAGAACGTGGCAACCAAAAGCATGGGA | 0.0 | 11 (1 del) |
| consensus dNRE | **KKYYYYWTMA** **YAAYRGYCAYYBC** **CCAAYARATT** | | /33 |

**figure S3. dNRE elements in promoters of *LgHbs* in different legumes.**
**(A)** dNRE elements (represented by single red boxes) detected in promoters of the *LgHb* homologs from *Phaseolus vulgaris* (common bean), *Glycine max* (soybean) and *M. truncatula* as detected by MEME. **(B)** Comparison of the similarity of dNRE motifs in *M. truncatula LgHb* promoters with the consensus dNRE motif and the expression of the corresponding *LgHb* genes in nodules. Data from (*21*). **(C)** Comparison of the similarity of dNRE motifs in soybean *LgHb* promoters with the consensus dNRE motif and the expression of the corresponding *LgHb* genes in nodules (FPKM: fragments per kilobase of exon model per million RNAseq reads mapped). Data from Soybase. Note a base was inserted (yellow highlighted N) into the consensus sequence to improve the alignment. In (B) and (C), conservation with the consensus motif is indicated in red. Mismatches are underlined. The numbers of mismatches, and insertions/deletions (indels) compared to the 33 bp semi-palindrome motif (indicated by arrows in B) are listed.

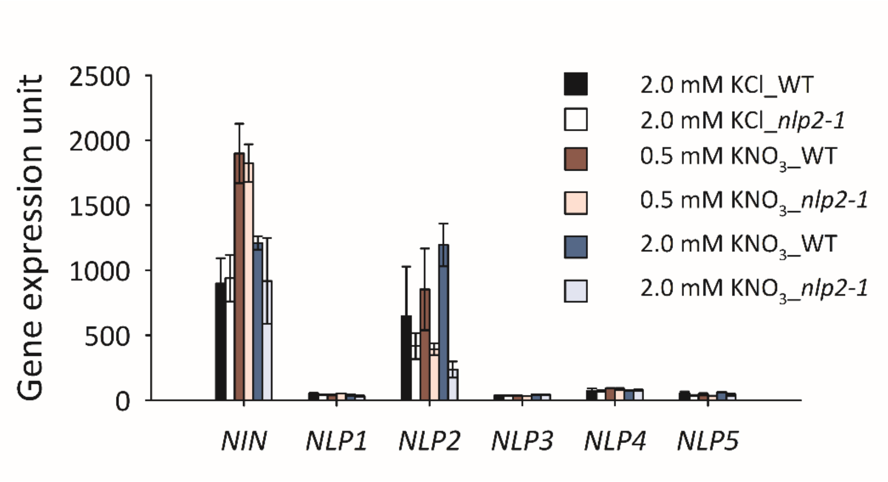


**figure S4**. **Expression of *NLPs* in WT and *nlp2-1* nodules under different N-regimes.**
Affymetrix GeneChip. Nodules from WT and *nlp2-1* were harvested at 21 dpi with *S. meliloti* Rm2011.

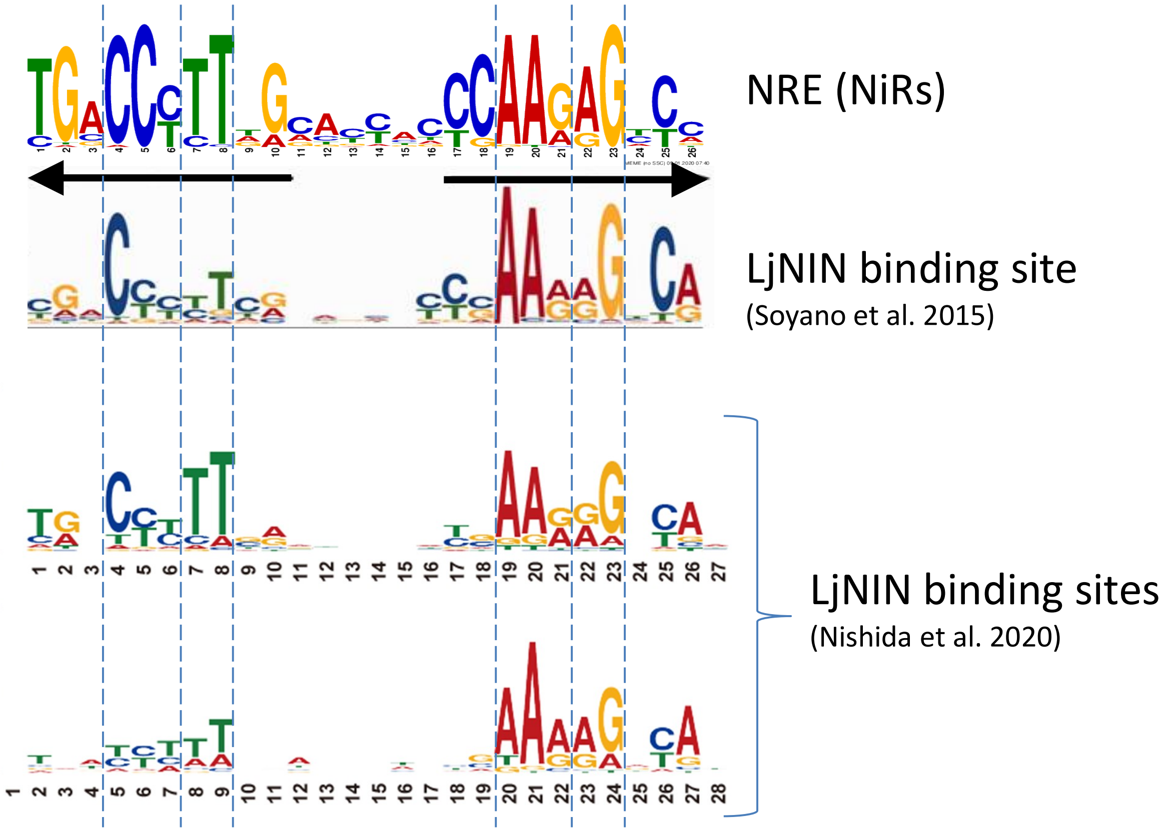


**figure S5. Comparison of NRE and NBS motifs.**
NRE that was determined from 38 *NIR* promoters using MEME is aligned to the reported NIN-binding sites (NBS) (*9, 10*). The semi-palindrome is indicated by arrows. Dotted lines are provided to aid in visual comparison.

RWP-RK (NLP2 WT  1861-1878bp from ATG)
AGATGGCCATCGCGCAAG
 R  W  P  S  R  K

RWP-RK (NLP2 mutation)
AGAGCGGCATCGCGCAAG
 R  A  A  S  R  K

RWP-RK (NIN WT  1966-1984bp from ATG)
AGATGGCCTTCAAGAAAA
 R  W  P  S  R  K
RWP-RK (NIN mutation)
AGAGCGGCTTCAAGAAAA
 R  A  A  S  R  K

**figure S6. Mutated versions of NIN and NLP2 used for EMSA.** The predicted changes in protein sequence are shown below the coding sequence.

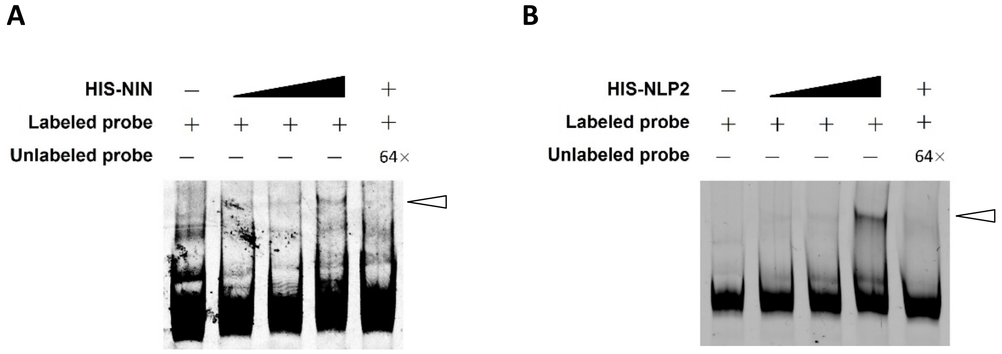


**figure S7. EMSA to test for interaction between a *LgHb1* promoter fragment with NLP2 and NIN.** (**A**) CY5-labeled *LgHb1* promoter fragment containing the dNRE with HIS-tagged NIN fragment containing the RWP-RK domain. (**B**) EMSA using CY5-labeled *LgHb1* promoter fragment containing the dNRE and a HIS-tagged NLP2 fragment containing the RWP-RK domain. Arrows indicate a band that was abolished by addition of excess unlabeled probe.

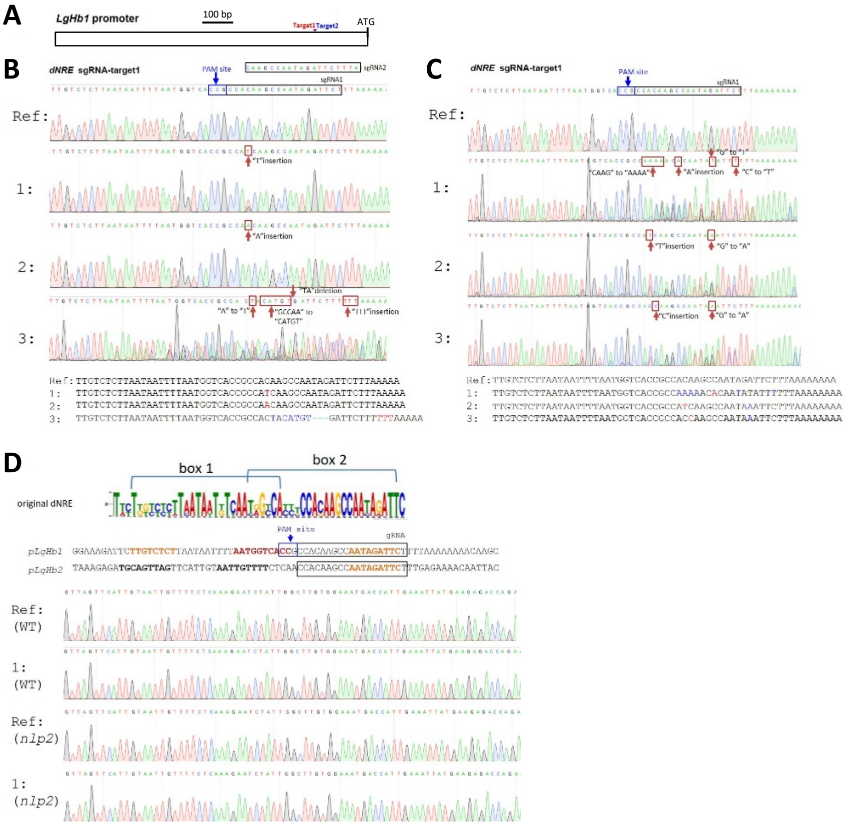


**figure S8. Targeted mutation of the dNRE element in the promoter of *LgHb1*.**
**(A)** Location of the dNRE target in the *LgHb1* promoter **(B, C)** Editing events in WT (A) and *nlp2-1* (C) nodules of composite plants transformed with *p35S:Cas9* and two gRNAs targeting the dNRE of *LgHb1* (dNRE-KO) using *A. rhizogenes* hairy-root transformation. DNA was harvested from nodules harvested 21 dpi with *S. meliloti* Rm2011 and the target region was PCR amplified and sequenced. In (B) and (C) Ref indicates sequence of target region of nodules from WT or *nlp2-1* roots transformed with empty vector (EV), and 1-3 represent sequences of the target region from individual nodules of dNRE-KO plants. Mutations are indicated in the chromatogram with red arrows and boxes, and are shown below in red (insertions), blue (substitutions) or green (deletion) text. **(D)** The *M. truncatula* R108 genome was searched for other potential binding sites for the gRNAs. Only one other perfectly matching site was found in the promoter sequence of *LgHb2* corresponding to the dNRE; gRNA1 was found to be perfectly complementary to the *LgHb2* dNRE but lacked a PAM site, so this site was used as a negative control. The original dNRE motif is shown for reference. The *LgHb2* promoter was not edited in either WT or *nlp2-1* nodules of the dNRE-KO plants.

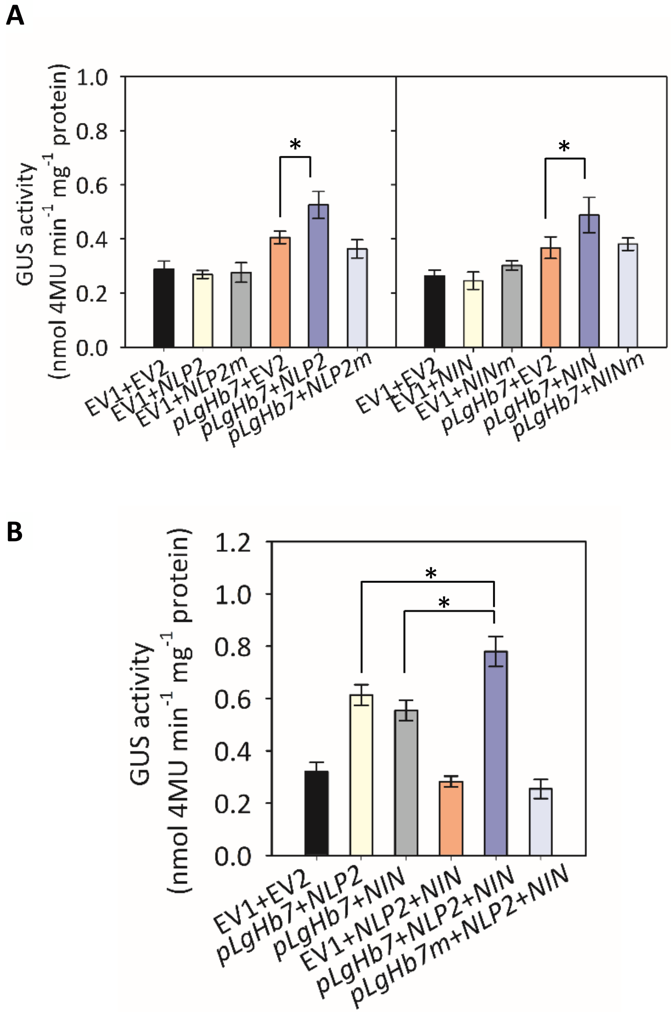


**figure S9. Trans-activation of the *LgHb7* promoter by NLP2 and NIN.**
(**A**) The *pLgHb7:GUS* reporter was co-infiltrated into *N. benthamiana* leaves with either vectors for expression of *NLP2*, *NIN,* or versions of these vectors in which the RWP-RK sequence has been mutated (*NLP2m* and *NINm*). (**B**) *A. tumefaciens* carrying vectors for expression of *NIN* and/or *NLP2* were co-infiltrated into *N. benthamiana* with a vector in which the *GUS* gene is driven by a -342 to -1 bp fragment of the *LgHb7* promoter (*pLgHb7*) containing the dNRE or a mutated form (*LgHb7m)* with the dNRE deleted (see supplemental methods for details). EV1: empty vector no-promoter control, EV2: empty vector no-*NLP2* control. *NLP2* and *NIN* expression were driven by the 35S promoter. n=3, ±SD. Each replicate consisted of six leaves. α=0.05, means were compared using a Student's *t*-test.

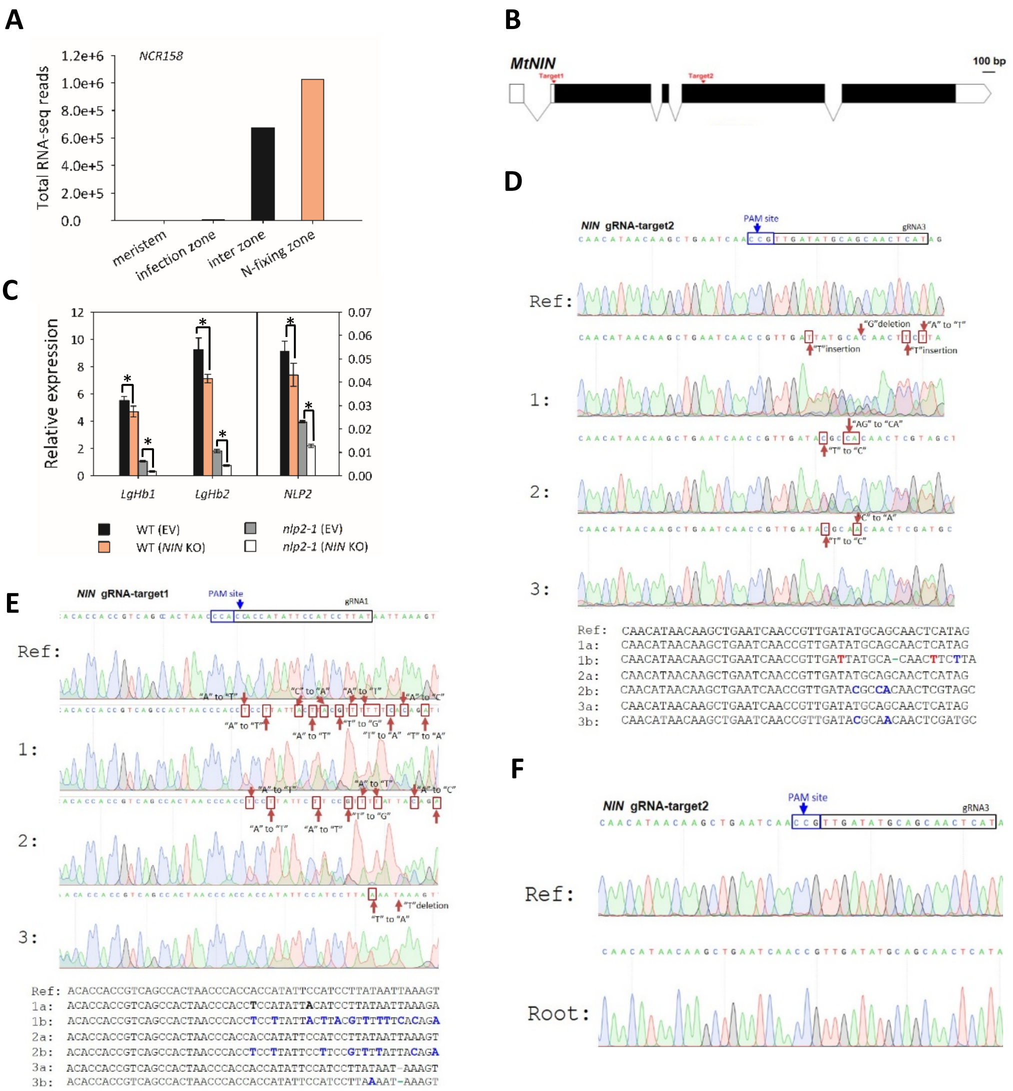
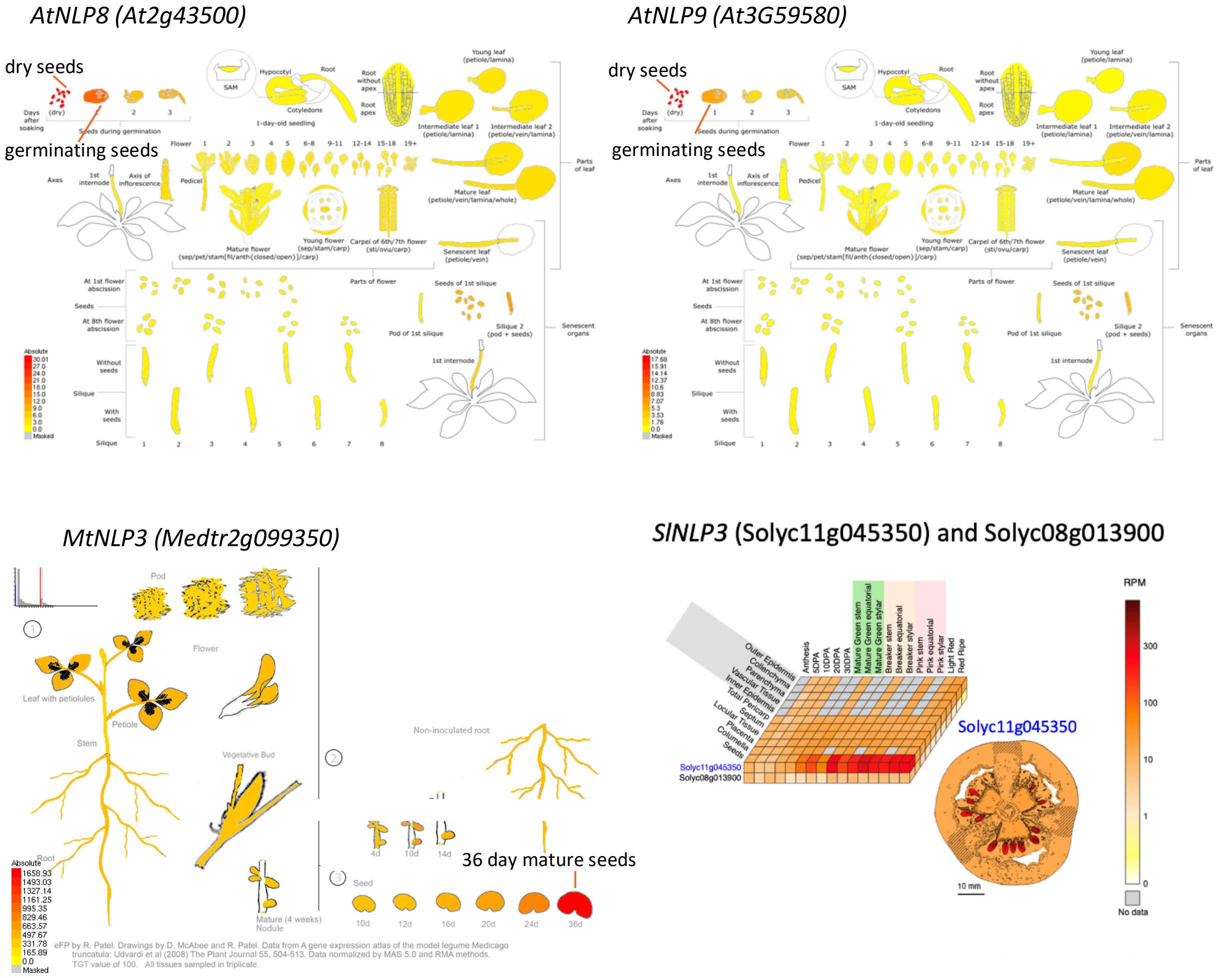


**figure S10. CRISPR-induced editing of *NIN* in WT and *nlp2-1* nodules using tissue-specific knockout.** *NIN*-KO constructs containing *Cas9* driven by the *NCR158* promoter which is expressed specifically in the nodule $N_2$-fixation zone, were introduced into *M. truncatula* roots using *A. rhizogenes* hairy-root transformation. **(A)** Expression of *NCR158* (Medtr7g027180) across nodule zones. Data from (*29*). **(B)** The position of the gRNAs targets used to edit *NIN* in WT (target 2) and *nlp2-1* (target 1). (**C**) Relative expression (qRT-PCR) of *LgHb1*, *LgHb2* and *NLP2* in nodules of *NIN*-KO and empty vector (EV) transformed roots of *nlp2-1* and WT. **(D)** One nodule was harvested from roots of seven separate *NIN*-KO composite plants. In five of the seven plants target 2 was edited and no editing was seen in nodules of the other two plants. Three examples of editing are shown for gRNA2. **(E)** A construct with a gRNA targeting *NIN* (target 1) was introduced into *nlp2-1* by hairy-root transformation. Three examples of editing are shown (1-3). For (B and C) nodules of EV-transformed roots were used as a reference and mutations are indicated in the chromatogram with red arrows and boxes and are shown below in red (insertions), blue (substitutions) or green (deletion) text. **(F)** Roots from nodulated WT plants transformed with *NIN*-KO were not edited at target site 2.



*AtNLP8 (At2g43500)*

*AtNLP9 (At3G59580)*

*MtNLP3 (Medtr2g099350)*

*SlNLP3* (Solyc11g045350) and Solyc08g013900

**fig S11. Some orthologs of nodule expressed NLPs are expressed in seeds.** Expression profiles of *AtNLP8/9* in *A. thaliana*, Solyc11g045350 and Solyc08g013900 in *Solanum lycopersicum* (tomato), and *MtNLP3*. Data from the Medicago Gene Atlas v3.0 (presented with the eGFP browser (presented with the eGFP browser (*32, 34, 35*) and from the "Tomato Expression Atlas": tea.solgenomics.net".

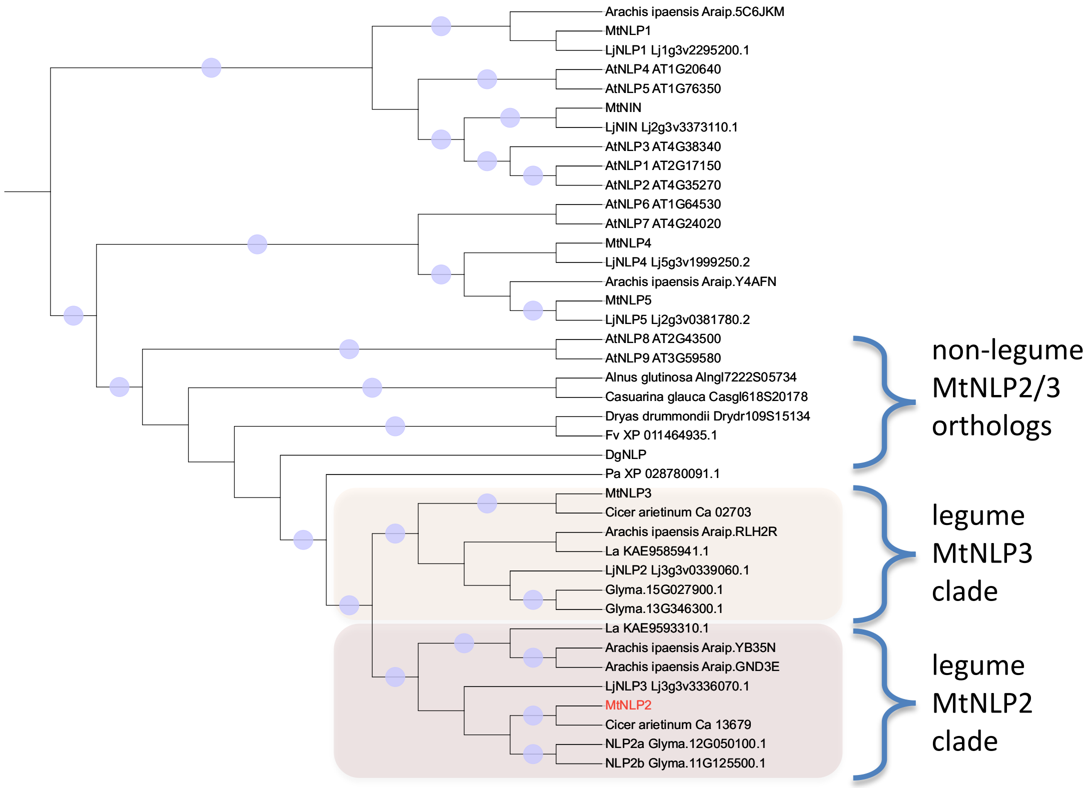


**figure S12. Phylogenetic distribution of NLPs.** MtNLP2 belongs to a legume-specific clade and has orthologs in papilionoid legumes but none in the early diverged legume *Prosopis alba* (Pa) (*Caesalpiniodeae*), and the nodulating non-legumes *Datisca glomerata* (Dg)*, Casuarina glauca,* and *Alnus glutinosa.* MtNLP2 is indicated in red, the MtNLP3 and MtNLP2 clades are indicated by shading. Glyma: *Glycine max*, La: *Lupinus angustifolius,* Fv: *Fragaria vesca.* At: *Arabodopsis thaliana,* Lj: *Lotus japonicus,* Mt: *Medicago truncatula.* The circles indicate >85% bootstrap branch support.

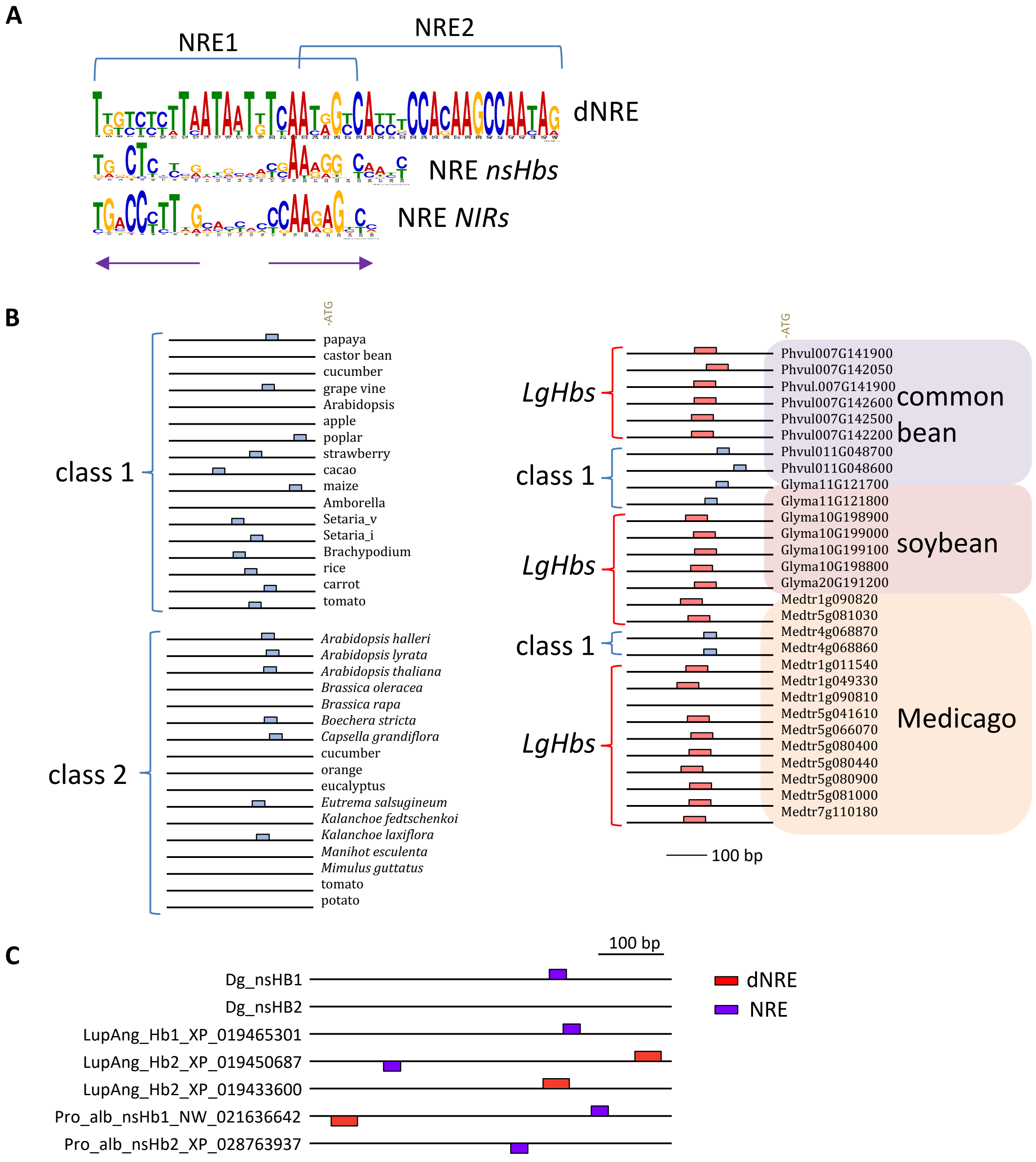


**figure S13. NRE and dNRE elements in promoters of *nsHbs* and *LgHbs*.**
**(A)** Motifs of NRE elements detected in promoters of 18 of 23 class I *nsHbs*, and 8 of 17 class II *nsHbs*. The dNRE and the NRE detected in promoters of 38 *NIR* genes from different species are aligned for comparison. **(B)** Positions of NREs and dNREs in promoters of class I and class II *nsHbs* and *LgHbs*. The motifs were detected using MEME. (**C**) dNRE and NRE elements in *Hb* promoters of *Datisca glomerata* (Dg), *Prosopis alba* (Pro_alb) and *Lupinus angustifolius* (LupAng). Motifs were detected by meme using an an extended dataset of NRE/dNRE containing promoters.

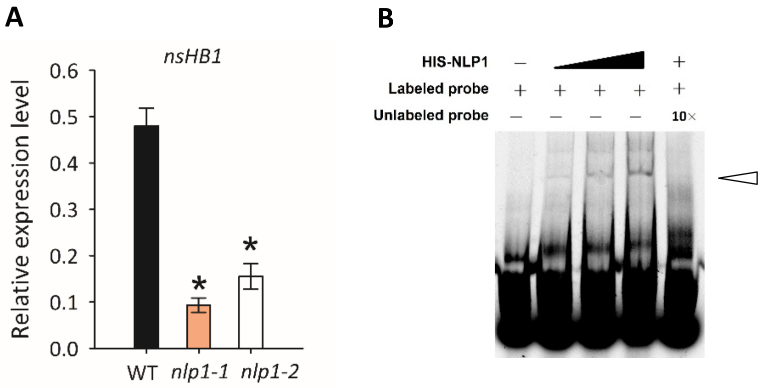


**figure S14. NLP1 regulates the expression of *M. truncatula nsHB1* and binds to its NRE motif.**
**(A)** Relative expression of the *M. truncatula nsHB1* under nitrate replete conditions. n=4. The means are shown ±SD and were compared using Student's *t*-test, n=3, α=0.05. *EF* was used as a reference gene. **(B)** EMSA showing binding of NLP1 to a probe with 2×NRE from the *nsHB1* (Medtr4g068860) promoter. Details on the probe design are provided in the Supplementary Material. Arrow indicates band that was abolished by addition of excess unlabeled probe.

**Table S1. Expression of heme biosynthesis genes and nitrate transporters in nodules of *nlp2-1* and WT of plants nodulated under different $NO_3^-$ regimes.** Affymetrix GeneChip. P value, *P<0.05, **P <0.01.

| | | ***nlp2*/WT (gene expression units)** | | | | | |
|---|---|---|---|---|---|---|---|
| **Gene model** | **Encoded protein** | **2.0 mM KCl** | | **0.5 mM $KNO_3$** | | **2.0 mM $KNO_3$** | |
| | | **Fold change** | **P value** | **Fold change** | **P value** | **Fold change** | **P value** |
| Medtr5g098800 | coproporphyrinogen III oxidase | -1.39* | 0.04 | -1.3 | 0.33 | -1.56* | 0.02 |
| Medtr2g076340 | uroporphyrinogen decarboxylase | 1.07 | 0.57 | 1.1 | 0.53 | -1.32 | 0.06 |
| Medtr3g072330 | uroporphyrinogen decarboxylase | -1.14 | 0.17 | 1.31* | 0.01 | -1.24 | 0.11 |
| Medtr1g052525 | uroporphyrinogen III synthase | -1.43 | 0.40 | -1.38 | 0.17 | -1.16 | 0.26 |
| Medtr3g101350 | uroporphyrinogen III synthase | -1.03 | 0.86 | 1.1 | 0.57 | -1.14 | 0.36 |
| Medtr7g034345 | uroporphyrinogen-III C-methyltransferase | -1.07 | 0.70 | -2.19** | 8.4E-05 | -2.67** | 8.2E-05 |
| Medtr3g026780 | uroporphyrinogen-III C-methyltransferase | -1.04 | 0.81 | -3.37** | 4.8E-07 | -2.61** | 6.2E-06 |
| Medtr3g075110 | oxygen-independent coproporphyrinogen III oxidase, putative | 1.52* | 0.04 | -1.01 | 0.95 | 1.44 | 0.21 |
| Medtr3g118070 | glutamate-1-semialdehyde 2,1-aminomutase | -1.02 | 0.16 | -1.31** | 0.01 | -1.57** | 3.6E-05 |
| Medtr1g085730 | protoporphyrinogen IX oxidase | -1.16 | 0.15 | -1.39 | 0.06 | -1.89* | 0.001 |
| Medtr3g088610 | delta-aminolevulinic acid dehydratase | -1 | 0.95 | 1.29* | 0.04 | -1.53* | 0.01 |
| Medtr4g057890 | *NRT2.2* | -2.56 | 0.12 | -5.55* | 0.004 | -6.28* | 0.05 |
| Medtr4g057865 | *NRT2.1* | -3.27 | 0.07 | -6.87* | 0.003 | -5.40* | 0.05 |
| Medtr4g104730 | *NAR2.1b* | -2.42** | 0.003 | -2.38** | 0.0001 | -2.13* | 0.02 |
| Medtr4g104700 | *NAR2.1a* | -1.92** | 0.01 | -2.03** | 0.001 | -1.92** | 0.01 |

**Table S2. Expression of JA biosynthesis genes in nodules of *nlp2-1* and WT of plants nodulated under 2.0 mM KCl or 2.0 mM $KNO_3$.** Affymetrix GeneChip. P value, *P<0.05, **P <0.01.

| **Gene model** | **Gene name** | ***nlp2-1*/WT (gene expression unit)** | | | |
|---|---|---|---|---|---|
| | | **2.0 mM KCl** | | **2.0 mM $KNO_3$** | |
| | | **Fold change** | **P value** | **Fold change** | **P value** |
| Medtr5g013530 | *Jasmonate ZIM-domain Protein 1A (JAZ1A)* | 1.07 | 0.29 | 6.43* | 0.03 |
| Medtr2g042900 | *Jasmonate ZIM-domain Protein 2 (JAZ2)* | -1.25 | 0.26 | 2.09** | 0.01 |
| Medtr4g054920 | *CYP94C1* | -1.1 | 0.07 | 2.23 | 0.06 |
| Medtr5g053950 | *Allene Oxide Cyclase 4 (AOC 4)* | -1.01 | 0.42 | 2.07* | 0.05 |
| Medtr7g068780 | *Fatty Acid Desaturase 8 (FAD8)* | -1.02 | 0.43 | 2.28** | 0.002 |
| Medtr8g067280 | *MYC2* | 1.08 | 0.44 | 1.51 | 0.10 |

## Table S3. NRE and dNRE elements detected in promoters of the 55 most strongly downregulated genes in *nlp2* nodules.

| Gene Model (Mtv4.0) | Mtv4.0 annotation | NRE/dNRE |
|---|---|---|
| **Medtr2g100710.1** | 6-phosphofructokinase | |
| **Medtr1g026070.1** | anaphase-promoting complex subunit 11 RING-H2 | |
| **Medtr3g089005.1** | basic blue-like protein | NRE |
| **Medtr3g089125.1** | bidirectional sugar transporter | NRE |
| **Medtr2g099790.1** | calmodulin-binding family protein | NRE |
| **Medtr1g093850.1** | CHUP | NRE |
| **Medtr3g079850.1** | *SymCRK* | NRE |
| **Medtr4g130780.1** | early nodulin-like protein | |
| **Medtr4g099400.1** | expansin-B1-like protein | |
| **Medtr4g086020.1** | ferredoxin-nitrite reductase (NiR) | NRE |
| **Medtr1g069825.1** | G1-like protein | |
| **Medtr5g067410.1** | glutamyl-tRNA reductase | |
| **Medtr5g081555.1** | hypothetical protein | |
| **Medtr8g044120.1** | hypothetical protein | |
| **Medtr4g107630.1** | integral membrane HPP family protein | |
| **Medtr7g099870.1** | ion channel regulatory protein UNC-93 | |
| **Medtr7g114870.1** | IQ calmodulin-binding motif protein | NRE |
| **Medtr3g030420.1** | late nodulin | NRE |
| **Medtr6g091605.1** | late nodulin | NRE |
| **Medtr1g011540.1** | *LgHb7* | dNRE |
| **Medtr1g049330.1** | *LgHb4* | NRE |
| **Medtr1g090810.1** | *LgHb11* | NRE |
| **Medtr5g041610.1** | *LgHb6* | dNRE |
| **Medtr5g080400.1** | *LgHb3* | dNRE |
| **Medtr5g080440.1** | *LgHb1* | dNRE |
| **Medtr5g081000.1** | *LgHb8* | dNRE |
| **Medtr5g081030.1** | *LgHb2* | dNRE |
| **Medtr0693s0050.1** | Myb/SANT-like DNA-binding domain protein | |
| **Medtr4g128540.1** | Myb/SANT-like DNA-binding domain protein | NRE |
| **Medtr7g103390.1** | Myb/SANT-like DNA-binding domain protein | |
| **Medtr0003s0210.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr0111s0050.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr0115s0030.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr1g041915.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr2g048203.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr2g450630.1** | Nodule Cysteine-Rich (NCR) secreted peptide | |
| **Medtr3g006650.1** | Nodule Cysteine-Rich (NCR) secreted peptide | |
| **Medtr3g031320.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr3g436500.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr4g059900.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr6g090485.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr7g071310.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr8g039130.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr8g064070.1** | Nodule Cysteine-Rich (NCR) secreted peptide | NRE |
| **Medtr8g069310.1** | Nodule Cysteine-Rich (NCR) secreted peptide | |
| **Medtr2g041100.1** | Nodule-specific Glycine Rich Peptide | NRE |
| **Medtr5g069890.1** | Nodule-specific Glycine Rich Peptide | NRE |
| **Medtr7g098180.1** | peptide transporter | |

| | | |
|---|---|---|
| **Medtr8g090375.1** | peroxisomal acyl-CoA oxidase | |
| **Medtr1g080020.1** | phosphoenolpyruvate carboxylase-related kinase | |
| **Medtr6g060570.1** | phosphopyruvate hydratase | NRE |
| **Medtr2g085065.1** | Serine/Threonine kinase family protein | |
| **Medtr5g027460.1** | transmembrane protein, putative | |
| **Medtr4g088350.1** | triose phosphate/phosphate translocator | |
| **Medtr4g094335.1** | *SEN1* | NRE |

**Table S4. Accession numbers of genes mentioned in this manuscript.** Heme and JA related genes are provided in Tables S1 and S2, respectively.

| Abbreviation | Gene name | Gene ID |
|---|---|---|
| *NIN* | *Nodule Inception* | Medtr5g099060 |
| *NLP1* | *NIN-LIKE PROTEIN 1* | Medtr3g115400 |
| *NLP2* | *NIN-LIKE PROTEIN 2* | Medtr4g068000 |
| *NLP3* | *NIN-LIKE PROTEIN 3* | Medtr2g099350 |
| *NLP4* | *NIN-LIKE PROTEIN 4* | Medtr1g100970 |
| *NLP5* | *NIN-LIKE PROTEIN 5* | Medtr0022s0430 |
| *LgHb1* | *Leghemoglobin 1* | Medtr5g080440 |
| *LgHb2* | *Leghemoglobin 2* | Medtr5g081030 |
| *LgHb3* | *Leghemoglobin 3* | Medtr5g080400 |
| *LgHb4* | *Leghemoglobin 4* | Medtr1g049330 |
| *LgHb5* | *Leghemoglobin 5* | Medtr5g080900 |
| *LgHb6* | *Leghemoglobin 6* | Medtr5g041610 |
| *LgHb7* | *Leghemoglobin 7* | Medtr1g011540 |
| *LgHb8* | *Leghemoglobin 8* | Medtr5g081000 |
| *LgHb9* | *Leghemoglobin 9* | Medtr5g066070 |
| *LgHb10* | *Leghemoglobin 10* | Medtr1g090820 |
| *LgHb11* | *Leghemoglobin 11* | Medtr1g090810 |
| *LgHb12* | *Leghemoglobin 12* | Medtr7g110180 |
| *NCR158* | *Nodule-specific, cysteine-rich 158* | Medtr7g027180 |
| *NIR* | *Nitrite Reductase* | Medtr4g086020 |
| *nsHB1* | *Non-symbiotic Hemoglobin 1* | Medtr4g068860 |

## Table S5. Oligo DNA sequence for PCR primers used in this study

| Primer name | Purpose | Sequence (5'-3') |
|---|---|---|
| NF7812-F | *nlp2-1* mutant screening | GTTTAGTTGTTTGTCTTGTTTG3 |
| NF7812-R | | CAAAGTATATGGTGGAAGTAGC |
| Tnt1-F | | TCCTTGTTGGATTGGTAGCCAACTTTGTTG |
| NF8520-F | *nlp2-2* mutant screening | GAGCTGCATTAACAGAGATA |
| NF8520-R | | CTGACTGTAAGCAGCTTGAT |
| qMtEF-F | qPCR | CTTTGCTTGGTGCTGTTTAGATGG |
| qMtEF-R | | ATTCCAAAGGCGGCTGCATA |
| nsHB-qF | qPCR | TAGCACGTAACCCTTACCCTTCAG |
| nsHB-qR | | GTTTTTGAGCTGATGGAGCAATCTC |
| qMtNIN_F | qPCR | GCAAAGAGCATCGGTGTGTG |
| qMtNIN_R | | GCCAATCTGTATGGCACCCT |
| qLgHb1-F | qPCR | CTGAGGGGATTTATGCCAAA |
| qLgHb1-R | | TTCTCCAGTTGCTCGGAGTT |
| qLgHb2-F | qPCR | TTCAACTTCGAGCAACAGGA |
| qLgHb2-R | | TGCCAATCCATCATAAGCAA |
| qLgHb7-F | qPCR | TGTTCAACTCCGAGCAACAG |
| qLgHb7-R | | TCCCAAATTTAACCCATTGC |
| qLgHb8-F | qPCR | GAGAAAGCACCTGCAGCAAA |
| qLgHb8-R | | TGTGCAGCTGAATCGCGTA |
| qNLP2-F | qPCR | GCACAAGTTTGGGCACCTAT |
| qNLP2-R | | AGGTCCATTCAGGAACATGC |
| qNiR-F | qPCR | GGACGATGGAGTTGAGAAAG |
| qNiR-R | | ATGTTTTTGGACGAGTGATG |
| pNLP2-F | Construct pL0V-PU-41295 | GAAGACTTGGAGAGAGTTGGGTCGTGAAATAT |
| pNLP2-R | | GAAGACTTCATTGTTCAGGTTGGTGAGAAAATTAG |
| NLP2-F | Construct pL0V-SC-41308 | GAAGACTTAATGATGGAAGGCCAATATTCGTC |
| NLP2-R | | GAAGACTTACCTTCAACTGTCACTGCCGCCATGTATA |
| PUB-NLP2-F | Heterologous expression | TTGTTGATGTGATTACAGTCTAGAATGGAAGGCCAATATTCGTCAAAGG |
| PUB-NLP2-R | | TGGAACATCGTATGGGTAGGTACCACTGTCACTGCCGCCATGTATATTG |
| PUB-NIN–F | | TTGTTGATGTGATTACAGTCTAGAATGGAATATGGTGGTGGGTTAGTG |
| PUB-NIN-R | | TGGAACATCGTATGGGTAGGTACCGGAGGATGGACTGCTGCTGCTG |
| NLP2-attB-F | Truncated NLP2 for EMSA | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTACAGTCAGGCAACAAGTGG |
| NLP2-attB-R | | GGGGACCACTTTGTACAAGAAAGCTGGGTTCAACTGTCACTGCCGCCATGTA |
| pNCR158-F | Amplify *NCR* promoter for CRISPR construct | gcatgcctgcaggtcaacGTGGTGTTTACCATTCTCTG |
| pNCR158-R | | ggtccttgtaatccatctagaAATTTTTTCTCTTTTGCATGTG |
| NRE-gRNA-DT1 | CRISPR dNRE Knock out | GAGAATCTATTGGCTTGTGG |
| NRE-gRNA-DT2 | | CAAGCCAATAGATTCTTTAC |
| NIN-gRNA-G-DT1-F | CRISPR NIN Knock out | TAAGGATGGAATATGGTGG |
| NIN-gRNA-G-DT2-R | | TTGATATGCAGCAACTCAT |

## Supplemental References

22. D.G. Barker, T. Pfaff, D. Moreau, E. Groves, S. Ruffel, M. Lepetit, S. Whitehand, F. Maillet, R. M. Nair, E-P Journet, Growing *M. truncatula*: choice of substrates and growth conditions. In *Medicago truncatula* Handbook, ISBN 0-9754303-1-9 (2006),

23. S. Marillonnet, R. Grützner, *Curr. Protoc. Mol. Biol.* **130**, e115 (2020).

24. K. Xie, B. Minkenberg, Y. Yang, *PNAS* **112**, 3570 (2015).

25. H.L. Xing, L. Dong, Z.P. Wang, H. Y. Zhang, C.Y. Han, B. Liu, X.C. Wang, Q.J. Chen, *BMC Plant Biol.* **14**, 327 (2014).

26. M. Padidam, Y. Cao., *Biotechniques* **31**, 328 (2001).

27. H. Liu, Y. Ding, Y. Zhou, W. Jin, K. Xie, L.L. Chen, *Mol. Plant* **10**, 530 (2017).

28. M. M. Bradford *Anal. Biochem*. **72**, 248 (1976).

29. T. A. Larue, J.J.Child, *Anal Biochem.* **92**, 11 (1979).

30. K. J. Livak, T.D. Schmittgen, *Methods* **25**, 402 (2001).

31. E. M. Lodwig, A.H.F. Hosie, A. Bourdès, K. Findlay, D. Allaway, R. Karunakaran, J.A. Downie, P.S. Poole, *Nature* **422**, 722 (2003).

32. J. Liu, T. Bisseling. *Genes (Basel)* **11**, 7(2020).

33. V. A. Benedito, I. Torres-Jerez, J.D. Murray, A. Andriankaja, S. Allen, K. Kakar, M. Wandrey, J. Verdier, H. Zuber, T. Ott, S. Moreau, A. Niebel, T. Frickey, G. Weiller, J. He, X. Dai, P. X. Zhao, Y. Tang, M.K. Udvardi, *Plant J.* **55**, 504 (2008).

34. D. Winter, B. Vinegar, H. Nahal, R. Ammar, G.V. Wilson, N. J. Provart, *PLoS One* **2**, e718 (2007).

35. R. C. O'Malley, S.C. Huang, L. Song, M.G. Lewsey, A. Bartlett, J.R. Nery, M. Galli, A. Gallavotti, J.R. Ecker, *Cell* **165**, 1280 (2016).

## Appendix

**1. The promoters of the 55 most downregulated genes in nlp2 nodules that were used to identify the dNRE motifs shown in Fig 3A and S13A.**

```
>1_Medtr0003s0210
GATACAACAAATTTTCATGATCCAACGATCCTCTCTTTAAAAAATAATGATGAAATGGTT
TAAATTAAATAAATTAATAATGTTTACGGTGGACCACTTACAATGTTAGTACACTTAAGC
CCTACAAAACCTAAACTAAATTCTAAAAGCAAACATGGAAGAATTTTTGGGTGTTACATA
CATCGTCGTCAATCATAAACTTTTTGGGGTGATAGTCATTTTTGTCTTTAAATGTGCAAT
GAATAGCCATACTAGTCCCTCCTTAAAAAAAATAGTCACAATAGTCCCTGTATGTATCGA
AATTTTAAAATAGTCATTGAATGTGCACTTTGCTAGTCAAATAGTCACTGACGTTAAAAT
AGTCCCTCGATGTTGACAATAACTTTTTCTTATATGGATTGATACGAAAATAAGTAAATA
TTTATAGGGATTGATATTATACTAAGCAAGTATATTTAGAGACTGATATGACAATATTGA
CGAAATATTTTTACGCGAGAAACTATTTTGACTAATAAAGTTCACAATCAAGGACTATTT
TGCAATTTCAATACATTCAAAGACCATTATAACTACTCCTTATATATTCAAGGAACAAAA
AAATTATTACCCCTTAACTTTTATGCTCTATGGAGGAATTATTAATGAATTTGAAATCTA
AAATAAAGTATCTATAGATTTGCTATAGGCTAAGTCTCACATTAAATCAAACAATATTTC
TCCTTTTTTTGTTATTACCCTTTTCTATCATGAAAGGTTATAAAATTGTTAAAATAAGAA
TAACTTATTATGATTGGTAAGTTTTTTTTTTTATGTAACATCTTTACTATAAATGCGAAG
GGACAAACACAATACTTATATGTAATTGTTATTTTAATTATACGATTTTTTTTTACAACT
CCGTAAGGGCTAAAATCTATATAATGTATGTATGAAGATGGTATTATAAATTTTAAATTG
```

```
ATAAACATACAATATTTATTACTTGTAAAGGGGAGAAAAT
>2_Medtr0111s0050
AAATTCATATTAAATTCATGTTTTGTTCAAAAATTCTAATTTATTTTGGGAGAGATTATT
ATAATATTATGGATTTTTATGCCAAATTTCATTCAAAAATATGAATTCTACATATGAGTT
TACTTTAAAATAGGGACCAAAAGTGAAGATGGAAAATTTTTAAGGGACTAAAAGTTGAAG
AAATTTTTTAGAGGGACTAAAACGAAAAGTTGACATATTTATAGGGACGAAAAACATATC
TAACCCTAAAAAAATTAATAATAAGTTTTTTGCTTATAAGAATTTGAGTATAGCTTTTTG
TGTTTTCTTTGGAATAAATTTTATAGAAAGTTGTAAAAACAATTTTCATTGGCTATTGGT
TATAGAAAAAATGGATGAGAGAGAAAGTTGAATGTAATTTGCATTTAATTTCATGAGAAT
AAAAAAGGTTTTTCTTGGAATAGATGGTATTTAAAAAAAAAAAAATTAAATATGATAAAAG
TTTGTCTGTTTTTGCTTACATTTTAGGAAAAAAAAGTTTTTTTTTGCTTACATTTTAAGA
TGATGTAAAAAGTTTTTTTTTTGCTTACATTTTAAAAGTTTTGTTTTGCTTAAATTTAAA
AAAGAAGATTAAAGATTAAAATAGTTTTTTTTTTTTTTGCTTATATATAATTCCTTATAT
ACTTCTATTACTATCTATGTTCATATCTCCAGGTACCTCCAAATTAACACTTGGCGCTAC
ATAATTTTCCCTTCTCTATCATGAAGGGCTAGGAGAGAAATAATAAATTGTAGTAATTAA
AATTTCAAGAGACATTATTAGTGAACTTTGATCGTCATCAAGAGACATTTAACGAAAATT
TCAATTTTCTTTCTTGCAGCCCGAAAGTGCTAAGATATATTTAATTGTAATTAAAATACC
ATATAATATTTACTTATAAAGATTCTAAACAATTTAAAATTTATTAAATCCAAAGAAGAA
AAATATGGCTCAAAATCTTAAATTTGTTTATGCTTTGATA
>3_Medtr0115s0030
AATTAAATAAAAATTAAAAGTATTAAAAGTTATTTTATGATTATAATTGATTTTTTTTTA
GAAAAAAATTGACTCATTTATGACATTTAATGTTGTCAATGTATATTACATTTTTATTGG
TTGATATCATGAGAGGGTAAAATACCCTCTAAATAAGATAGGGTAATCTAGAAAAAAAAC
CCTCTTATGTTAAAATATACACTTCATTTTCATGAAAACTATAGTTTATACTCCCTTCCC
TTAATTATAAGATTCTTGAATTACAACCTCCTCCCTTAATAATTAAATATGAATTACACT
TTAACCTATTTTCACATAAATAAATACTTGTATGATGTTGAATTACCATTTAAGAATTTT
TTTTTTTTTTTTTATGATTTAAATGTCAATGATAAGTAAATTTAAATATTTTGCTATTAA
CTTTATAATTTAGAGTATAAAATTCACTTTTATCAACTTTAGTAATTTTTCACATATTTT
AATTTTATTTTGTGTTATTTCATATTTTATAATTTAAAACTATATGTTAATGAAATAGTG
AATAAAAAAATAGTGAAAAATATGCATTCGAACTTTTAATTATATTAAAATAGACTCACAA
TACTATTTTTCTTCGAAGTTTGTTAGATTATATTTTTTTGCTAGATTATTAGATATAAAA
AAAAATATTAAGGGAGTAAAATGTAGATTTTAAGATAAGAGGGATGGGAATTTAAACCAA
GCTTCTTTAGGGGAGGGGAGTGAAATCTTTTGTAATATGTTTTAAAAAAGAAGGGAGGGG
AGTGTATATTTTAACATAAAAGGGGAGGAAAGTGTAATTCAAGCATCTTTCAAGGGAGGG
GAGTGTAATTTACTCTTATATATATAGTTCTCTATATTTAACATTAATTATAATTTTATT
AATGGCCTTTGGCAACTCTAAAATTCTAAAAACTATTTATTTGATGTTAACAAAGCATGA
CATTTTTACTTATAAAATTACTAAAACGATTTTGTTTATT
>4_Medtr0693s0050
TACACTTGCGAAAATCGCTATCACCTATTAAGCCTAAATTCTTAAAATTGTTTGAAAGTT
TTCACTTTTTTTTTCATTAAGTTTTCAATTGAGTTTATTGGAAAAGAATTGAGTTTCTTC
CTAAAAAATTCAGAAACAAGAATTAGAAGAAAATATGTTAATGATGGTAGTTTTTTTATT
TGTTGATAAATGTTTTGATTTATATATATTTGTTGCGTTTTTAAAACAAGTGACATACCT
TACCAGTTATCAGCAACCAGTGAAAATATGAGGAAGTTGTTTACACACAAGTGTGTGTGT
GTGTGGGTGGTTGAATTGAATGTGTTAAGCTTGATAGCAGCACGTTATGAGTTATTATTT
TTTTTTTAGCAATATAAGGTGACATTGGCATTGCTTGATTAAAGAGATGTAGAGGAAAAA
TATTTATTACAAATTATTTTTGTTATAGGAAATTGATCCATTAGTTTGTTTTTGATAAAG
TTATAAG
>5_Lb2_Medtr1g011540
GATCCCGTACATGAATCTCCTTCCAATGATAATTGTGTGGAGAATGTCATGAATGCACAA
GCGAATATTCAAGTTTCAGACTTTGAGGAGACTCCTTCTGATACTCACATTTTGGTTGGT
AATACATCAATTCCTCTGCATTGTGTGGCGCCTATACAAAGAGACATTGATGTTATTAGC
GCAGACAAGGTCTCAGGCGCCCTCCTTGTTCTACTATTCTTCAAAGAGATAATGATTTAA
TTGAGCAAGATTTGGCTTTGGACCGTTAACATCATTCACCCCTTATCTATCCAACAATCG
TGAGAAGAAAATCAATCATTGCTCCGAAGGAGGCCTATAACACTTGCTCCAAGGGGGAAC
ACAAGGGCGGTTGAGTTGTACCATTCAACCTTCTTTTTCGTTACTTTTATTTTTGCTTTA
TTTTGGTTTTTTTTAGGGTTTTGGTTTGGCCTCATTTTGTATACATTTTCTTCTTTTTTT
TAATAATATATCTTATAAGTGAGAGAGAGAGGATGAGGGTTCTTTGGAGGTGTTTACCTA
ATTTGGGATCTTCAAATTTTCCACTTATGCTCTCCACGCTCTTAGTTCACAAAAATAAAA
CTATATAACTCCACAAAAATATAATTAATAAAATTATTTTAGTAAATGAAACTCATTTTG
TTATTAAAATCAAAATAACTCACATTATTAAAAATTGGGCACAACTCAAGTGAAATATTC
ACCATGAGACGAGGAGAATAATTATTTTTATTAGATTAATTAATAATTACATTGATTAAA
TCAAGTTAAACAATACTTTCATTAATGATTAGTTTTTTATTTAAAATATTTATTAACTTC
AATGAAAACTTGTTAGATTAAGTTTTTTAAATGATTATTGTCTCTTTAATAATGTCAACA
GCCATTTCCACAAGCCAATAGATTCTTTAAAAAAAAAATCATAAAAAATCAATTGCATCT
CTTTTCACCTCCAAGTTTTCTATATAAACAAGTATTGAAT
>6_APC11_Medtr1g026070
TGAGTTTTGATTTGATCAATATTTAAGAGAATTAAACTCAAAAGACTAATTGCATACTTG
TCTTATAATTTAAAGGATTACCAACACAACAATCCATTTGAGTTTTGATTTGACCAAATA
TTTAAGAGGATTGAACTGAAAAGACTACTTGTGTATTTGTTTTATAACTTAAAGGATCAC
ACCACAACGACTCATTTGAGTTTTGACTATGTCAATATTGACAGCTCAACTTAAACTCTT
TATATTCATTTTTTATTATAATATTTATTTTTCTTTAAGTTTTCATAAATTCCATCCAAT
CAATGTAATCATATTCTAGACACCATTAAAAAGTAAATAAGAACAAATTTTCGAGTTGAA
ATTCAAATATTAATTGGTTACATGATGAGAGTCAAATTTCTTTCACTAAATTCAACCTTA
CTTGATTAGAGAAAGTCTTTATAAATTTTAACTTGAAGTACAGTCTAAAATAAAATAATA
AAGATATAAAAAAGTTTATAAAAAGCTTTTCAATTTTGAAAAGTAAATAAAGTTAAAGCA
TGATTAAAATACAAAAATGCATGATTAGTTTAAAACAGAGTTGCTAAGGTGTTATAGAAT
GAGAGAAACAAGCGGCAGAGTGTAAGGCCTTGTCCCAATAAGAAAGCAGATTATTACCAA
AACTCAAAGCGCACACATATACATACACAAATTAAAAGAAGGAAAATAATTCTTCAACAC
CCAAAACCATTTTACTTTACCACTAGAGAGAAAAAGAAATCAAGAGAGTAAAATAATACA
TATGGTGAATATATGATAATAAGAAAGAGTGGTAAAAGGAAAATGAGATGTTGAATATTG
GTATGAAAAATAAAAGGGGGTGTCTATGTGAGTTTAATTTAGCTGGTATGTACAATGCAT
```

```
AATATATAGTTAAAGGTATGAGATTCAAATCTAAAAAAAATAAAAAAATAAAAGGGTGTA
CAAATATAATATAATTGCTTTTAAAAGAACACCAAAATTA
>7_Medtr1g041915
TTTCTTCTAACCATTATATTTACTTTTGTAAAAGAACAAAAAAGGATAACTATTTCCATA
ATTGGTTTATGTTTAGTTTTCTTTCTTTGTTAAATAAAAAAAGGTTAACTTTTTCCTACA
GACGGCGCCACTGATCTGGCCTGGTCAATAGAACGACATCCGGTATTATGAGCTTGCTTT
GTGGTAGATGGTCGGAACAACAGGCGGTGGTGTACTTCCAAGCACTTTGACGCTGAAGTT
AGAGTTTGAGAAACGAAGAAGGTTAATGGTTAAATATAATGTACAATGTTATATAGTGAA
GGTGGTGCTTTCCAAGTTTATATGGCTAAGAGTTGTTGGTTAATAGTTACAGATGGCGAA
CAAATCTCCTTGTGATGGAAGAGTGTGTATAATGTGATTCTTCCAAAGTGAATTAGGCCA
TATTTTTTGACCCTGGGTTTTGAGAGTATGCTTTTGAGCAGTCTGGAAGACACTCAGTTA
CTACATATATCTATTTTAAAACAATATTAGTCATCTTTTGTGATTTTATCCTTTCACATA
AAAAAAATTATAAAATAATGGAAATTTAGTAAAAATAAATAAATGATTTCTTATGGATGG
TCAATTTTTATTTGATCCTTTTACAAAGAAAAATGATTGGTAGACCACTCATATTTCCAT
ACCACTCATACACCCCCACATATTTTAGAGGGGAGAGATGGATGAAAGAGAAGAGAGACA
AAGTGTTTGTGGGATCTAGGACACGTGTCACATGTTTATGTGGAGGTTTAAGTGTGTATG
CCACCTATGGGTGTACATGTATCACCACTCTTTACAAAGTATGATGACTTACAAAAAATA
AAACATTATTATACTTAAGAGACATTTAATAAATATTATAATATTAGGATGACTTTTGAA
TCTCCATAAAGGCTAAAATATACTTAATTAAGTTATGGAGAGCATGTGTTTTATAAATTA
TTTAGTGATAAAAAATTTCTTATACGAAAAGGGGGTGAAT
>8_Lb5_Medtr1g049330
AAGAAAAAAAAAAATCCACAAAATGGATCCTACCCTAAAAGATAAAGCAAACGATGAATAC
CAAGAGGGTCTAAACTGGCGTATGAATCATCTGCCAGCACAGAAGCGATTCAAGATCCTG
CAACTCCCTGCCAAACAACAACAAGCTATAATTTCCCTTTCGAGCCAAACCATCTACAAA
TCCATTAAAACACGATGCCAGCGCACAACGAAAGAACCAACCGCCAACATAGATAGATTC
AAGAACCAACAGCCCTTTCAAACCAAAATTCAGAGCGATGAAATATCCCATCAATGTAAA
TTGGTTCGCTGATCATTGCGGCAACATCGATGATGTTTAACCGAATCTCAGATTGCGACA
AAATCAACCATAAACATCAATGCGATCAACAACCGTGAACCACCTATAACAACATATTGA
AAATTCGACGGTGGTAAACAACCATCGTGAATAAGGAACCAGAGACCGATTCACGAGAGG
ACAGACACACGGTGTGACCACAAAGCAGCACAACTGTCGTAAACCCCATCCCTAGGTCGA
CGCAAATCAACGTGTCGCAAACGCGAAACCAAGATTAGACCATAGAACAGTACCGCAAAA
ACGCGGCGGCAACAAGAAGCCACCAAAAGAACTCGATTGCACCACCAAAAGCCATGGTCT
ACATCGTGGCGGCCAGAATTGAGGTGCGGCAAAGGGTTGCGGCTGTGGAGGTTTGTTAGA
GAGAGGGGAGCGGAAACGAGAAGAAAAGAAAGGCGTCTATGACTATAGCTTTATTAGGAA
AAGGTTAAAAAATCAAATTATGAGCGGGTAAAAAATTAAAAGTCGATAAAATTTAAATTA
GATGTTTACTTCTTTTTGTTTTTAGAATTATTCATTAGTTTAATTAATTAATTAATTGAT
TATGATTTTGAAAAGATGATTGTCTCTTAATAATACCAATAGCCATCTCATAATTAAAGG
CCAATAAATACTCTAATTGAATAACCACGAGAAAGCAACT
>9_Medtr1g069825
TTTAATATGATGTTGAAAGTAATTGTCAATTTTTAAGTTGTTCAAATAACACCACTCTTC
TTCTCGTATTGCTCAACAAATACGACTCACTTGTAGATGGAATTTCTAACACTCACACAT
TCAAGAAATCAATGATCATTTAATCAAAATCATACACTCATGCTCCAAAACTCAATATTT
TAATTTATAAAATTAAAATATGTATTTTCAGATCAATTGATATTGATTGGATGATCATCA
ATGTCTCGAATGCAAAATTCTCCATCATGTTAAAAGGATTAATCCAATTTCAGTGCTCAA
CAAAAAATACAAGAGTAATACCCTTTAAATACAAATACAACTTGGTAAAATTCTCTTTCG
ACACCTCTTATCTTCACAAAATACAATAATATGAGTAATAATATTTACACAACCAATTTT
AAAATGTTTTTAATAAAATACAACGTAAGTACGAGAGACACGTCATAAAATAGATGCATA
ATTATTATTTGAGAAATGATATTTGTACAACCAATTTCTAACAATTTTTTGACAAATTTC
TCTCTCATATTCACGATATATTCTTATTTTCTCTTTACCTTTTTCTCTCTATTGATTTTG
ACCAATCAAAATAAAGAAAAACAAAGTTGTCATAAAAATTATTTAGAAAAAGATGGACAA
ATATCATGAAAATTATTTAAAAAGAGATGTACAACTATCGCTGCTCTTATTATTTGTCCA
ACAACAAATAGTAAAGAGGAAAAGTGGAAAAGTTCAGTCACCCCTCCCATCAATCTCATC
ACATATCCTCAATTAACATAAAAAATCTTCTTGCTTCACCACTTGAAGAGGTCCATACCA
TATACCATCTATCTCTTCAATTTCTCTCTCTACACAAAAAAACAAACTAGTCTGTAATAT
CAATCAACCAAATCAGAGTTTGTTTCCTTTTCATTTTCTTTCAAGGTAGTTTTTTTTTCT
AAAAATGAAAAGACATTAAAAAATATTATTATTATTTTGT
>10_PEPC_Medtr1g080020
ATTTAAAAATATCGATCAATTTTTATACACCAAAGGACATATTTATAATTTTTTAGAAAT
AAAGGAACTTGCGATTGTTTTTTTTTTAAGGAACTAGCATTGCTTTTTTTGTTCTACAGA
TTTTTTACTAACTTATATTGCTTTATTTCCTCTTTTTTTTCTTAGAAAATTGTTTTATCTA
TACTATATATAAAGAAAACAACCATTTTCAAGCACTTTTGTGCTGATTTTTCTCTTTCCA
AAAGTGCCGTCAATTTTAAATACAAATAAATAACACATGTAACAAATATATAAGAATAAA
TTTAAATATTAAAATAAAAGTGTAACAAACACAAATGTCTAATTTTTTCTTTTATCATTT
AAAAAGTGTAACAAATTAAATGAGTGTATTTATACTTACTTTTTCATTATCTTAAATCTT
AATTATACTTTTAAACAACAACTTTTTTTTCTAATAAAAATTGTTGTTGGTGTCATTAAA
AAAATAAAAATGTAAATTATATTTTTTTGTTATATCGTTAGTTTCATTGAAATAGATAAT
CATGATTAGTTTAATTTGTTATAATTTGAAAGCAAAAAACATTTATATTTTTACTTTTAA
TCAAAATTAAAATTTCATAAAAAATACCGTTGATATTTTTCAAACGTTAGCACGATAAAA
AGAGCGAAAAACGGATACGTTAGTGCCCACGTGTCCGTCTTTTCGCTAGTTTGTATGATT
TAACTATTATATTTTTTCAAAAAAAACTTTTATTCAATTACAATCGATCAAAATATTAGG
AATTGGTTTGTTGTTCACTTGTTCCAAATTTCCTAACCTTTTTTTCCGCAGAGAATCTCT
ATCTTTCTGTTTTTTTTTTATTCTTTTCATATTTATAAAATAAAATATTCAAAAGATAAAC
CCTATTTTATTTAAAAAAAAAAAAAAAGAAACTAATAAAACGGATGTAACGGTAGTTAGTT
TTTTTGATAAGCAAAATATAAACAGAATAAACCCTCCCGT
>11_Lb8_Medtr1g090810
CCTTGAGTCCGACATCACGACTTTATGCTACACTGCAAGATGCTCCAAATGCTAACTGTT
AGAAGCTTGAGAAGCCAAACCAATTTTTTGAACTATTACTATTAGTTGAATGAGAGTAAG
AGAGTTGAGACATCCTAAATGTGTTTAAAATCCTCTTAAATTCTGCTCGACTCTATTTAA
GATCAAGTCTATTAAAGATGAGATTGAGTTGATTCAAAATTTAATTAAAAATAAAATTAT
AAGTTAATCATTTGGTTATATCTAAAACTAAATAGTAAAATATCGTTTCAAAACAAAAAG
TAAAATATCAATATGTGTTAGTGCAAAGGGACTCAATATTTTAAGTAGGTGGTTTGAGGT
```

```
TCGAGCCTCAACTCTTTTTTATGTATTGAGAAAACTTTATTTAAAGATGAGAATTAGAAC
TTTGATCATACCTTTTACAATATCTTCACCAGAAAAAGGAAAAACATCAATCTTAGATGA
GTACTTTGCTAAAATATTACCAACTATTAACATAAAATTCAATATTAATATTTTATTTAT
AAATGCTTTACCACAAACAGATAGAATTTAACATATGAGAATAAGACATGTTTTTGAGTT
GGATTGAGTCTTTGCTAGCCTCTAAATAATTAAATAATGTGCTCTTAGCGTACTTTTTTA
AGTTTTGATAGCTACCATTAGTTACATGAAAAACTGAAAAGAATCATATTCAAGTTTCAC
AACCACAGTCTAACGCAAGTGATTAATGTGCGGTGTCTGTTGTAAATCTTAGAGTATTGT
GTTTGATTCTTGATTAATATTTAAAGTTTTAAATGTTTTGAATTATTTAATGAGTTTAAT
CTTTTAAATAAATTATTTAAGTATATATTATCTCTTAATAAAGGTTGCCTCCTCTCAAAG
GCCAAGAAATACTTCAAATAGCAATCACAATAAAACCAACATTAGCTCTTCGTTGTCTCA
AAGCCTTTTTTATGAACAAGTGTTAGATGTTGCGTAAATT
>12_CHUP_Medtr1g093850
TGTGACACAAATAATAGATAAAAAATATTTGATTATTTTTGGTCAAGAAACCATCAATTC
TTTAATCATTAAAAAAAATTAAATTATATAAGAGAAATTATAGAAAATTCATAAAATATA
AGAGAGCTGTAGAATCAAGTAAATTATTTTTTGTAAGCATGCAAGGTAAGCCTCATCACA
TTTACTCTTTCGTTTTCTTTCTAAAATCTATATTTCATATAATTGAAGTGGTGAAAGAAG
CATATAAATAAAAATAATCAAAACATTATATTCTTTGGGAGATTGTAAACAATAACAACA
ACGACAAGATGGGTTAATTTTTTTATTTAAAAAAAAGTTTGGTTGATTTTTAAACCTAAA
AAAATAAATAACCTCAATATTATTAGGAGGACGAAAAGACAAAGGCATCAAAATAAAATG
ATGGAGAATAAAAATAAAATGATTAAAGCATATTCTGGTCATTTACAATTTCCGTGCGAT
CCAATTCACTCCCGACAAAAAATTAGTCCATCTGTTACAATAAATCAAAAACTAAATTTA
AACCGTTATAAAAATGTAGAAAATAATTTTAAGTATTTTTTATATAGTCGCAGATTAATT
GTGAAGGACCAACAACAATAAAACAAACGAGGAAACAAACCTGTCTTTGTTCTTTTCCTA
ATAAAAGAAACATGTCTTTGTTTTATTCTAGTTTTACTATAGAAGGTTATTTTCAAATTT
CAATCAACGTTGTTCCTTTGTTGTATTTAAGGGTAACTCTAAATTCTCCCACTTCTCTCA
CCAGATCTTCTTCTAAACATATTAATATTAGTCCAAACCCATTTGTGATTTTTTTGTCAT
TAATTGTTATGTGTCTCTTTGGTATCAAACTTGTGTCTGTGACTATGAACATGTTTTCAG
TTTCTCACTGAAAGGTGGTGAAGTTGAAAATTTAGTAGTATTTTCCTATGAATTCTGCAT
GTAAATGAAATTTAAAAATCTGGATTTTTCCTTAACAATT
>13_Medtr2g041100
AAGCAAATGCAACCCAAGAAGAAGAAACTCAAGCTGAACCAATTGTTGGGAGAAAGAGGA
AAAAAACAAGTATAATTTGAAAAGATTTCGATGAAAAGGAAATTACTAAAGGTGTGTTTA
GAGCTGTTTGTAAACACTGCAAAGCACAATATACTACTGGGACGGTAGGATCCAGCACGA
GTCAAATGACCACTTTGTCAGGTCACGTGGTGCATCAGATGATCCAAAATCTTGCACCAC
TTCAGATCAATGCACAAGACTGGACAACAAACTCTTGCAGTATACTAGGAATAACCTTGT
TCCACCTCGAATATCATCAAAGCACCTACATGTTCCACTGCCTTGAAACATAACCAATTC
CCAAATAGTTGTCTAGTTGTTCCTATATATAAGGACACGATTTGAACCAGATAGGCGCCA
CCACTTATCTTTTGTTTAGTTACACAAAATGTAATTGAGTCATTCCTAGTTATAAAGGAC
TCACTTAAACCAGATAGGCATCCTTGCCTATCGCGGTTGTGTTACTCTAATAGTTTAAAT
TCACCTTGTGACTAACTATTCTAAAATTAAGGTGAGAAGATAAAAAACAATAAGAAGCCT
TACACTAGAACAACTACTATAAATAGGGGACCTCTCCCACAGGTACGATAACATCACATT
TTTATCACCTTACTTTTAACCATTTTCTACTTGCTAGTACCCAATCTTACTTACGCGTTA
GGTACGCCCCCATATTTTCATATTTATGCTTATACAAATAGCTATAGCTATATTTCAGAT
TAAATAAATTATCGACCATCTTAATCACTAGTTAGTTGTGCTAAAATTCTTCTTACCTCG
TTTGATAAATGTGTCCAAATTTAACCTTTGATTAGCTCAATAGATATTGGTAACCTATTA
ACCTCACAACGAGCTAACTTATAAATATATTATATTTTTTATTGTACAATATATATTATC
TAATCAAGTTCTTTCTATTCTTATTTTCTATATATATTTA
>14_Medtr2g048203
AAAAAAAATAATGAAAAAGAGTAACTTGATACCTATTATATTGTTGATCTTCATTTCATC
TTGTGTTTTTTTTTTATATTCACTTTTTCTGCAAAAATAAAGTAGAAGATGAGTTAGAGA
GGATTGAGTTAAAAAAAAATTAGAGTTTAGAAAAAAATTACAGTTGAAGATGAGAAGGAA
AGATGAGAGAAGTTAAGAGAAAAAAAATGAGAAAATAAGGAAAGATGGAATATAAAATGA
ATGAATGAACTTTTGAAATTTGAATTTTGAAATTGGTGAATAGTTGGTCAAAAAAGCACG
AGGTAGCAATAATGGCATGGGTAGAAACGTAAAAATGGTAGTAACTTACTATCATATATC
GTAAGTAGATAGATAGATGGGTTATAACAGAGAGTAAGGTTCTTGGGGAGATGATGCACG
CCAAACTACTAAAACTCCCGCCTCCTCAAGTTGGTGCGCTTTCTACTGAGAGGACGATCA
TGACACAAATGAGTGTCACCAGTTGAAGAATAAGATGGATGACCTGATCCAAAATGGACA
CCTAAAAAGATTCGTGAAAGATGATCATTGAGCTTGAGCTTGATTACCTCTGATTGCCAA
CAAGGTCACCTCCTTCCTGTGTCTTCAATGTCAAGTCGATGTTCACAACTAAGACACTAG
TCAGGGGCACCCACAGACATATTTCTACGTGTAAGTCGATGTCCTTGTCTTTTGATTGGT
AAGTTTTTATATGACCTTTTTACTAAGTGTGTTGAAGGGACAAACACAATACACTGATTT
AGTAGTTGAGAGGAATTTAATAATGAGTCTACCTGTAGAGACAACTTTTCTAACTTCATA
AGGGCCAAAAGAATATTTAATTAATGTAAGACATTTATATATTATAAATAGATAAAATCA
AATTTTATATGATACATGCATAGGGGGGAGTATGGAAGATATTGTCAAGTTTGTTTATGT
TATAATTATCTTTCTTTCCATATTTATTATTGCAACAAAC
>15_Medtr2g085065
ATTTTTTTTTTTTTTTTTTTTGTAAATGACTCAGTTATTTAGGGATGGATAGAATATATA
AGGTGTTAAATGAGTAGTAAGTGTTTCATTAGGTAACCTCTTACGCTCAAAATTCAAATT
TCATCTAAGAGACATGAACAATTGTAAAATGGTTAGTGTCATCGATATCAATAATAATTT
CTCCTCCAATTTTTTTTATTTTTTTTTAAATAACTTTGCTACAATGCATTTTGTGGTTCA
AAAGTTTGCAACTTATGTACAGAGGAAAATGTTCATGGTTACTTCCTCTGGTCTCGGACC
AAATATATGGGTTACTTAAACTTATCTATGGGCATTTTGTGTAGTGAAATTCATGGTTCC
TTAAACTTATCTCTTTTACATTTGTGGAATCTCTTCGATATAAAAATATAAGCAACAAAA
TTGTTGTGTTTGTTGGTCTCAAATATAAGCAAAGTTAATTATTTTTATTATATTTAATAT
TTTTATTTCAATGTCACAAAATAATTGTATAAATATTTTTTTTTTTTTGTGACATTGATG
TTTTCGAACACATATATTTATACACCTCTGCATTTTTAACACTCATTCAACACCTAATTT
TATTTCTCTCGTGTTTTTTCCTATCATATTATCAAATTTCATACCACTTTCTATATCTCT
CTCAACTTAAGATGTACACCTTCGTCACTATTATAAACAAAAATCTATATTTTAAGTTTA
CTAAATACATGATGTATCTGCTTTGTATTATAAACTATATACATAATTTATTTAATAAAT
TTAAACAATAAATTTTTGTTTATAATAGTTCTCAAAGTCGAAGGTATAGAAGTGCGTACG
AGTGTGCGAAATCATTTTCCAAATATATTTCTATTACCTTGAAGACAATGTTGATATCAT
```

```
CATCATTACCATATATAATAATTTACTTATAATTAATTGTAATTATGATCATGAAAAAGC
AATTTAAAAATGGAAATGAAGAAATACCCATTTTCAAAAT
>16_Medtr2g099790
ATAATATGAATATGTGAAAGAAAGTTGTCACAAAAATTGTCGAAAAATGATTGTACAAAT
ATAATTTCTCTTAAAAAAAATGTATACAACAATTTAGTCATTAAACTATCACACTCTATA
ATTTTATTCATAAACTATAAAAAATATATAAGTATTTCCTAAACCAACCTCACATTCAAA
CATCTAACACTAATAATAATGTTTACAAAAATTATGAAAATAATTGGAGCAGTTATATTA
ATTTTTATTAAAATTAACTAAATTAATTAGGCCACCTCTATTAATTTATTATTTTCAGCC
ATATTTCTCGATCATTACTTTACATCAAATCAAACAATTGAACCCAAATGTTAATAAACT
TTAATCAAGACAAATCGTTTAGGATAACTTGAGTTCAAATTCTTACAAAAATAATTTTTA
TTAGATTTTTTTTAACAAAAATTTCTATTAGATTTTACTTACCTCCCATCCGAATTATAT
ATTACTATCATAGTTCCATCTCCTAGAAAGAGATGGGTTAAAAACAAAATAAAAAGCCAA
TGACAAATAACTTGAGATGCTAATCCTTTATTCCTCCTAATTAATCGGTGACATCAATTA
ATTACCAATGATTAGTGAATGCGTGCATAGAAAGTTTATCACACAAAATATTACTGTACT
AATATTACAGCATGCAGGGGGTGTGACCTGTTGATGTCATCTGTGGCTATTTCTGAGAAG
CAAAAGCCATTGATCACAATCACAAATATTCCAGCAAGCAGGGGGTCACCAATATTTGAT
GTTACTTGTGGCCAGTTCTAAATAGCAAGAGCCAATTCAGGTCAACTTATCATTCTATTA
CTTACATTTTTATTTGTACAAATTCCTAGTGAATGTGAGCTTAATTAGTTAACAAAAATT
TGGAACCAATCAATTAAAAAAAAACAGCTTTGATAATGCCATTTGCCAAACACAAACTTT
ATAGTATTATAAATATATAACACCTGAGAAGTGATTCGGT
>17_PFK_Medtr2g100710
TTAAGAAAGCCTATTGGGAAAAAGATAATAAAAGGCAAAAAAGAAAAAACGCTAGACAAC
TTTATATACGTCATACTATGGCATTTATTTATTTTTGGTGAGCTCACATGACTTTATCTA
TCAGCATTCAGCAGCCAATACTCGATCAATCAATTTACGTTTGTTACTTTCCAATCATAA
TATAAAGTCAGTTTGTTACAAAACCTATTTTGTCTTTGTTAGTAAAAATCATTTTTAACA
GCAGATCTAATTTTTTTATTTTTTTTTATTTATAAGTTTTTATAGATAAGTTGGCTCTAT
AAGCTATTGATTGTTTGGTGGCTTTTGAACTCAAGTTAAGGTAAAAGCTCCAAAGTTTGA
TGATGTTAATTTGTTTTACAAACAAATATTTTCGTGAAAAAGGTCATTTATATTCTTATT
TTAAAACAAGTTCTAATAATTACTTTAGACCAAGAACATCAGGTGATTATTAAGTTAAAT
GTTTATAACAAGTTTTTTTATGTAATATATTAATCCTTTTTAAGTTATGAAATGTTTGCA
CCATTTCCCTCTACTTAGTTAAAAAATAATTGTGACAGTTTAATGTTAAGTTGGTAGGTT
TAAATTTGCCGCTCTTAAAATTATGGTTCATAGATGTATTATTTAATCTTGGGTTGAACA
AAATTGATATTACTATATTTATAATGCGTTTGATAAAATATTTATATGTAAAATTAAGTT
TAAATTCGAAAACAAGAAGAAAACGTGTTTGTACGCAAATAATCATATTTGATGTCATAA
ACATCTTTTCAATCTATCACCTAAATGTAAAAAAGTCATGTATGCATTTTTTAACGAAGA
TGGATGAAAAATGATAACTTTGTTAACATTTTGACTTAAAAATAATTTTATATTGCATAA
AAAACTTAAAATGAAGGTATTTTGTTTTTAATTTATATTGTCATAAAAAAGTACTAATTA
ATAATTTAATTATGATTGAGACAACAATTTAGCTAGAGAA
>18_Medtr2g450630
ACCATCATTAATTATTTATATTTTTTAAAAGTTTGTGTCTTTATCTACTCAAATAGGACA
AATTGAGAGGGTCTAAAATGGAGGGGATCCTAACTTGTGAAAATTACCCATTGGATCGGC
TTGCAGAGTAACTTCTCTTATATAGTTTAGAAATTTGGATTGTACTACATGCCTTTTCAT
CAATAAAAATATACATAAAACCAATCATAAATTCCCCCTTATGATTGTAAGTGATTACTT
AAGTGTTATGAAGGGCTTGTTCAAAAAAAAAGTGTTATGAAGGGACAAACATAGTATATT
TATTTACCTAAGAGTCATTTAATATTAATTCTAAGTTCAAGAAAACTTCTTTACTATTCC
ATACATGCTAAAAACTATTTCTTTAAGTAATAAAATCATGACATTTATAAATTATAAATA
TAGAAATCAACTTTATTTGATACATGAAAATGGGGAGAAATATGGCTGATATTGTCAAGT
TTATTTATGTTATTATTATCTTTCTTTCCATATTTTTCTTTGCAACGAACCTTGAAGGTA
TGTCTTTTTTATCCTTTGACAAAATTACATTTTTAGTCTCTTAACTTAATTTTAGATAAC
AGTTTGGTCCTTTATTTTTTTTTTCATTTCAATTTTGTCCTTTTGATCCATTTACATATG
AATTTTTAAGCTTAAAATTCATGATTTTATTTTCTTATAGTCTTTGAGTCTTCAAATCTA
TTTATCCTCGTCTATGTTTGCATAAGAATATTAGATTTGAAGCTTGAAAATGTATATGAA
AATGGACAAAAGTACCAAATTAAAATGAAAAAAGATAAAGAACAAAACTGTTACCTGAAA
TTAAGTTAAAGGACTAAAACTGCAATTTTGCCTTATCCTTTTTATGTTTTCTCTTATTCA
ATTTTTACGCAGTTTTATCCCATTTTATTACAATACTTTTAATTAATTTTGTATTATAGC
GGGTCCTATATGTCTTGAAGATTTTGATTGCCCAAAAAGT
>19_Medtr3g006650
CCAGGAGGAAAACTTGACGTTTTGTGTTAAAATTTCCAGTTGGCATATTCATTGGCATCA
AAAAATAAACTAAAGTGGAAAATGGATAGAAAGCGAACAAATTAAAGTATTTGTTGTTTA
TTTATTTGGATCATTTTGGCTGAAAAATTTGTATAGTTTCCTTAAAAAAAAAAATTGTATA
GTTGTCGCTCATTGGTTCACGCCGCAGAACTAATTAATATATTAATATATATATATATAT
ATATATATATATATATATATATATATATATATATATATATATATATATAAATAGATAGAAGAAT
TTTGTTACTTAAAAAAAAAATAGAATTGTAATCTATTCTATATATAAAGAAAACAACTCA
TTTCAGCTCTTTTCGACTGGTTTTTTTTTTTTCTTCTAAAAATGCCCTCAATTTTTAATA
GAAAGAACACATGTAATAAATAGATAAGAATAAATTTAAATATTAAAAAATGTAACAAAC
ATAATATATATATATATATATATATATGTCTATTTTGTTTTTCCTTTATTATTTAAAAAA
TGTAAGAAACTAGACGAGTATATTTATGCTTACTTTTTCATTATCCCAATTATACCTTTA
AATAACTTTTTATATAATAAAGATTGTTGTTTGTGTCATTAGAAAAAACAAATTTAGATA
ATTTTTTTTTGTTATATTTTTTGTGTCATTGAACTAGATAATCATGATTAGTTTGGTTTG
TTATAGCTTGAAACAACAAACATTTATATTTTTAGTTTTAATCTAAATCAAAATTTCATA
AAAAATAAAAACATCGTTCAAAATTTTCGAACGTTAGCGCAATAAAAAGAGCGGAAAGAT
GAGCACGTGAATTCCCGTCTTTTCGCTAGTAAATATAATATTAAGAAAAACTTTTTGTGT
AACCACTTTTTTAGTATACTGAAAGGACAAACACAACCTATTGATGTACTTAATGGATAT
TTTTCAATATTTATAATGTTAAGAAGATGATGAGCTTTAT
>20_Medtr3g030420
ATTGGCTACAAATTTGCTACTAATTAGCTACTAAATGAGTAACAATTTTTAAGTGTTGCA
AGTTTTACTCTCAAATGTTACTGAAATCATTTAGTAACACGGGCTTTACATAACATTTGT
CAAATGTTACAAGATCTAAATAGTAACAATTGTTTTTGATTTAATTACAATTTTGAAGTG
TTACTAAATCTTATATATGTAGTAGTGTGACCAAAGACATTGGCCAAATTATGTTGATCA
AATATCTTAAAATCGAGTAATTATCCAAACAATTTAAAGAATTGAATCAGACATGACCTA
GCCGCCTTTTAATATATCTCTTAAGTAAATAAAATCAATCTAGTTGTTAAGTACAACGTC
```

```
CTTTGTGTTTACCCTGATTTTTGGTAAACAACCGTTGTTTTAAAATTAATTGTTTGCAAG
ATCAGCCAATAGTAGTAGTTAAGTTTTTTTGACAAAGGTAGTAGTTAAGTTCAAATCATG
TAATCCCTCTGAAGATGAATTTTATAAAGTTTACTCCTTGTCGACACTATAGTACATTTG
TTAAAAAATTGCCATCGATGCGGTGTATTGAATACTCACACTGGTTATTGGTCTTCACAT
TGAAGGATATGGTGTTTTCAGACCCAAGAATGAAAGCAACAACAAATTGTGTCTTTGTTT
CACCAATTAGAAAAAATAGAGAGAATTATAAACAGAGAGGATCCTAACTCCAATTTAAAC
CTTAGCGTGACTATTATTTATTGTGAGTTAATTGCTTCTCATCCAAATATCTTAATTAAA
CACAAGACATAAAATGTACCAATTGAAGGTGACGCTATGAGTTTGTTCCTTTAAATAAAC
GAGATAAATTGCGACATCTTAATATACCTAAGAGACACTTATAATAATGGTGGTTATATG
AATTGTCTCTTGCAATTCTAGAAGGGCTAAAAGAAATTTAATTAAAATTAAAAAAGCATG
AGGTGTTTTACATATAAAATGATCAAATATTTGTTTGTTT
>21_Medtr3g031320
AATTTTAATAGGGCTTTTTAGCCCGGCCCTTAAAAGCCCGGTACCCGTTAGGGCTAGCCC
GAATGGGCCGCGGGTAGCCCGTTAGCCCGCATAACAAAAAAAAATTCTATAAATTATATA
TATTTTTCAAATAATTATTTACTTAGTTGATTATCAAAGTAAAAAACAAAAGTGAAAATT
CAATGAATTAACACAAATATAAATGTAATATAAAATTATATTTATTCATTTCGTTATTTA
TAGTTTTAAAGATCGAATATATAAATATCTATAACCATAACATATCTAATTTATTTAAAA
TTATTTTCCTCGTCGTTTTAGTTTACGACGTTTGTACGAAAATGATTACAATTTACTTTT
GTGCTAGACATTATTTAAACATATTAAAAATATTATTTATAAAAAAATTATAGCAGTGTT
TTATAGTATTTTTTTAAAATAAAAAATATATATAATTTTTAATATGGGCCGGCCCATTTA
GAACCGGGCCGCCCATTTTGACAGCTCTGCTTGGGATGCCATTTCTTCCATGCAGTGCCC
TTTTAATCTAATTTTTGCAATAAAAAAATCTCCATGAGTTTTCCTTTCCTATCACGAAGG
GTTATAAAATTATGAAAATTTATTAATTTTTTTTTTTAATTCTTTTGATTGGTAAGTTTT
TTAAGTGGCCTATTTACTAAGTGTGCTAAAGGGACAAATTAGAAAATACAATTATTTGTC
CACGTGACCATAACTTAGTTGACATGAACAGTACATTGTTATATGCAGGGGTCGGGGTTC
GAATCCCGAACGCCCTACGTACTTATTTATTTTAAAAAGTGAATTCTAACTACTAGACTA
CTTGACAAAAAAACAATAATTTACTTAAAAGATATTTACTGACCGTGGTAGTTTTAAAAA
TAGTTTTTCACAACTCCAAAAGCACTTTAAGCTATTTAATTGATTTTTTTTTTTTTTATAA
TTAAAGCTATTTAAATGGCGTAAGGAGAGTGTTAGTTTCA
>22_Medtr3g079850
CCCGAAAGAGACCTATATAAATATTGACAATGAAGCTCTCCACTTGCATAATTTTTTAAA
TTACTCCTAATATTCCTATCATAAAAATAAATAAAAAATCCTCATATTCTTCTTAACTGA
TGAGAGAATTTTTATTAATAAATCTATGATAAAATATATCTCATTTTAATTGTGTATCTT
GTTCTTGTCCCTCATGACAATTCTTTTTAATTTATTTTCTTTCACATTTACGCCAATCAT
CACATTTAAAAGTTTTATGTATATGAATCTTTTTAAGATGTTTGAGAAGTTTTTGCCTCA
AACAATAAAAAATGTTTGAGAAGCCATGTACGCTATCAGTTTTTATTACTATTTATTTGG
CACAATCTCAAATCAAAAATATTTGAACATACAAATAATAAAAATTAATAGAAAAATAAA
TTAATTGATGTGATTATATCGTTTCTTTTTTATTTTATTTTTTTGTATGTGGACGCTCAA
ACATTTTTTATCTGGGGTTGTGCCCAATTAAAAATTGTCGTTAGTGTTTAATACTTTTTA
TTTACAAATATTAGTATTTACCTTGTTCAAAAACTTGTCTCTGTAATTAACATATATTTT
TTCATATTCGCCAATAGAGATGGAACCTACAACTAAATGATTAAGAGACCTAAGGTTATG
TCACTTGTGTTACACCATATTGGCTAATTAATTATGTCTTAGTTGCACCGATCAATGTTC
CAAAGAAAATCAGAATTATAATTAATTCATATTTTAGTTGAAGTACATGTCACAATCACT
GCAATGATACCAACCTTATAAAAAATAAAAATAAAAATTGGACAGATTCTGATTAATCAC
AAACATGAATGGCCAGTGGTAACACATTAGCAGAAAATATTTAAAAGTAAGAATAAACTG
GCCATTAGTGACACCAAATGGCTAATATAACTTTGCTTAATTTCAAGAGTTAATCAAAAG
CTCTAATATTAACCTTCTAATTACCAAGACAATTTTTTTT
>23_Medtr3g089005
TGAACCAAAAGTTACAAGTGTTCTAAAAAAAAAAAAGAAAAAAAAAGTTATGTTTTGTTT
TTAAGGAAAAAGAAAACATTGAGGATATTTAAAACTCAAATCCATGATAAATTAGGCTTA
AGTGAGGCCCAATCATGTTGAAAAAAATAATTAAAGCAAAAATTTAATTTAGAATAGTTG
TTGGAGTTGAATCCATGACATCCATAATGCAAGCGTGCTACTCCACCACCAGGTCAGCTT
ATCATTTCTATTATATTAGCGCGTTTTTAATATATATTGGCTTTTTACATGTATTATACA
TTAAAAAAAATATTTTGGGCCTCAAAATTTAGGAGGCCCAACTCAGTAGAGCTGTTTGCA
CCCCCAAGGACCGACCCTATTAGTTACGCTATGTACCCTAGAGAATTGCTCCCAGAGGAT
GCACCTCAGAAGAACCCTGGAGGATCATCACAAAGGAACCCTAGAGGATCGCCCTATAGG
ATGTGTTGGTTCATTAGTCTAGTCTACAGTGTTATGATCTATATACTGTAGATTATCACA
AGTTTGTTAGAGTTTAGACAATTACTAACCAGCTGACACTGGTTAGAAGTTAGTTTGATA
AGATTGAGTTAATTAGAGATTAACTTAATCATAGAGTTAACTTGGTAGTCAAGTAACCAT
ACCCAAGTCTATATAAAGGGGTTATTGTACTGTGTATTTCATCAAGTGTTTTCACAACTT
CTCCATTGGATACTTTACATAGTTTTCATTCACTCATTCTTCTTCTAAATTGTTATTCAT
CTCACACTTAGTGTTTAGCTTTTCAACCAACAGTAACAACATCAAACTAATTTAATTAAG
AAATATTTAAGTAATAATCGACCTTAAATATTGACGAAAAGCCACCTACCTGTCCCTGCT
TGTATTATGTTCTCAATATTCCATCTACATTATATAATAATTAAAAACACTTATGGCCCC
ATAGCTCAACAGATTGATCAACCAAATTCTATTATAAAAA
>24_Medtr3g089125
GCTCAACACAGCCAGTGGCCGGTATCCATTAATCATAAATCATGTATGTAACTCAAAATC
GTATATTATACTATAATTTACCACAATATTTGTGTAGCATTGTAGAACTGTTGAAGAGAA
AGAAAATGTGAATGAGGTATTCTCCCCTCAAATTTTTAAACAAATGCATAAACAAATATA
GAATTTGTCTTTACCTAGCTTTACATTATATATCTGCTGCAAATCCCACAAATACACAAT
GTAATTGAATCCTCGGAGTTATCCAGCAAAGAATTCGTGGAATTCAATGATAAGCCACGG
AGAAACCTAGGAAAAAAAAAAGTCCAATTGATATATTTGAAGGACATTCAAATGTAAATG
TAAAATTGTAAATGTAAATAACCCCAAAGGTAAACATTTATGAACCATAGACCCAAACAT
AGATACCGGATACGACACGGACAAACACAACAACATCGATAATAATTTAAGAAACTGACA
TGATTCAATTTGATAATATTATCACGGTATTGTGTTAATCACATACAATATGTATCCGAT
ACCGGAACACGTCTAATCTGATATGTGTTTCATCGGTAAATATAGAAATATATACATTAA
TGGAGGCCAATCTTTTTAAAAAAAAATCGGTTAAATTTAAGTTTTTACTTCACATTGGCC
CTTCAGCCTATTTTTGCAATAATTTTAATATAAAAACAGACACTTGGTTTAATGAAAATA
TCAATAGAACTGTTACCTTAATGGTAAAATCTGCAGAGTAGCACACATGTTATGTTAATC
AACATTTTTTTCTGTAGCACATCCCGTGAGTTACGTACGTGTATGTATGACAATTGACAA
TAGGATGAGCATCTGAATGAATCCATTTGGGTTGGAAAACTAAGAAAGGAAACATCTCTC
```

```
AGCATCAAATTCAACGGATAAGATCATATTCATATAACAGTATCATTATTATTATGATCA
TATGTGTTTAACCTTACACTTGTTATTAATGATTCCTTTT
>25_Medtr3g436500
TTTTGATGAAGCCTATGTGCCAAGACTTATTTATGATGTTGAACTAAATCCGCTACAATG
TTTTTGAATTGAATTATGTTTGAAGGATAATTAGTTAATTATCTCATTTATGATTTTTAG
AAGTGTAGCATTCCGTTTATATGTTGATTACTCTGATAATATGTTTGAAATTTTTAAGTT
GGGAAAACGGGGTGTTACAACAGAGACGTATTTTTCAGTGCAAATTTGTTCCAAATTTTA
AGCGATACTTTTACTATAAATATAGAACTTGTATCAGAGTAAACTTTGAGATAGAGGGAG
CTGAAACAAGAGTTTAGAAAGACAACAACAAAAACTAGGGTTTGTATAGAGTTTGTCTCT
TTGGCAACAAACACTAGGGTTTGAGGGTTGATTCACCTTCGGAAACACTATTAGGATTGA
GTCTCTTGGTTACGGGAAGAGATTGGGGCTTTGGGTAGAATTAGGCGGGAGACTTATTCT
TGTAACTCTTTTGTAACGTTACTCATCATATAGTGAAACAGAGGGCTGCTCTCTCCCCCA
AACTAGATCAATTTGGACCGAACTGGGTCAACAAATTCTTTTGTGTTCTCTCTTCTTTCT
CTCTCGTTGTTTTCTACTTTTGGCTTGTGTTTATAATTGTTCCATACACTTTGTTTTGGT
TGCTTGACTATTCCCTGCTCCACACATCAAATTTGTTATTGGTGTAATTTTTCATCGAGT
TCACAACAGCTTGCACATGAGAATGCATTTTCTGCCATATTGAAGGTTAGGCTTTGATCA
AATATGTTTATTAAATAATGAAATCTTATGATTGATATTTTTATTGTATGACCTTCTTTC
TATGTGTGCTGAAGGGACAACTTATTGAAGTACTTAAGAGGCATTATATAGATAATATAT
TTTAAGGAAGGCCTTTTTTATAACTTCGTAGTTTCTAAAAACTCTCATAAGCGTTATAAA
GAAAATATAATCTATTTCATTCATGCAACTAAAGTAAACT
>26_Medtr4g059900
CTCAACCTAGAGGACTTAATAATCACAATCCAAAGATTTACAAAGAGTGATTTTAAGAGA
CGATACTATATACGTGTATCTCATTTGGTTTTCTTTGACTGTTTATAAGACAAGGTATAC
AATGTAAACATTGTATAGATAGAATGCAGTATATTGTATGTATTCACAGATTCGTTACCT
TCTCTAGTCAAAATTGGTCCTTCATATAGAGGTGAAGCGAGAATTAATGCTGAAGAAATC
TTTGTGTTTGAATGTCGTGCTGAATATTGAGATAATCTTGATTAAACGTAGTCGATAAAT
CAACCATTCATTTTCATTAATGATGCTCTTAATGAATTTTACTCAATTGCAAAAGACATG
ACAAAGTTTTCTTTCTTAGCATAGAGTCCTTGCATGCATCAACGATGTGTACTTAAGTCA
GGCATTTAGTACAGAGGATCTTAGTCCATAGGCTTGAACCACAAAGCATTGACTTCCACA
AAGCTTGGCCTTGATCTTGGTACAAAGTCTTGTCTTTATCCATGACACAAAGATTCACTT
CAAGGTCCTGAGCGGTGACTTGATCCATAGGCAGAACTTCATTTACTTAAAATTTCTTGA
ACACTTAAACATATATTAGATTGTCTTTATTGTTCTCGTATGTTTGTAAAGCGCTTCTAT
TCTTTAACCAATATTGGTTTAACAATATGGCCTAAAAATCATTCAATTGTACTTTTTAAA
TTATAAATAAATATAATGTCACATCTTCTATGTGTTATTTATACAACTAGTTGTTTTCTT
AATTTATTATGGTTTATCTCTTCATGAATTTTTCTTCTATCTATTACGAAAGGCCAAAAT
GTGTTGTGTGTGTTATGCTCTTAATTTAGCGAGAAAAAAATAATAATTAACTTTAATAAA
TGGCTTCTTGAAACTCCTAAAGGATAAAAACTTCATAATTGAAGTTAACAAAGCATAAAT
TTCTTACTTATAAAGTGATCAAACAAATTCTATATATTAC
>27_NiR_Medtr4g086020
TTAGAACCCATTTGGGTCTGGTGGTATGCTTGTGATCTAGAAGTGTGCTTCTTCCTAAGG
TCTCGAGCTCAATTCTTTCCGATGTCAATTTGAATGGGTTAATTTAGCTTCAACTTTTTT
TCTTTCTTGTCTTCAGTGATGCAAGTTTCCCTTTTCAATACTCATTCACCATTTGATTCA
GATGTGGTGACTAGAGATAAAGCATAATTTGACTCTTGGTAACCCATAAAGATTATTTTA
ACCGTGTGCTGGACCAATCCACTTGGGGGCATGACCTGTGTCTATGTGTGGTTTGGTTTC
CATTCTGATTTATGCGGCGACTTGTAATTTTAAAAAATCCAAGAGAGGCAAACATTGAAC
AATCCCAATATTTTAATAACTTATGCATGATTTTTTTAACTAACGATGATATGTTTGTGA
CTGAGATTGAGTCATACATCTCACTAAGAAATGGTTCAAGTACCAAACAATCATAGCCAA
GTTGTAAACATGACGTTCAGGATTGGTCACTTTGATGGTTCAATTTCATCTTGGCTTCTT
ATTCCTTTTTTAATTCTAATTCTTTTTGAGTAAACTATTTCATGCCTTCTTTTTCTGATT
TTCTTTGGTATTACATGTCATTTGTTTTGCCTCATTAACTAATAATGATAAATGTTATTT
TGACTCACAAAATTCATATCAAATTTCAACGTCGTTCATATTGATTGGATGATTAACAAA
TATAACAAACTATCCAACTAACTAACTAACCACCAACTGAATACATTAACTATAACTGTA
AATGTTGTTAACCATATTTTTCAGATGTATATATATCATCCGTTGAATGTAATTAATCAT
ATATTTAAACTAAATTATCCTAAAAACTCGGTTCGAGACTTGGGCAGAAAATGGACGAAA
ATGTAATGAAACTTTAATTTTGACTTTGACACCACACTACATGTGCCTTTTACATATAGT
TTATATGACAAGTAATGACAATCATGCTCTATTATAAGGT
>28_Medtr4g088350
AATTCAGTCCAAAGACCGACTAAGGTGGTGACCAATCTTATTGTCCACTAGTAGAAATCC
GCTAAACCAGAGCTCCAATCTCATGGTCAACTAGCGGGAGGCCGACTAATTCAAGGGATT
GACTTCATCATCTACCAACAGGAGACTCATTTAAAGTCAGAACAAAATTTTGTAAATATC
AGTCACAAAAATCTCTCCCATACCATCCTACTCTCTCACCATCTCTCCCATCCACTCCTT
CAGAAAATCTTCATTAACGACAATCAAGCCATCAACCACACCTAACTCACGTGTTTGTGA
AGCACGTAATTATACCCAACACACAAGATTAAAAAAAAAAAATAAACAATCTTTTTTATC
TTTGAATGTCTGAGACATTGGATATATTCTTAAACGATGAATGTCCATAACACTCACTTT
CAACACTCTTTTTAACATTCTTATTTTTATTGGATGAGATCAATGTATATTTCACCACTT
TGTTGGGTTCAATTTTCAAAATAGAGTCACAATAATTTAAACTAATAAAAGAGTGAGTAT
TAAAGAGTGTTCAAACGAGAGTGTTATTAGCATTGTTCTTCTTAAATATATCAAAATTGC
AAATTTGATATCCGTACCGATTAATGTGGTGATCAATTTTATTGAGGGCTGATTAATTCA
GTGGATCGATTCTACTATTTCCAACGAAAGTTCTATTTAAAGTCAGAACAAGATTACATA
TAAACTAATCACAAGAATTGATGTTGAGAAGATTCAAACTAGAGATATCTACACCACCGG
ACCAATCTATGAGGACCAATCCAAGTCTATATTTCAAACTAGAGATATCTACACCACCGG
ACCAATCTATGAGGACCAATCCAAGTCTATATTATTAGTCATTTTAATCTTGAAATCTAT
TTATAAATAGTCAATTTATTTATACAAATTATTTTTGGTTGTTTAATTTAGTATATTTTT
GCAAAATAGTGTTTTTAAAAAATTTGATACATTTAAAAAC
>29_Medtr4g094335
ACAATTTCAGCATTTTTGCCAGTTGAGCTATGACTTATGAGACTGCTCCCCTAACATTCT
ATCGTCCGAAAAAAATATCTTTTATTCTTAAATTGCATATTTGCCGATTTCAACTAATAT
AATAATTATTTTTATATATAAAATATATAACAATTAATTACTAATATTGACCGTTTAAAA
AAAATGTGAAATGGAGTGGTGTGCTTTTTTTCAATTAGTTGGAAACCCGCTATTTTCTTT
TTGAGGAAAAACGAACATATAGTAAGCTGGACTTTCTGCCCTCCTTTGGGATATACTCAC
ACTTTCGTCTCTTTTGACTTGAGACTAATCCTCGGTGCGACCGCATAAACCAACTGAAAT
```

```
TGATCTTTTGAACCTTTACGAACCCTTCAATTAGGACATTTCGTTGAGATCTAAGTCCAA
CTCCAATGATTTGCAGAATATAATATTCTAATCTAAAATACATGATGCCTTGTGTATAAA
TAAATCTAATTAATTAAGCTGCACAAATTAAACTTTTTTTGGCCCATTCACAATTCCATA
CTCCACAACTACATAGTTTACCCCCTCGAAAAATAAATATTACATAGTTTACCATCTTTT
TTGCCCATTCACGTTGGCCTATGTATACTGTATAAATATTTTAACTATCATGAAAGGTCA
TTGACGAAATTGTTTATTTTCTTTTATTTACAGATTGTACTACCTGGGATTTAATTTGTC
CTTGAAGTATACTATATGAAGATTAATCTATAATTTTTTTAAAATACTCTTACATTTTATT
ATTGAATACATTTTTATTAAGGTCGCATTAAATATACTTAATAAAAAAAATATCCCATGAT
ACAAAAAAAATGTATCTAGTGATTGCCATTTTATACGTGACGCCAGGTTATTAATTTTTTT
CCTTTGCCCCTTCGAAATATCCAAACCATATATTTTCATTTTTACCAAATGCATTATTTG
TATAAATAAACAGACCAATCACACATTCATGATTTATCCC
>30_Medtr4g099400
CTATATCCATTTGATTTCTACTTAAAAATAATGTATTCATTTTTATATATTTTTTTGTCA
ACAGCTAAGTAGCAAGAATCATAAAATACACATAAGTGGAGAATTCCGAATTTAAACTCG
AGTCTCAACACATCAAATATCTTTACAACGGTCAATTGAACTGTAGTTACATGGAACTAT
TCATTTTTTTATCAATACTTTTAAGGGTTGAACACAATGTATGATTACAATATTTTTGATA
CTTTTTGGACACAATTTGGTAGGTATAATTAATTAATTAACGAGTCTGTGGGAGAATTGC
AGCAAAAACATTGCTAGCAGCCCTCATAGCTAGTTATCAAAAGTAAAATATTACAATTCT
TGAATAAATGGGGGGTGCAATCTTGAATAATAATCAACCACGGTTGTTTAGACATTTCTA
TGTCTATCTATCTATGTACAAGAGTATTTCTTTATTCAAATGTTGATGAAGTCTGTCACA
TATTCTATTCACATTCTCAATGTTCACTCTGTTAATTTATTGTTTGCTCCAAGCCTTTGG
TTTTGACCCAGCAAACGAACTTTTGAAAATTAAATATATATTAACCATATATATATATAC
TATTAGACAAAGTAAAATTCAAAATAACTATGGTCTATGCATCATGTAGCAGTAGAATAT
TGGATAAACACAATTGACAAATTTAGCTAGACCAAAGTTGGAGAAAACGGTGGCTTCATT
ATCACTGGTATATGCATAAATCTTGACCAAAGGATTGTACTATAAATACTTCTATCTTGG
TATCGTTGTTTCACATCTCATAAAAGCATTTTCTTCTTAATTGCTTACCTTTCACAGGAA
AAATAAAAATCATCACATTTGATAACTTGTGTTACATTATTATATATTTAAGGCATGGAT
CTTAGTTTTAAGCATCAATTTAGCCTTGTCTGTGTTATTCTGTTGTTTCCAGTACTATGT
AATTGCAAGGAATATTATACCAAATCCAGAGCTTCCTATT
>31_Medtr4g107630
AACAGATTGGAAGACGTCGTAGACAGAAAGTGCAGTGACGTAAATGCTGAAACTCTTGAA
GTAATTTTAGAGCTAGCGGCAAGATGTACCGATTCAAATGCAGATGATCGACCTTCGATG
AATCAAGTATTGCAGTTGCTGGAGCAAGAGGTCATGTCGCCTTGTCCAAGTGAATTTTAC
GAGTCCCATTCAGATCACTGATAATGGCTAATGTGTTCCTTCCCCCTGTTTCGTGGCTTG
GTTTTGTAGAAAAGTTTTGTACCAAATGTGTAACAACCATTTGTAAATAGATTTTCTCCT
TGTCCAAAACAAAACCAAAAAAAGGACAACGTGGAGATGGCAATGGTTTTTTGTATTTGG
AATGGTCTCTCCATATTCTTGGCTTCTTGCTATGGAAATGTAATCATGAAGATGATTAGA
TGCTATGTTCAAGTAGGGTCTTCTTTTCTTCTTGGATTTGAGGTTGATTTTGTTTTTCAA
GAGATTCTTCTTGTGATAGTATTCTTTAAAGAATATGTTGGGTTTGAAAAAAAACTTTAT
GGAACAAAGACTTGAAGTCAAGAGGAGGAACTTTTCCTACTCCCAATATATTTTTGAATT
TATACTTTCATCCAAAAATTTATTTGCAATGCTTTTAAAAAGTTTTTTTTTTTTTTTTCA
TTAAAATAAACCGTTTGGAAAATTATTTTTAAAAAAGTGTGGTTTTAGAAATATGAGGTG
AAAAAAAGTTGTTCCCCAAGTGATGATTATGAGAGTAGAGAGAGAGATTGGATCTAGTCC
ACAAAGAGAGAAGAGGTGAAGGTACTCCTAAAAGTATTAGTCGTGCGTATAAGGTAAACC
ATTATTTTGTTTTCTTATAATTTATTTTTTTTGGTCTGTGTAAAAAATAGTGACTATAATT
CTGTATTGGAGTGTGTTTTTACCGGTCAATTTAGTCCTTAAAAAAACTTTCAATAAAAGA
AATATTATCGATTAAGTTAAGGAAATAGTAGAAAAAGTGG
>32_Medtr4g128540
ATGCTTTTATTGCTTGGCCGCCTTTAAATTGATATACGGTTTATATTGTGAAACGGAAGT
TGAATTTATAAACGCTTTGAAATAAGAAATTCTTGCCTTTGTAATCTAGGAATAGATGCA
AACCCTTGAAACCAATTACGAATTCTTGTTCAATCATGCTTAATTCTAATCTTAAATTCA
CTAAGGAATTAGGGATTATTTTAGGATTAATGATTTTGTCACTAAAGAATTAGGGCAAGA
ATAAATTAGAGAATTCGGTAATGATACGATAAGTCTATTCAATGGTCATATTCAAGTTGT
TAACTGAAGATAGTCATCTGGTGAAACTCATCCCTGACAATCTTATTATTATAGTTCAAT
CCAATTTTACTTTACTTAATTTTCATTATTGTAATCACACAAATTCAGTACCTTTTTGTT
CAATTAAGTTATAAATACTATTCGACAGTATAACGCGATCCTCGAGTTCGACACTCGGTC
TTACCGTTTTAATTTTACTACTTGCACGATTTTGTACACTTGCGAAAATCGCTATCACCT
ATTAAGCCTAAATTCTTAAAATTGTTTGAAAGTTTTCACTTTTTTTTTCATTAAGTTTTC
AATTGAGTTTATTGGAAAAGAATTGAGTTTCTTCCTAAAAAATTCAGAAACAAGAATTAG
AAGAAAATATGTTAATGATGGTAGTTTTTTTATTTGTTGATAAATGTTTTGATTTATATA
TATTTGTTGCGTTTTTAAAACAAGTGACATACCTTACCAGTTATCAGCAACCAGTGAAAA
TATGAGGAAGTTGTTTACACACAAGTGTGTGTGTGTGTGGGTGGTTGAATTGAATGTGTT
AAGCTTGATAGCAGCACGTTATGAGTTATATTTTTTTTTTTAGCAATATAAGGTGACATTG
GCATTGCTTGATTAAAGAGATGTAGAGGAAAAATATTTATTACAAATTATTTTTGTTATA
GGAAATTGATCCATTAGTTTGTTTTTGATAAAGTTATAAG
>33_Medtr4g130780
GAAACGGAAACACAGTTAACCTCTCAGAAAATGAACAGGTTTACATATAATATCCTCTAA
TTGTATTTTTGGAAAGGGATGGAATTATTACCACAAAAAAAGGTGTAAAAGGTAACTTAC
TAGCAGCGTCGAGCCATATTACTATTTTTCTACAACCATAAAGAATCCATTCCCGTTAAT
TTTTTAAAATTGATTATTATCTATTTCATCCCAAAGGATTATGAAAAGGTGCAAGGAATC
TTTAGGTAATGGCTCGCCAGTTCACCAAACCACGTTGTTGCTTCCAATTGCCAAGTAGCT
TACGAAGATTCTCAAGCTACCATTGTTTACTATTATTATTACCACTTTTTTAACAATAGT
AACAAAAGTCTTGTTAACGAGTGCTCTAAGGCTATTCTTTAAAGATTTTTTTAATAAAGA
AATTATTTTTTAAAAAAGTTAAAAACTCCAATTTTCAATGCATTGATTTCATAAACTTTG
ATAAAAACTTACTATTTAAAGTTACTAAACAATGCCCTAAAGACATTGGTTTAAATTTCC
CTAATAATAATAAGGGTTTTTCTATCATGTGCAAATTTGCACATGTTAAGAAGTTTAAAG
TGGTAACTTTAACTTGGAACTTGTGCATTTCATATTAAATATGTAGCTTTTTAATAAATA
AGTTTTGTTTTTCAAAAAAAAATTACTACTTTTAGATATCTTAACATGTGCAATGTTGCA
CATGTTAGACAAACCCTAATAATAATAATAATAAGGGATATGCTAATTTGTGCTTAAGGA
CACTCTTTAAGCACACTTTAGATAATATTTTTTCATGAAAATTTGTGTATTCAATGAAAT
AAAATTATTAGACTTTTCAAATATATAATTTTTAATTGTGAATCCTTAAATAGTGTCATC
```

```
GGGACACTCTTTAGCAAGACCCTAATAATAATAATAATAATAATTTGGATATTAATGTCA
CCTAATTAATAATAATAATTTCACATTCCCTAATTGTTTT
>34_Medtr5g027460
TTATATACAAATTTCCATGGGCCAAAAGAAAAATAAAGTAAGATATAATGATGGTAACTA
TGTTAATAGATAGACTATTTTATTATTCCGTAAATTTAGCTAAGTTGATAAGGATATCAT
ATTTTTAAAATGTGTGAGCTAATTGATAGAAAAATTATATTATTTGTAAACGACAGTATA
CATGATAGTCTTTTTTCCTAAAAAAAAAAAATGATAGTCTTTTTTTTTATCTTCGACATC
TTATACATGTTTTAAAATTGCACACAAAAATTAAAAGACATTCGAGAATCAAAACGAGTG
TGTACTCATTTTTATAACTAAACTCTTTTTTATATTCAAAAATGAGAACACATTCATTTT
GAGCATCTAATATGTATTTTTGATTATAAGAATGTGTACAATGTTGCTACTATAAATATT
AAATATATTGAAATTATAAAACCCCCTCTAAATGTAATCAAAAAATAAAAATAAAAACCC
CTCTAAATGTCTCTTCGTTTAATTTAATAATTTGATTTCTAAACGTGTCTTCTATTTATT
AGTTTAATTTTCAAATATACTTTTCATTAGTGACTTTAATGACTTTAATTTGTGAAATCA
CTTATTATAAACACAACTAAAGATGTTTTTGTAGTTTCAATACAATCAAGAGGAAGAAAA
TGACAATTCACTTGTACAAAAATTAATTTCTTTTATTAAATATAAGGTTAAATAAGTACT
ATTAATCTCTACACTTTATCTCTATCCAAACAAAAGACAAAATTAAAATCACATGCAACT
TAAAATCACATTTTAAGTAACTAAACACCACAACCATATTATCATACATTTAAATTCTTA
AAGTTTAAAACATTATTTTTTATATATTCCTTAAAACTGAATAACTTTTTTTTTTGTTGT
AAAAACTAAATTAAAAAATTACTATAATTTAATAAAAACAAAATATATTTAACCCCTTAA
ATATATGTACAATATGCCCTTAACAAAAGCCCCATTGACA
>35_Lb7_Medtr5g041610
GATGCGCTTGTTCAAATAACATCCCACATTGTTAAATGTGTAACCAATTGTGCCTTGGTT
GCTCTTGCCGAGTTAGGCGAGATGTGGGTAGGAGCTATCCAACTGAGGGTGTCAGTTGAG
ACATAGGGTTAAACGTGTGATGTTCGGGTCATGGGCCTAGGGGCCTTAGATATGGGCCTT
AGAGTCCATGTATGTTAGGGGATGAGTACCAACCTAAAAATTACAATGATACACATGCAT
TCAAAGCACGAGACCTATGAATGTTTGAGATTAATGTTGTGGTGCCTACGATATGAGCCA
TGTATTGTCCAAACAAACATAATTGTACAGAGACTAAAAACAATCCTTTTGTAAAACGGT
TGTTAAACATATGCTAAAATTAAAGAAGGCAAAAAAAACATATGCTAAAATTGGATGTAA
GGAGTATATTTAAAAAAATTACAATTTTATAGAGACGAAAAAATATATTTAGTCCTTTAT
TTATTTATTTATTTTTGGTTGTAATTTTTATGTCTTTCTTCTAGTGTTGTATTCTACATT
TTTAATATACCTATGAAATTTAACAAAACTGCACAACCAGTTATGATATTGTTTTATTCA
ACATTATTGTTGTAAGAATTTCTACAATGGCTCGGGCAACATGAAATTCTAATAACATGG
TAATAATAGGATTGTGTGAAACAGTAAAAATGTTACTGTTTCTTAACATTTTCTATTCGA
CTAAATTAGTGTTTATATTGTTCAAATGAAAGGCTACTTATATCAATGTTTAGATTTTTT
ATTTTTAAAAATAAATAGTAAGATTAATTAAAAAAAATTGAGGAAGATTAATTAAAAATT
GTTTTGTTAAGTTCTTGAAAAGATTCTTGTCTCTTAATAATGTCAATAGCCACCTCCACA
AGCCAACAAATTCTTTGAAAAATAAATTAGAATAAATCATCTACATCTCTTACACCTTCA
AGTTTTCTATATAAACAAGTCATGAATGTAAATTTATTAC
>36_Medtr5g067410
CCAAAACTTTTTATTTGCAAGAGGAGAGACCATCAAGTAAAATAACAATTTGTCAAAAAA
AAAAAAAAAAAAAGTAAAATAACAATTAAATTAGGGGTGTAATAAGGAACTCCAAATTCA
CCAATAATTTGCATAAAAGATATACATATTTGTGACTTTCAAATGTAACAAACTCAACAA
ACCATATTTATATTTTTCAATTATAGCAATAATATAACAAATATAATAATTTTTATATTT
TTTTATAGAGTTTTTATATCTTTTTTAATAACACGAATAGGTAATAATATAATATTATTA
TATAAAAATTTATTTAAATGTATAATTGGAATAACAAAAAAGTCTGCCTAAAATCTCTCC
GAAACCCCCTATTTTCTTTATACATATTATATGTATGTATGTGTGCGTGTGTGTGTGTGT
GTGTGTATATATATATATATATATATATATATAATTTATATCATAGATTCAGCAAAAGTT
TTTAAGTTAGATAGTCATCATATCATAAAGATTTGATTGCCTAACACGACGCTCTATTCT
TTTATTCATATTTGAGACGGACTATACATAATATGACTGATATTACAAACACCAACTTTA
ATATGAAAAGATAATATTAAAACAAAATAATAATTAAAATAAAATATGGTATATTCCTTT
TTTTTAAGAAGAAATATGGTATATTCCTAAGAATAAAAAATCAAACTTGAAAATAAAATA
TAAAAAAATAAAGAAAAACTTAGTACTAAAAACAAAGTAAAATATGAGACCAAGTATATA
TATACAATTTTTTTATTAAACATGAGATTCATATTGGAACAGAGAGGGATGATGATAGAA
TTGAAGGAGTGTGGATAAAACGTGTACTCACCAACCAATCAGAACACAAAGAAATCACAC
ATTCTCAAATAATTCCAAAAACTTTCCATACGCCACATCTTTTGCCACATCATTTCATTA
TTCATTTCCACACAAAAACTCCAAAAACCAAACCAACCAT
>37_Medtr5g069890
AACACCACTCATTAATAACATGATCAATCCACTATTCACCCAACCATCTCTACAAACCTC
TCACCTCCAAAAAATAAATAAATAACTCAACCCAAATCCCTTAAACATTGTCGCCTCCTT
CGAACAAGACTCACCTCTGCCACAGTTCTCTTCCCCGTCCATTTTGGCTACGAAGAACAA
TTTGTTCTTGACATTGATTATCATTTTCTATAAAAAAAATGAAAATTTTGATTTTAATTT
AGGAAACGTTCAAACATCTTAGTCACTTGCTAGTTGTTGTAGTTATGAAAACATGCTTGT
CGTGATTGATAAATATGCTCAAATTTAACTTTTGAATGGTTCAATACATGTAGATGACCC
CATTATCCTCACAATCATCTAGCCAAACAATATTACTTCATGAAAAAAAAAAACTAGCCA
AAACATAATTTTTTATAATTTTTTGTTGAGCAATACATAATATCTTTTTTTCCTTGTATA
ATATATTGAGAATTATGACTTTGATATTAACAAACATCTTAGTCACTTGTTAGTTGTCAT
AGCTTTAAAAAAAAAAAATTACATTGGCTTAAAAAATATGCCCAAATGTTAACCCTTAAA
TGACTCAACAGGCATTGACGAATCATTACCCTCGCAATCAGCTAGTCAAATTTTATATTT
TATGTTGTACAAATATATTATTTTGTATTTGGTTTGTCTAGCATGTGCAATATTGCACAT
GCTAAATAAACTAAAATTAACAATTTTTTTATTGAAAAATAACAATCAATTATTAAAAAT
TTGTATATTTAATGTAAATGGCATGAGATTCAAAACAAATTTATTATTTTAAATTTATTA
ACACGTGTAAATTTGCATATGATAGAAAAAAATTTCTTTATTTTATTATTTCATGTGGTA
ATGTTTTTTTATTCTTCTTACATTTATCATGTCGAATGATAATTATAAATACAGACTGAT
TGTATACATATTTCACAAGAAACTCTTACTTATTGGAAGT
>38_Lb6_Medtr5g080400
CCAACACCCTCCCCAAGGCCCTCACGCACACCACACATCTCCCGCCACCAAACCGAACAA
CGGCTACCCTCATCCTTCAACCTCCCCCCTCCTCATTGTAACGAGCTTTTAAAACTCGAT
ATCACAACCCCTCCTTCTCCACCAACATCCTCCAACACCACTTTTCCATCAAAGATAAAT
TAAAAGCTCCCACCCTCCTCACCCCCAACCCTTCATCCTCCTTTAGAGTACAAATAGACT
TCCAATTTATCTATGCTATTTTCCCAATTGTCAATTTGTTACAAGGGAGGATTATTGGCA
CAATATGTTTTAGCATATGGGTGCCTCAACATAGCTATGAAAATACATACGTTTCTCTTG
```

```
CATTGGATTAAAATTAGACAAACTTTATGTAAGACTAAATTTTATTTTTTTTTACGGATG
TAAGACTAAAATTAAAAAGTCGTATTTTTTTTTACAAGAAAAAGTCTTAATTTTATAGAG
ACTATAAACATATTTGAACATTTATCTTATATATTGTTTTACTGCAATATATTTTTTGTTG
AGTTTGAAATTTTTATGTATTTCTTATGGTGTAGAATTCTTTACATTTTATACAAGTAGG
AAATTTTAAAAAACCTACACAGTCTTTGGTTATATTTTTTCAACTTTACTTTTTTTTTC
ACCTCGTTTTTCAACATTATTGTGATTAGCATTTTTTTTATTAGATTAAACAATGATTAC
GTTGTAATTAAAAAAAACTAAGGATTACATTGATTGCATGGAATAATATTTATATTAATG
TTCAGTTTATTTTATTTTAATTTAATTTTTAAAAATTTATTACGTCTATTTAATTAAAGT
TTGTTTGAATAATTTTTTTAAAAGATTCTTGTCTCTTAATAATTTCAATGGTCACCCCCA
GAAGCCAATAGATTCTTTGAAAAAACAAGCACATGAATTTACTTCATTTCTTTCCACCTC
AAAGTTGTCTATATAAACAAGTATTTAATGTAAAATTATA
>39_Lb1_Medtr5g080440
CTTGTAAAATTACTATTAAAATTTTATAAACTTGAAAATTCGTGATTTTGTAGGGACAAA
AAAATATAATTAGCCCTTTTTTATATTTCAAAGAAATTTTAAAAATTACGCAATATTGTT
TTAACGTTACTTTTTTTTTTTTACAAGTTGTTTTAACGTTACTATTGTTAATGTTTTGAA
TTCGATTAACTAATAATAACGTTTGTTTAAATGGCTTGTAAAAAAAAAACGTTTGTTTAA
ATGGAATAATATATATATATATTAATGTTTAGTTTATTTTAAAAAATTTAAAAGTTGTTA
CTTTATTTAATAATTTTTTTTTACAAAAATAAACGATATTCATTCATTCAAATTGATAGA
GTACATCGATACAATACAAATCCAAATTCGCTAAAAACAAAAAGGATGAATTTGTGAACA
AACTCACAAGTCACAATATCCAAGTTAATAGCATAAAACGACAAAGTACTTATGCCTACA
AATATGACTAATATGACTATATTGATGTCTCTGGAATACTCATGTCTCCAGATCTGTAAC
GTTGACGACGACAAGTCATTGTCATTGATCGGATCTGCACTTGCTCAAAGTTAATCTCAA
CCTGAATCAAATAAACACCGCACTAAGACAGGAAATAAAAAACACACCGCACAAAGACGG
GAAATCAAAACAAACAGAGTCGTTCAAAGACGACGAAATCACAAAATCAAACGAAAGAAA
ACAGAAAATATGTGAAATCACTTATTCAATTAAAATAGAAGAAAAAACAATTCGAATGGG
TGATTTTTGGGTCAAAATTGACCCTAAATCATGAAGAAGAAAAGTTTATTTAATAAAAAA
TATTATTAAGTTTTTGGAAAGATTCTTGTCTCTTAATAATTTTAATGGTCACCGCCACAA
GCCAATAGATTCTTTAAAAAAAAAACAAGCACATTAATATACTACATGTGTTTCACCCTT
CAAGTTCTCTATATAAACAAGTGTCATATGTAAAATTATC
>40_Lb4_Medtr5g081000
ACCTTGGTAACTAGAGGTTTTTCTTAATTAGAAATTGCATGATCAAATAGTTCTGATATG
AGATCAGAGAACGTTCTTATCTCGCACTTAGATATAGATCTAGATTTAGTTACACAAATT
CTTAGTTAGTTATATGTAATTTGCTTATAGTTAGAGGCATTTACCATATTTAGATGTACA
TGGGAGTATGAATATGAGCTTTAAGTTCTCTCATTAATGTAACTTGATAATTTTTCATTC
ATCAACAATAATGAAAATTTGCTTTCCTTTATCTCAAGGTCGAGATCAACACATATTTTT
AGCATAGAAATCTTCATTTTTATATATTTAGAAGTTTTTTTTGACAAAATATACCTAGGA
AATTTAACTTCGTTGGTGATACTTTTTTCACAATTATTTTTGTAAGAATTTATGTGATGG
CTTCGACAACATAAAGTTTCTATTTAAGAATATAATATGGCAATTACTGTATTTTACAGA
ATGGTAAAACGTGTAGTAGCTATTTGTTGATTCCTTTGTCTCATTATAAGTGTATCATTT
GCAATATTTTATTCTTTTAAATTAATTATCACTTTAAAATATCAATCAGCATTAATTATT
TTTTGCACTAATATATCTTTATTTTTTCTATCTTAACCCACTGAACTATAATTAATAAGG
ATATTTTAGAAAATCAAACTTATTTATCATTAAAATTAAATAAAATAATCATATTTATCA
TTAAAATGAGATATTTACAACGAGACAAATAAAACTTATTTTTATTAGATTAATTAATTA
TTACATTGACTAAATGAAATAATATTCATATTGATGATTATATTTGATTTTTTCTAAAAT
ATTTATTAACTTTAATGAAAACTCGTTTGATTAAGTTTTTGAAAAGATTATTGTCTCTAT
AATATTGTCAATAGCCATTTCCACAAGCCAATAGATTCATTAAAAAAATAAATCGCACGA
AATACTACATCTCTTTTCACCTCTAATTTTTCTATATAAA
>41_Lb3_Medtr5g081030
TCGTCCTTGAATTATATAACTTACTATCACATCGACCATTGCATAATTTAGATTAATATC
AAACCAATCCTTTGGAATGACCTCAAATGACTCATCATTGATAGACATAATTATAAATCC
ACAAACAAAAACTCATTGATCTATTAATGTCATTGAACAAAACCCGTGGTAGTCTGTACT
TACGTCCTCCACAATATCTTTTTTCACTCCTTATCCGCGATAATTCCTACCCCCATTTTT
TGATTTAACTTTGCGAGTGTACCAAAGTTTAAATTATGAACTATCTAATTCTTTCGCTTT
CTCACACACCCATTTGGTTTGTTGTAGGCACATAAAATTGATCTTTCACCTAGTCATAGT
GTCCAATATTTCCATACATTTTTATTCAAGAGTGCATATATTCCATCAATTAATAATTGA
TTCTCTTTAATAGTCGGTAAAGAAGCTATTTTATGAGAGTATGATTCTTTTTGGAACATG
GAATACATGCATTCTCACAGGTGGTTCACGAAGATGTAGAAACCCTTATTTTTCACTGCA
TTCAGGCGGCGATGCAAATTCAAAAATAGTAATTGTTTTGTCAATATTACACTATATTAT
TTATTTCAAAGAGAGAAATAGATAGAATGAGATAACAATTTAACATTTGTTATTTATAAG
AAAAAGATAAAAAAAATAATAATAATAAAAAAATAAAAACAATTAAAAGTAATAAAATTTA
TTGGTTTCCTTGATGTAAGTATAAAGCCCCAAAATGAAAACTAAAAAAAAAAATAGAAGGA
GCATTACCTTTATTTGATTTGTTTTTTTGGTGAGAGGTAGGGGGCAAAGCCCCCGCCATA
TTTGATTAAATTTTGTTAGTTTTAAATGTTAAAAAAGATTCTGGTCTCTTCATAATTTCA
ATGGTCATTTCCACAAGCCAATAGATTCTTTGAGAAAACAATTACAATGAACTAACTGCA
TCTCTTTACACCTATAAGTTGTCTATATAAACAAGTATTG
>42_Medtr5g081555
CCCCGGTACCTCCACTTGTGTGTGTGAGTTTATAATGGCTTTGCCATTTCGTCTATCTAC
CAAAAAAAAAAAAAAAAAAAAGAATTGATTAATTCTCTATATCTTAAATTCTTAGTCTAT
TGATATCCATATTGTTGTTCAACTTAGAAACAGCCCTACCATACTAATTATTGCTATCAA
AATTAACTATTTATTTTTTCATATCATCAATATTATAATATAAACCATTTTGATTTTCTC
ATATTTGTCCAATTCATGAGGCAAAATATAAAAGATGACTCATGAATATCTATAGCAATA
TATTTATGTGTGCGTCAAAATAACAACTCATCTTTTATATGTCATATACAAATAATTTTT
ATTTTATTTATAATATTAGACAATTATATAATAATAGAGGATTGTGCCTATTTTATTTTA
GTGGTTTTTTGTAAGATTGTTTTTTTGATGGTAGTATATAGTATTGATATAAAAGCCAGT
TACATGTTTAGTTAAGTGCTATACATATATTCTTATTCCATTGTTGTTACATAACTATAC
ATTTAGTTTCAGTTAAATAATATAATTTTTAACAAGCAACCCACGGATAATAAAATAACA
TACACTACAATTTTAACCACATAATTAATCAATTAATTAGTTACAATTTTCAATAATGTC
TATTGTCATATTATTTTATTTTTTAAAGATATTGTCATATTATTTAAAAAAAAATAACTT
TAGCACATCAATTACCTTTAAACTACAATTTGATTAATATTCTTTAACTTCCATTTAAAT
CCTTAAATATGACTCTTCATTCACTCATTGTATTAATACAAAATTTTAATAATAAGTTAG
TTAATTATAATTTCATGTTCATAGTGATGATCATTGTGTATAAATATGTGGCAAGATTTT
```

```
GAGATCTTTTCATCTTATCCTCCATCCCAACAAAACGAGTGTCCATAGAAAATTAGAAGA
AAAAAATGGAGTACGAAGAAAGGGCTCAATTGTGAAGTTT
>43_Medtr6g060570
TATTGAAAATGGGTTTTGTTCTTAGTAAATTTGATAATACTTTGTTTAGCAAACTAATAA
AAAAAGATATTTTGTTAATTAAAATTTATGTTTATGATATCATCTTTGATTCCACTAACG
AAATTTTGTGTATAAAATTCTTAAATCTAATGTAAGGAGAATTTGAGATGTCCCTTGTAG
GAGAGCTCAAATATTTTCTAGGATTGCAAATAAAACAAGTTAAAGAAGGAATTTTTATTA
GTCATACTAAATAGGGTTATGCTCTTCTCCCATATGTTTTTCCTCATTCTCGTCAAATAG
TCGAGCATGAATTAACTAACACATGTGTATTACAACGTGAGTTAATTCACGTATGTGTAT
TTTTTAACGTTGGACAAGAATAAACAAAAACAGATTAAAGAAGAGCAATTCTTAACTAAA
TATTTAATCGGAGTTACTAAAGAAATTCGAAATGCAAGATTACAAGCTCATTTCAACACC
TGTAGCCTCTCATGCCAAAAAATTAATTTTTAAATACATTGAATGATTATTGTATCTGAC
ATATAATTATAACTAAATACATTAGTTATTTAATGACTCTAAAAATTTGTTTTTTAAGGG
ATGAAGTATTTAAATACCCGATGAAGCCTCTTTGGATGGACAGGGAAACTTGCAGCGTAA
GCATATTCTTCTCTATGTGCTTGGTGATGGGCCCCACACCAGCTTCTAAAACAAGACTCG
CATGATACCCAAACCCATCTCAACCGTACATTCATCTTTCCTTTCCCACCACCAACGGTC
CAGATTCACCACACTTTCCAACAGCCACGTGTACAACCACCTATATATATTCCACTCCCA
CTCAGTACTCACACACTCTCACTCGATCATTTTCCTTTTCATTTTCACCTCATTGTCACT
TTCTCGAGAAAAAAATAACTCAGAAAATCACAGAAACACCGAGAAAATTTTCAAAAAAAC
CGTAAAAAAATAAAAACACGAGATTTTGATTCTTACAACA
>44_Medtr6g090485
CATCAAACTAACTACCTTAAGAAAACCCTTATCTCATAACAGATAATTGTTGAGAATTAA
TCCAACAAAAAGAAATAGAGAGGTTTCAATAAATCATAATGACCGCGAAAATCTGCTATC
AAAACCTTTAATTTTCCATAATCAACACAACCGCTTTTCACAATCAAATCAAAATTGCAA
ACTCCCGCCGTTTTATATCTTTTATTTACTGCTTTAATAAATTTTGGTAGTTGTCTAGGT
GAAACAAAAATTTCCATAAGAGACGATTTTACTTACTACTTTATTACTAGAAAACAGTCT
GGTGCACTTCCTGAATTTAGTTATCAATGATTAGGGTGGATCTTTCATTCTTTTGCATGA
TGTTGATGATTGATGTTTAAATGTTGTGTATGGTTTATATAATGTATAAGTAATATTCTT
TACTCTACCAGTGTTATTTTCTTGGATATGATTTCTCAGTTCATATGTTTCCACGTGGAT
ACTTACTCATTGTTGCGTAGGTAGATGAAAAAGATGGAAACACTTGAAGAGGAGATGGAA
TTGGAATTTTAAGCTACTGAGTTATTTACATTATTATATGAGGTAGTTGATGAATGGATC
GGTTGGCATATAAGTTACTTTAGAGTCAATTTTGAACACTACGATAAAACAGTAATAAAT
TTTTCTTATCTTTTAACGAAGCATACGTCGTTTTGCTTACTTTAATTTCACATGTTTATT
ATGTGATAACTTATCTGTTTGTGACTGTTTTAAATTGAAATTTTTTCTTGTATTTTTTAT
TAGATTCTATATTTAAACTCTTTCTCGGTAAAGTGTGTCACATGTTCCCAAATTGTATTT
TAATTTTGTAATTAATTTTCCTTTTTCTATCTCGAAAGGGTAAAATAAAAATTATTTCAT
ATAAATGGAATTTTTTTTATGTGACCTCCAGACTAATTTTATAAAAAGGGTCAAACATGA
AACTTCAATATAGTTAAGAGACATTTTACTAAGTTTGTAA
>45_Medtr6g091605
AACAAAGAAATAGAGAGGATTCAATGAATCATAATTACCGCTAAAATCTTCATTCAAAAC
CTTTAGTTGTCCACAATATTCAACTCAACTGTTTTGTGCAATCAAATGAAACTTGCAAAC
TCGCCCCAATATATTTATTTTTTTATTATTTTAATAAGTTTAGGTAGTTGCCTAGTTGAA
ACGACAACTTCCGTAGAGACGATTTTACTTACTACTCTATTACTTGAAAACGGTTAAGTA
CACTTGCCGAATTTAGTCATTAATGATTCGGGTGGATCTTACATTCTTTTACATGATGTT
GATGATTAATGTTTATATGTTGTGTATGGTTTATATCATGTATAAGTAATATTCTTTACT
CTACTAGTGTTATTTTTTTGGATATGATTTCTCACTTCATATTTTCCACGTGGATACTTA
CTCATTTTGTGTAAGTAGATAAAAAGATGGAAACAATTGAAGAGGGGATGGAATTTTAAG
CTACACTGTTATCTAAAGTTCATTTTTCCACTCCATAGATATTTTATTTACATTATAATG
AGGTTGTCGATGTATGGAACTGCTAGCGTATAAGTTACCTTAGAGTCAATTTCCTATATT
ATGTAGCATGCCCCTTTTTAATTATGTTAATTTTATATGTTTATTATGTGATAACTTATA
TGTTTGTGAATGTTTTAAATTGAAATTTTTTTCTTCTATTTTTTATTATATTCTATATTT
AAACTCTTTATCTGTACTGTGTGTCACATGTTCCCAATTTATTTTATTTTTTTTTATTTTT
TTTTCTATCTTGAAAGGGTAAAATAAAAATATTTTATATAAATGGAATTTTTTTTATATA
AGCTCGTGACTAATTTTCTCCAAGTAACAAACATAACACTTCAATGTAGTTAAGAGACAT
TTTACTAAGACTTATAACAACTATTTAAAGAAATTGTGTTCTTTATACATTATAAATTCA
TCAAATCTTTTTTGGTATGTTTCATGCACAAGCGGAAAAT
>46_Medtr7g071310
GGAAGTTTGAAAGCTAGTTTTGCTGCTGCCCAATCCACAAGGCGCATGGCGTGATCATTC
CTTTGTTGTGTAATCTTGAGAAACCAAAAAGGGAAATGTTTTGAAACTTAAAGGATAAGA
GTTGAGTTCTTTTTTCTGCGATCATTAGTTTGAGAAATACATTTTGAACACCTTTGGTAG
TGAGATTTCGTTGGAGAAGGGATCTATGGTTTCTCTATGCATCTCTTGAGTGTTAATCTT
GTGGGTCGAGAGATTGGGAAGAAAGGGTAAAAATTAGGGTTTACTCTTATTCATATCTTG
AAGTTAATTTTTTTGTAACTCTTATGCATCTCTTGTAAAGAATCAGTGATGGTGAATCAA
AAAGCTGCTCTCTTGTATCTTTTGTTGCTTTGCTTGGATTGTTTTATTTTGGTTGCCATT
ACTATTTGATCCACTCCGTTTTTTCTTGTTGATTCAAACAATGTTCGAATTCACAACAGA
TTTCATCATCCTTCAAATTTGTTATGCTCTTTTGAGATTTTTTCCTTGTCTACCACGAAG
GGCTAAAAAATGGGGTAAATAATAACAATGTTAAAAGTGAACTTTAATGACCAAAAATAT
ATTATGTGATATCCTTATGAAGTTAAGAGATCAATATTTTTCAATAAGTCCTAACTCAAT
TGAAAAAGGACGAAATTGATATAATGAACGCTGTGATCCATATTTTAATATAAAATTTTC
CACTTGTGTGTGGGAGTTTTCAAATGGTTATTATCTTTGCGTCAGGATTTTCCCACGTTG
TAATCAATATTTAAAAAGAGAGTCAGGTGATCAATTTAATGAGCAAAAGACAAATATGAT
ACATGGATATAACAAAGAGACATTTAGTGTTCTAATTTTAAAATGGATTAGGATTGACAT
AAATAGAGTATAAATTTTTTACATTATAAAATGATCAAACAATTTTTCTTTTTTTTATGC
AAAGCAATTTTCCATTGATTATATGCAAATGAGACAAAAT
>47_Medtr7g098180
AAACATTCTTTGGTCACTATTATAAGAAAAACTTTGCATTTATTTTCAATCACCCTATGT
ATTATATAGCAGTTCAATTGAAAAAAAAAATATATTATATGTAAAGTTAGTTTATTGTTAA
AAAAAAAGTAAAGTTAGTTTAATAGTCTACAAAATAGTAGACGGTAACACACATATGCAT
CTAAAGGTCCTTTAAGTTTTTCATTTTTCCCAAAGTGATCCTTTAAGTTTCAAAAGTCCC
AAACAAATCCTTTAAGTTTCAAAAGTCCCAATTAAATTAGATGCATTTGACCTTTAGATG
CATTTGACGTTCTTTTACACATGCATCCAATTAAATTACATTAGGATACACTTTATTTTG
```

```
TTAGTTCTTAAAATAGTTTCATAAATATCTATAGGGAAATGTTAACCAGTCCCAAGGCAC
TGGTTAAGGACCTAAAATGGTAAACTTTTATGAAAGCTAGTATAATCAGTCCATTGGAAA
TTGAAAAGGTATGACATTTTTAAAGAATAATTGCTTAATTTATAATGCTTAAAAGGGCCC
CCAGTAACACTCGTCAACAAGACTCCTATCTATATGGTAATATATAATTGAATGTATGTG
AAAAAAATTACATACAAATTAAACACATTGTGAAAAGGCTGAAAAGCCATTTTGATGTAT
TTACTATTGTAATCTCATGACTTTTGTACCATGTCAATTTCAAAGATCTAATAGCTTGGT
CGAGAAGGTAAACACCAGCTAATATGCTAATAAAAAATTGATGAAATTGAAAGGAAAAAA
AAAATGACATGAAGAAGGACAAAATAAAAACGGGGCACATTGTTTTTCCTAGTAGTCTTG
TTATGTTAGCTCATACAAAATGCATGATGTATAACTACTTTAGTTTTGGAAAGCTAAATA
TTTAAAATGTTTCCTTAAATATATCAAAAGTTCATTGGAAAAACAAGCTTGTATTAGTCT
CCACACTGAAGAGTCCAAGTGGCTGTTCTTGTTCAATATC
>48_Medtr7g099870
AAATAAAATGTTGAATGATTATATGAAATTTAGAGATGTCCAAATATGATTGTTGCAATT
AGTTTCCTCTTTTATTAATGACTTTTATGTTTAGAGTTTATTATAAGACAAGGCAACTCG
CACAATAGGTAATTAGGTATCCCCCAATTCATGTCCTTACGTTTAGTCATATCCACTTAT
TTATTGGTATTTCGACGATATTCAAGACATTATTACAAAGAAAATGAGGTGTTAAGCGTC
GTAAAATAGTTTTTATACATTTATGTTCAATGAAAATACAATATTATGCAATTATAGTTA
GAATTTAATTGGTATACACTTTTAGTGTAACATTATTTTTAACATACATGTAGTGATGTG
TTGTCACATGAATTGTCATATTCATTTAAAATTTATTTTACATAAATAGCTAATGAAGTG
CTGTCAAATGAATTGTCATATTTATTAATTCACGAGTCAATACAGAACTAAATGCATATG
TAAAATAATTGTACATTATCATATACAATACTTAATATGTTACAAAATGAGTATCACTAG
TATTTATAGGTTGCGAGTGTAAAATTATTTTACACCAACAAAACAAGGTAACATCTATTA
AACTCTTAGTATCAACAAAATGTTACAATACTCCTCTCATTCAAAACAAGAGAAATGATA
TTTGAACAACTATTTTTTACAATTTTTGTGAGAATTTTCTCTCTCATATTTATATTATCT
TTTTACTTTCTCTCTCTATTGCTTTGGTTTTTGTGCCACTACCTTATTTCCTTTGTAAAA
TTTTGATTGTCAGAGAAGTTGTTACATATATGGATGTTCAAATAACACAACTCTAAAAAC
AATCATAGCTCAATAACAAGCTCACGAGACACCAAATTGTTTATCGAAAAGTTTTTAGAG
AACAACATATTTATCCATTCTATCAATTTTAATTCTTAGACACGACAAAAGGACACATTT
ATCTATTACAGAATTCACATTAAAATTGTACGTATTCATC
>49_Medtr7g103390
TTCTTTTAGAGGAGACAATTTTATTTGAGGCGTTTTGGACGTTAAATATATATATCCACT
CTGTTGAAATTCAATCATTGAATCATCACATTATGGTAGTATAGCCCATCTCGCTTAACT
TTGTGGTTATTTGTTTAGAGCTTCTGCCTACCTTACATTAGTATTTTTTTTTTTTTTTTG
AGAGGTACATTAGTAATTTATTGACTAAAGCCTAATACCACTATTCTTTTTGTAAACGGG
TTTAAGGAGAATACAAAATCAATGTAAACCACGACTACACACCATTAATCCTACTCTTGT
AAATTAATTTTAAAATAAATGATGAAAAGTGGCAAAATAAAAGGCTCATGTTTCCTTAAA
AAAAAAAAAAAAAAAGGCTCATGTTAGATGTTTGTGTAAAAAAATTAGTACTCCCTCTGT
TACTATCACAAGAAGTCTACGAGTAGTTTACAAAAATCACATATACTGAGAAAAGTTGTT
GAAACATTAATTTTTTTTTAAAAGTATGTAACTATTAATAAATAAATTGTCATTTATGTA
TAAATAGATTCAATGTTGATTTTTCATATTTTAAGATAATTAATATTATTAATAAGAAGT
ATAGTTGAAGAAAAAAAAAACATCAAAATTTATAAATGGTCTTATATTTAGATACAAACA
TTTTTTTTTAAATGATCTTAGGACACATGAATGACCACCCTTTATTAACATTATTTTGAG
GACTGAAATTACGCATGAGTAATTTTTGGCTTTAACGCTGACTTTTGAAAAGAAACACGT
ATACAAAATCCATCATTGTTCGCGCTTTGAAACGGCGTTTCAGTGGACGTGGTAGTAAAA
GTAGTAAGCTCAACCAGTTTTACACAAAACTTGTGGCTTATTATTCCAAAACACACACTA
ACCAAAACTATGTGCTGTTTCCTTTCTCACCACCTACAAAACCAGTTGCTTCACTTCCCT
TTATATACATACTTTAATACACACTACACATCTTATTCCC
>50_Medtr7g114870
ATGTTAATGTTTTAATTATTATGTTTCTATTTTTCTCATTGCGATGACTTGAATCATGAA
CATTTAACTCTTCAACTCTTTTAATCTTTTACTCAAGCTAGTTGAGCTATTAAACCCATT
AGTTGATGTGAAAGTGGATGATGCTTGGAATCTTCAATTTTCGAAATAGCTAACGATGGT
TAGTTAATATCCTTGACTAGATCAAAGCCAAACCCAAAACACTACAACAGCATAGAGATA
AAAAATCCAAGTCAACACAACATTTTTAATTAATTTGTGCTAATTTATAATTTATGTTAA
CATTAAACTGAATTAGAGTTAAGCAATTAAATACTGCATTTAATCATCAAAATCATAATT
CATCTTCTAAGACTAAAATGGTCACTGATATATGATTAAAAAAATGTTAACATCTTAATT
AATTTTCTTAGCACTTGCGTGCAAGCAATGTATGGCCCTTGAATTCCCGTTAAAGCCTCA
AAGGAACAAAAGAAAACGTTGCAAAACAGTGATGTTTTTACCTTCTTCTTTTATAGAGAG
ACAGTATCATTGAATATTTAAATGTAACCCTCAAACTAGCTGTTAGACCATGATAAAGAA
ACATCTTTTGATTTTAACAGCATGGTCCCCTAAAAATTACTACATACACTAATGGAAATT
TTGTTGGCTCTTGACTAGGTTAAAAGTGTAATAGTCATTTTGATTCTAAATATGTGATAC
ATGGACACAATAATCACCGAATGTATCAAAATTATATAAAATATTTGAGTGTGATTTTAG
TTAGTAATTTTATGAAACAAATTGACTAACAAAAGATACATTTGAGAACCAAATGATTCT
TCAAAATAAAAAATTAATGAGACACAAATTTGAAAACTAAACAAACTAATGAAACATAAA
TTTATGTATATTTTTGTAATTTTGATACATTCAAAAACGACTGTGACAGCGCAACGTTTG
AAGAACAAAATTACTATTTACTCTCGTCGAATTATATATA
>51_Medtr8g039130
CAAAAGTGATAAACCATCTTAGTTTAAAAAGTCGATAAGAGAAGCGTTAAACACTTAACA
CTTTAATTAATCAGTTGAGGCAATATCTTTAAGAAGGTAGGGTCGATTCAGATTAAGTTG
ACAACTTTACAACAAAAAAACAAAAAACTCAACGAAGTAGTTGAAGCCAAGTTGTTTAAT
GCAAAAAACATAAGAATTAGATACCTTGTAGAAAATCCAGTGTATAGAAGTCAAGCAGGA
CATACTCTGGTAATTGTAGTTGGGTTATTATAATGGAGTACTAGTGCAAATCAGGGATAC
AAGACCTGTATTTTGGAGATTTAAGATCGTTTTTACATAAAAATAGGTCTGAAATCTTAA
AAAAAAAAAAGATCTTGTATCCCTGATTTGCACTAGTTAAATCCTTGAAAAGTATCGAGA
CAAACTCTCTCTTTAGAGGGTACTAGAGGTAGTAGTCAGGTGGATTGTGAACTAGTATAA
ACATATTGTGCAATTTTATTTCCCTTTTGATCTATTTTATTTAGATGCACAATGTTCTTT
CTGACTTAACTCTAACACTAAGTTTTCAACATAGTTAAATGATTCGATAAATATTTTAAA
AAAATCTGTTAAGAAATATTTTCACACAAGATCTTTCATAACTTAATTTAAAAGACAAAA
CTTTTAAATTGTAACTATAATTCGAACCCACATTTTTTTTGCTCTCTAACCTTCATGAGT
TTGAATGTACGTTGACAATAATTTCACAAGAGCACACAATTAAATACATATACTTCACAT
AATACTCCATATTCAAAATTATCACAACACGTATAATTGCTTACAATAATCAGGGTGCGT
GTCTTGTCTATCATGAAAGGGTAAAAGATGAATAAAAATAGTAATAATGAATAACTAATA
```

```
AAATAATGAGCTATTACGACCATCAAGAGATAATTAATAATAATTATAATTGGAAGATTG
AAGTTTTTCACACTATAAAGTGATCATGAATATTTATATT
>52_Medtr8g044120
CCCAAATATCATGGGAAAGATCAAAAGAAAGATAGTGAACATGAAGATTATCATCTGTAA
TACATCTAACATGTTCTTGGACAAATGAATCAAGTTAAATCTAAAAGAATTTTAAAATGA
CAAGTTAGCGTGCCACCTAATAAATTCTACATACCATGTAGGCTTACTTACATAGTCACA
TGAAATAAGAGTTTAAAATAAAAAAAATATTATATTATAGGGTTGAAATATAAACTTTAT
TTTTACATAGGTAAATCGAAAATGATCTTTATTAAAGGAAGTAAATATATATTAACAAAA
AAACTAATTATGACAAACACATACTAATATTTAATAAGTAAATTTGTAGTCATACAAAAC
CAAGCTTTCTTGCTTACATAGAAAGAATCTTGAAAAATCTCACTCAATCAAATAGTTATT
TAACCTTTTGTTTGCTGTAAAAACATGTGATCACGATGATAACCGATAAACATCTCTAAA
CAAAATCAAAATCCAAAAGGTGACACAAAAACAAATAAAAAATTTAAAATAAAAACAATA
AAATAATATTAAATAATAACAAGGCCAATGTTATGTTGTCCCTCAAAAGTTTATTGAGAC
ATATTTCAGTTAAAAAATAAAACAAATAAGTTCTTTGCATTTTCATGTAGGAGATGAATC
AGAATTTATGATTACTTAAGCAAATAGTCCCTAAAAATCTTTAAAATTTTGTTTTTAATT
CTTAAAAATAAAAATTACACTTTTTAATCCTTGTAATATTTTTTTTTAATTTAGCATTTA
GAGACAAACGATGTTAAATGAAGTTTTTCATAGAGACTAAAAATATTTGTGAAATTTTTT
ACAATAACTAAAAATGTTAGACGAAATTTTTTATTGTAAGGACTAAAAACGAAATTTGAA
ATATTGATAAAGAACATTTACATTTTTGACCCTAGTTTTTATTGTTAAAATTAAAGTCAA
TATAAAAATGTGATAGTTTACTTTTGGTTATTATTTATTT
>53_Medtr8g064070
TATACTATTGGGTATTTGGCTTCTACCGTTGCAGGATAGCAACCAATGAATATTAATAAA
GATCCATCCAATCCATATTTTTGGTAGTATTACCCATATTGGATGGGGAGAAGCCAATTC
TGAATAATAAGAGAACTTTCTAACCCTACAAGAAGGAAGGTTGTAGATGAGACTTTTCTT
TGGCTTATGAAGGCATCCCAACAAAACATTGAATATTTTGCAAGAATTTAACAAGGATCA
AACATCTTCGTAAACCAATGTTAATAACTAATTCATTGTTCTTAATCATGCTTTGTCATT
TGAGATCTTTTAAGTATTTTACTACTTTTTCAACTTTTATTGTCTTCAACTTTATTTTAA
GTGTTGTTACTTTTCATGATTTATTTTATTTTTATGCAAACAAAATTTTGTTCTTGTTAT
ACTATTACTTGAGATAATGTATACTTGCACTAGTTAGACTCTTGATATAAAGTAGTTTTT
TCAGTATTGGCAGTTAACTCATACAATTTCTCCATTTGATATCTATTAACTTTTATATAA
ATTAAATGTTGGAAAAAATTGATCTTATGGTACTTTTAATTTATAAATAAATGTGTTGTC
ACCTCTTCTAGGGTTATTGATATAACTAATTTTTTTTTTTTTAATTCATCACTATTATTTC
TCTTTTGAATTGTGTCCCTATCTATTATGAAAGGCTAAAATGTAATATATATATATATAT
ATATATATATATATATATATATATATATATATATATATATATATATNAGTGAGATCTT
TACTAAGGTTAAATGGTCAATTGAGTGTTATTGTGGATAAACATGACATATTGATTAAAG
ATATTTTATATTAAATTTCAATTTAATAAATGGCTTCTAGCAACTACGAAGTACTAAAAA
CTATTTAATTGACGCTAACAAAGCATGACAATTTTACTTATAAGTGATCAAAAAGAATAA
TTTTTATTACATACAAACGAAGAAAAAATATGGATTTGGT
>54_Medtr8g069310
TTATCTTGGGATTGTTGGATAAGTCTAATGACATAATCCTAATCAAACCAAACCATAAAC
CTTAGTTTTATATAAATCTACTTTCTAATCTAATTTCACAGGTTTTATAATGAATTAGCA
ATTAGTAAGCTAGGAAGCGAACACGGAGGGTTTAGCACTTGTTAATTCATCATCACAAAT
ATTAGGCGAAGGATCGAAAGATGAAGTTTTAATATTGGAACAAGTGAAATTAGAAAAGGG
TTGCGAAACATGAGAATAGTCTAGCAAACCTTGAGACATATAAAGTTTTGTTCTTCAAGG
ATTCTAATAGCAATCAACCATGAGAATAGGCGTGGTTGCAAGAAGAACACACTCACTGAG
ACCGAAAGGAGATGTGATAGGCTAAAGAGAAACTTGTCTTTAGAAAGTAGGTCAAATATA
AAGTTAGAGGTTTAGGGTTTGTTTGTGAAGAATTCGTTGTTATGATGAACCAATAATCCT
AAGGCTTCTTTTATTATTATTTTCTTCCGTTAAAAATCAAACCTTTTTCAACTGAAACTT
CTTCTAATCATTCTCTAAAAGTTAGTTAAATTGAACAACTTCCTATCAGAACGATACTCT
TTTATTACTACACCGGTATAACCGTGCACTTGCGGTTTTAGATATCAATCGAGTGAACTA
AATGAACAAAAACTCATTGCATTTTTGCACAATAAAATAATCAAGTGTTTTTTATCTGTT
TAATAATATGTAGAAAAAGTAGGACTAAAACTGTCAAGTTTATTATTGTTTTTTCAAAAG
ACAACGATCAAGTTTATTTATACATTAACTTGTATGATAGGATACTTTTGAATCTGCTAT
AATAATATTTAATAATGATTCGACTTTTAGGAAGAGTTTTAAACTAAAATCTATATATTT
AAAAAATAAAAGCAAATATAGCACATAGTATTTTTCTATTGTTGATGTAGACATATTTTT
TACATTATAAAGTTATCAAACAATTTTTATTGTTACATAC
>55_Medtr8g090375
ATCCTTGAGACCTCTTAACGTTACACCAAGGCATCGAACTTTCAACCTCTGAGGTTCTCC
GCTGTTGAATTGTGCATGAACTTCGATCCACCCGGCCATTGTTTCAAATCTACACAATTA
GAAATTAAAAATAATCAATGAAAAACTCATTCAAACACTTGTAGTGAAAATTTGACATAA
ACCCTAAAAATTCTAAACAAACCCTAAAATTTTTATAAAATTCGAAAAAAAAATATAACT
TTAAGCAATATAACTTCATTATAACATGCATTCATAATATATGTTAATAGCATTTCAGAT
CAATAACATAACTATATTATAAACAAATATTCAAACCCTAAAAAATTTATAAAATTCGAA
AAAAATATAAATTCAAGCAATATAACTTCATTATAACATGTATTCATAATATATGTTAAT
AGCATTTCAGATCTACAATATAACTATATTAAAAACAAAAATTCAAAATTTAATAATTAA
ACAACAACTAAAATTATAAACATATAAACACATATACTAACTAAAATGTATGACAATAAC
ATAAGTGATAACTTGCTTGTGATTTAGTTGAAGAAATTTTGGATGATTTGATAGAGTCTT
TTTTAGAGAGTGACTGAAAATTTGAGAGAAAGACGGTGAAGGTGATGATAGAGTGAGTGA
GTGTCTGGATTATGGTTACTGAGTTAATATATCACAAATACAACGACAATTAACTGCCGC
TGACTTTCAGCGGTTATTAACTGCCGACGCATTGCCTAACGACAGTTAACTGCCGCTGGA
ATAAAAATGTCATTTATTTGTGTCAATAGGAAATATTTCTACATGAGAGAAGCAATTTTC
AATTGAGTATTCCCCTTAGAGAAAAGTTAAGTTTGGGTGGGGTGAAGAGGTAGTTGTAAT
AAAAAAGGGAAAGAAACACATTGAAAGGCGCGCCCGTAATCCTTCTATTCGTTACCATCA
TATTAAATAATGATTTTTTCTCCTTATATTCCTTCAAAGC
```

## 2. 38 Nitrite Reductase (NiR) promoters used to identify the NRE shown in Fig 3A, S5, and S13A

```
>Medicago_truncatula Medtr4g086020 NiR chr4:33643564..33646615 forward
ATTCATATCAAATTTCAACGTCGTTCATATTGATTGGATGATTAACAAATATAACAAACTATCCAACTAACTAACTAACCACCAACTGAATACATTA
ACTATAACTGTAAATGTTGTTAACCATATTTTTCAGATGTATATATATCATCCGTTGAATGTAATTAATCATATATTTAAACTAAATTATCCTAAAA
```

```
ACTCGGTTCGAGACTTGGGCAGAAAATGGACGAAAATGTAATGAAACTTTAATTTTGACTTTGACACCACACTACATGTGCCTTTTACATATAGTTT
ATATGACAAGTAATGACAATCATGCTCTATTATAAGGTGACCCTTAGCTCTAACCAAAGGACGATGGAGTTGAGAAAGAAACTCTTGCTTACTTGTA
AGGTCCACACTTCTTCACTCACCTCCTCTCAAATTCATCCTACAAAAATGTCCAAACTTCTCTTTCTCACAATCACAAACTCATTCCACACACACTC
TCTTCTCCAAAAATG
>Glycine_max Glyma.07G212800 | Chr07:38482588..38486164 reverse
TTACTTTTAATTCTTTAGACTAAAATTAAGAATGAAAATAAAAAATTAAACAATTAAAAGTAAGGAAAAAAAATTAAGTGACGTTGACTAGTCAAAA
ACAATTTATAGACTTAAAACAGAAATTTTTAAATATTTTAGAGATTAAAGAAGCAATAAATATTTGTTTTATATTTTAATGAAATTTAAAAGAAGAA
ACAAACTAACCAAATATTTTGGTGTAAAAAAGTATAGAGAGAATGTGTGTAGCATGCAATTTTGAGAGGGAATTTTAGTTAAAAGTGTTTTGTTTTC
ATGAGGAAGGTGACCCTTGGAAGCACCGAAGGGTCTATCGATCCGAGGAAGATCACCTCTGACCTCTTGCTTACTTGTAAGGTCCTCGTTCACTCTG
GTCCTCTCGTCCTACAAAATGTACACTCTTCTCTCAAACCGCCATTACCATAATAACAACCTCAAAAACTCAATATTCTCAACAACACACTACTTTC
CTTCTTCCAATAATG
>Populus_trichocarpa Potri.004G140800 | Chr04:16399188..16402333 forward
TCCTCCATCGTAATTGATGTTCAAAGCTCAAATTTCATCAATTCAAGGCTTAGGTATCCCATTTAACACGAAAGAAGAAGGGTTTAATAGAGTTTTG
TTTTCTTTTGCACAAGGCATACAAGTTTCAGTAGTTTGGACTCCAAAATTATCCACCGCCAAATATGCCCAATCAGATGTTCATGCTAAGCATCTTA
ACATGATCTCCATATATATATATATATATATATATATATATATATATATATATATATATTTGTTTGAAAAAAATAAATCAACTATCATCCTCCCTAGCT
CCGACCTCTACTAAACAAACAAATTAGATGACCCTTGGCACCTCCCAAGAGTCCCCTCCTTCAGGAAGAAAAACAGCCCTCTCCTCTCCTCCTAAGG
TCCCATACATCATCTACAAAATGCCCTCTCTCCTCCTTCTCACCACCATTACTTCTCTCTACAACACTAACAAAAATCCACACTTCTCTAGAAACTA
TCTACCATCATTATG
>Fragaria_vesca gene28260-v1.0-hybrid | LG3:21015246..21017561 forward
AGAACTTTGCGTCCCCAAAATCTATCCAATACAACAATCAATGATGGCCGAAAAGTCGAAACCAACAAAACAGCCAAACGCCAAACAAGTATCCTTC
CTCCCCAGACGCCACATCCAAACCACACGACTTGCCGCAGAGCCTGTAGCCATCAGGCTTCAAAGTTGAAAAGCCTTGGCAAATAATTAAACCAAGG
GGCACATGCCAGACAGACTCACGTTACGTTAAAACGTAAAACAATGTCGGTTAACCCAAGGGACAAAATGGTATTTACAGAGACCAACGGTGACCGT
TTCACCCCCGGGAAATAGTTTCTTAAACAAAGGCGGTGACCTTTGGCACAACCCAAAAGTCCATCCTCACCCAAAATGAACCTCTCCCTCCACCTAA
GGTCCTCACCGCTGCGCCCACATTCTTCCCCTACAAAAATGGCCACACTCTTCTCTCTTTCCTTCTTATCATCACCATCAACTTCTGAATCCAAAGC
ATTCAATCTGCCATG
>Prunus_persica Prupe.8G126000 | Pp08:15056933..15060044 reverse
TAAAAGAGAGATCGGGATGATGGAGAATTGTATCCCTTATAGTTTAACTATCGCTTGTAGTCTTATTTTCTATTTTTTGAAAAAATATATAAAAAAT
CAAGTTGGACAACTTGTCTACAAAACACATCTCCGAAACGTGGCACATGGTGGACAGGGGCCACCGTAACCCAAATGCATGAAAAAACAAAGGCAAA
CAACAAATAAAAACTTTGGAGGAAAGGGATATACAATGGTATATACAATATAACCGTCACCAGTCCCCCCATTTCCTCTCCAAGACAAAATACGACG
ACCCTTGGCTCCAGCCAAAAGCCAAAAGCCAAAAGCCCCTCCGTCACAACCAAAAATGAACGTCTCCCTCCTCCTAAGGTCCTCACCGCTAAGCTCA
CCTACAAAAATGCAAGCCCTCTTCTCTCTTTCAATTCCTACTTCAACAACGCCATCAAAATTTGGAGCTCTTTGCAATCATCCGCACCAGCAGCATA
CTCTGACCAGCCATG
>Ricinus_communis 29801.t000127 | 29801:703327..705962 forward
CAGTAATTAAGTTAAGAATTTATTTGATGTTATTATATTATATATATATATATATATATATATATATTCATGCAAACTACTTTGCACATAAACAATA
TCCACAATGAATAAGGGCCAGTTAATCTCATGCACATGCAATCAACACTATTTTAATATCTTGAATTCCAAAAACCAAAAGTAAAACAATAAGAAAA
GAGGCAGAAAAAATAATTAATTTAATAAAGAAAAAGGAAATTAAACAAGTGTCATAACTGTCAGCTTCTTATAGTTGCCACTCAATATCCTCAAAGA
AGCAAAGATGACCCTTAAGGCCTCTCAAGAGTCTCCTCCAAATAGAAAACAAAAACAATTACACAACAACACTCTCTTCCTATCCCTACAAAATGCC
CTTCTCTTCTTCTTCTTTCCTCCACCACCATTTCTTTCTTCTCTCACAACCATCACCACTTCTTCATATTCCCATTTCTTGGGCTTAGTTTTCTTTT
ACAGAACTTACCATG
>Eucalyptus_grandis Eucgr.I01148 | Chr09:22573933..22577360 forward
TCTCTCACGAGTGGCTTCATCACCTCCTTTTACAAATACTAATGGAGAAACTAAACGAAGCTTAGCTAGTTTTAAAATGACCCCTTGACCAATAGCA
AAAGTCACTTGATCCTGCAGCCCAAGCAGCAGCCACCTCCTCCTGACCCTGATATAGTAAAGACCAAAGAAGCAAAATGAAGCGGCAATAAAAAATA
CAGGAAGTACTTGGATAATACCATGCAGTAATAAGGACAACACAAACATTTTATCCCTTTGCTCACCGTGCATTCTCCGTCTGACCTCAAGGGGACC
AAAAAAAGGAAAAATAATGAATGACCTTTAGCAACACCCAAGAGGCCATTCAATCAGAGCTTCTCTTCCTCTCCTTTCCACCTCTCCTAAGGTCCCA
ACAATAAGCTTCATAAGGCCTTGAAAATGCTAACATCTAGATCTCACTCCATTGAAGCACATCACAAACAACAAGTTCTTGATCTTGAGCAACACAT
AAGCATTGCAACATG
>Malus_domestica MDP0000722117 | MDC002248.202:26732..29166 reverse
ACGTACACATTGGCAGAGATTATGCAGTTTTAAAGGCTTAAGACAATTTGCAAATTTCATACTTGGCCAACCACAACGGCTTAATCTTTTCCAGCAT
GATATGAAATGTCCACTTTAACCGACTTAAGCTTAAGAAGAGCATGGAAAAATAAGTTTGATAAACACATCTCCATAACGTGGCACATGCTTGACAC
ACCACCCCACAAAGACAATTATCCAATTGGCATTTGGTATGAAAAAACAAAACAGAAACAAATTGGTGAACACAAACCCAAAACACAGATGACCCTT
GGCACCACCCAAGAGTCCACTCTCAACAAAAAAAAAAAAKCRWCTCCCTCCTCCTAAGGTCCTCACCAGCCACCGCTGATCCTACACCTTCACCTACA
AAAATGAACACCCTCCTCGTCTCTCTCTCAATTCCCACTTCACCTCCACCAAAATCTTCAGCTCTTTGCTATCATTACCTGCGGCAGCATACTTCTA
GTACTATCAACCATG
>Arabidopsis_thaliana AT2G15620 | Chr2:6810476..6812939 forward
AAGTGTATGGTCGGTGTCAATAACTAATATGTCATATTTTTTACAGACTGTTCAGATATGATAAAAATAGTTTTTTTCCAACCAAATGGTTGGTCTT
CTGAGTGTATTTGAAAACGACATTGAAGAATTAATATATTTTTTTTTTAATTTTAGTTTTTTATAGTACAAATATTAAAACAAACAATCCTACCATAT
CATAACATTTGTAAATAACATTTTAAGTTTTGTTTTGAGTTTTAATTAATTTTCTATGACAAAAAAATGAAGTCAATAGACTAAGTGAATCATATAG
TATAAATAAACACAATTTAAATAGTTTCAAATAAATTTAGAAAGAATAAAACAAATAGAAATCAGAAGGTGTCTGTTTCCTCCTCGCAACATACGAT
CAAAGAGAAACAACTTGACCCTTTACATTGCTCAAGAGCTCATCTCTTCCCTCTACAAAAATGGCCGCACGTCTCCAACCTTCTCCCAACTCCTTCT
TCCGCCATCATCATG
>Brassica_rapa Brara.I00981 | A09:5519502..5521587 forward
CTAATATATTTATTTTAATTTTTAACGGTAACCAGCAGAACATTTATTAAATATTGTATAGTATTTTTTCCACAGAATATTTTTAAAATAGTATATA
AAGTTTTTTTTTTTGGGTCTAAATGTTAAGAATAGTAGTAGTAGTAGTAGTATATAAAGTTCCGTACTAATACCCTTAATTCATTTAAATGGAAGAA
AATAATAATATAATTTATTACCATGCCAATATGAAACAGATACAGCTAAATTTGGTTCGAAGTCAAAGCTTCCTCGATTAAGAAATTGTAGAAAATA
AAGAGTTTCATTAAATTCGAATCTAAAAATTGATTTCAGTTTCATTTATATTTAAAAACAAAAAATCTTGGAAGTGTCTGTTTCCTGCTCACAACAG
ATGATGAGAAGAAACTTGACCCTTTGGCTAGCTCAAGAGCTCATCTCTCCTTGTACAAAAATGGCCACACGTATCCTACCTTCTCTCACCATCTCAT
CCCTCCTCCACCATG
>Manihot_esculenta Manes.10G071700 | Chromosome10:11502524..11505231 reverse
AAAACTCTTCAAATCCCCCATCGTGTCTCCTTGTGCGGTCAGGGAAGAACGTCACCCGCTTATTCTTTTTTAATTTTTAAATTTTACATTTTTCGGT
CACTTCTTTTTATTTTAAAATATAATATTAGGCCTGTTTAATTAAGAAAATTACTTTATAAATGGAGCCGATCTCTATTTTGTAATGCACATGCAGC
CATTCACAAGAACTAGGCCAATATAGAAAACCAAAAGTAAAATAAACCTAATCAAATCTAACAGTCACCTTCTCAGTTTCCCTCCAAGTCAAGATCC
TTCTCACAGTGGTACAGCCAAAGAACATGACCCTTAAGGCCTCTCAAGAGTCCCCTCCATAAAAAAACAAACAAAAAATACTGCAACCTCCTCCTCT
TCTCCTCCTACAAAATGCCTTCCACTTCTTTCCTCCACCATCATCTCTATCTCCTTCTTCACTCCATTGCTTCTTGTTCTAGCATTTTACACAAGTT
AGATTCAGCATCATG
>Citrus_clementina Ciclev10028067m.g | scaffold_8:3069766..3072920 forward
GAAAAAAAAAAAGAAGAAAAGAAACAATTTTTTTTTTTTAAAAAAGAAAAAAACAAGTTCATTTGATCTAGAGCTTGAACACCAGTTATTCCCACGG
CGCATGTTGTCATCTGCCCACCAAATTTAGAAAGCAAAAGAGAAGTAAAACTAAAACAGACAGCAACAGTTGACAATGAGATAATTAATATAGTTGA
ATGACAAAATAAATAAGTAAATAAATAAATAAAACAATAGGCAACTGTCTGTCAGCTTCTCACAATCTAAAAACTTGGGAAAAAAAAAAAAAAAAAGG
TAGAAAAACACCAACAAACTTGACCCTTAGCCCCATCCAAGAGTTCCCCATGAACAAATCCTTACCTCTCCTCTGCATCTAAAGGTCCCTTTAGTGT
CCCTTACAAAAATGGCCACACTCTCATTCGTAATTCTCATAACGTTACTACAACAACTACAAGTTTCATCAACAAGGCTTCTCATCTTTTTCATCTT
GCAATTTCAATCATG
```

```
>Carica_papaya evm.TU.supercontig_62.8 | supercontig_62:53541..56287 forward
AATACCAAGAAAATCAGAAAGTGCCAAAGCAAAAAAGAGCACAAGGTATTTCCTAGAACCAGAAGAACCATGAAAAAACATTTGTTTTTCTTGTGGG
GGGTGGAGGGGGGGGGGGGGAAGGAAAAAAAAAAACCCTCAGCCACCACCCAAAAACAATCCACATCACATGCCACTTGTCGGCATATTCTTTTTACG
CCAACAAAACCGCGGGCTAAACGGTGAGGTCTATTAAACAAAACGACACAATTAAGAAGGAAAACGTCCACTTGTACGTCTCTCTATTTTCCGTCCT
CTGAATCTGATCTTCAACACACACAAAGAGAAAAACTGACCTTTAACACCACCCGAAAGGCCATTATTAATCCTTCTTAAGGTCCCCCCTCTCTCTC
TCTTCAAAAATGGTTCTCCTTATATCACAACCAATCGCTATTTTCGTTGGCTATAACACTTTCTGCTCTCACTCTCGCCTTGTTGTTTCTGCTATAC
ACGACCACCGCCATG
>Citrus_sinesis orange1.1g047572m.g | scaffold00065:114542..117209 reverse
GTAGGAAAAAAAAAAAGAAGAAAAGAAACAATTTTTTTTTTTTTAAAAAGAAAAAAACAAGTTCATTTGATCTAGAGCTTGAACACCAGTTATTCCCA
CGGCGCATGTTGTCATCTGCCCACCAAATTTAGAAAGCAAAAGAGAAGTAAAACTAAAACAGACAGCAACAGTTGACAATGAGATAATTAATATAGT
TGAATGACAAAATAAATAAGTAAATAAATAAATAAAACAATAGGCAACTGTCTGTCAGCTTCTCACAATCTAAAAACTTGGGAAAAAAAAAAAAAAGG
TAGAAAAACACCAACAAACTTGACCCTTAGCCCCATCCAAGAGTTCCCCATGAACAAATCCTTACCTCTCCTCTGCATCTAAAGGTCCCTTTAGTGT
CCCTTACAAAAATGGCCACACTCTCATTCGTAATTCTCATAACGTTACTACAACAACTACAAGTTTCATCAACAAGGCTTCTCATCTTTTTCATCTT
GCAATTTCAATCATG
>Theobroma_cacao Thecc1EG007067 | scaffold_2:5126076..5129440 reverse
ATATATATATAAAAGAAAAAGAAATAGAGCACTAGCCTCCTTAGTTCCTTGTCTTTCTTCGTCCTGAAGACAACCGCAACTCAATCTTCAATCCCTC
AAGTGACAATTAATTAGAAGGAGATGGAGATGGGAAGAACAAAGCCTAGTCATCACATGCCACATGTCGTGAAACTCATATTAACTAAATTTAGAAA
GTGGAAAGTTGGGGTATGGTAATAGGAATAGCAGTTGAAGGTAAATATTTAGTAATAAAAGTAGAAGACAAACTAACCCTCATCTCCTTCATCCAAA
GTCGTACTAAACAAACTGGCCCTTCACACCATCCAAAAGTCCCCTCCATTCCCCCCCCCCCAAAAAAAAAATAACCCTCTGCTATTCTCCTAAGGTC
CCCATCTTTCCACTACAAAAATGTCCAAACTCTTGTCTTCCTCCGCCAGAATTCTTGCTGTTCCAAACACTTTGTATCTCAGTTTCAACTTTTCGAA
TAACTATCTCAAATG
>Aquilegia_coerulea Aqcoe6G141000 | Chr_06:7990174..7993109 forward
GGCATTTCTCAACAGTCACCTTAAAGGCTAAAATGGTACTAAGTAGAAGTGGAGGGGTGAATCATCAATCATCTCAGGTAGCAAAAAAGCTAAATGG
GAAAAAATCCCACAACTAAAATGTCGGCATAGTATGCGAAAAGCCTATACTACCCTTTGTAGAATATTTAGATTGAGCATTTGAGCTACATAACCTT
TGATAGAGTCATGACTCTTGGTTATTTTATAAAGCCAATATCTCTCTTCCACACACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAGGTCAAGGAGGGTA
TTCAAGTAAATTCATTATCATCACTTTGTCACTAACTCACCACACGCCACTTAGCAGCTAAGTAAGCTCACAGCGACCTTTAGCATCACCCAAGAGA
CAAAAAAGAAAAATTAAACAAAGAAAACAACAAACCATCACCACACTACAAAAAAAACCTCTGATCCAAGAATGAACTCTGCAGACTTTCTTCTACC
TTTCTTTCAACAATG
>Amaranthus_hypochondriacus AHYPO_012823 | scaffold_159:37637..41037 reverse
AAAATTTATATTAAATTCACTATGAAAGGAAAAAAAAATGACGTTTCTAATATCCCGCACAAGGTAGAGTCCAAGTGGGGGTTTTCCCTTTTTCCTAA
AAGATGCGGATAAATCATACTCATTCCCCTAAGTAAAAATCGAACTGGTATCTACACCCGGTAAAAATCGAACCCCAAACCATTTGTTGTACGTGCA
TGCTCTATAGATCACCAAAATAGTAGAACCACAATCAATCTCAAAAGAAGTGTCCCTTAGCCATGTCCAAAAGACCTTGTTTGTGCTTACTATTCCC
AACTTCCAAGCAATCAACCATGGAAACATGTCGCCACACTCCCTCTCTTTGGCCTTTTCCCACCTCTTCCATTCAACCATTTTGTCTTTCTCTTTTC
TCATTTTCTGACCTCATTCTTCAAATCACAAACAATCTGATTCTCACCACTATATTATCTCTTCATCTTCCATTTCCACACTCCATAATCATCATCA
AAGTAATCAACAATG
>Eutrema_salsugineum Thhalv10022613m.g | scaffold_11:5741964..5744643 reverse
TCGTAATTAGCAATGATTAATATTAATTACCATAAATATCTCAGAGACGTTGAAGTGTGAATATTTTATAGTTTAACTAAAATACTCGAATATTTAA
ATATATTACAGTGTTGTGTTGTCCACTGAACTATAAATATTTGTGTCTACAGTAATGTCCGCAAGGTTTTTCATTTGACGTAAGAAAAAAAAAAATAT
TATAGTTTTATGAAACAAATTATACCGCTTATTTTGGTTCGAAGTCAAAAGCTGTCTATCAAGAAAAAACAATGGAGAAAATGAAAGAAAAAAAATCA
TTCAAATTTTTTAAAGAATTGATTTCAGTGTAATTTTAAAACCAAAAACAAAAAAAAAAAACAAATAGAAATCGGGAAGTGTCTGTTTCCTCCTCAC
AACAGATGATCAGAAGAGACTTGACCCTTTGGATTGCTCAAGAGCTCATCTCTTCCCTCTACAAAAATGGCCGCACGTCTCCAACATTCTTCCACTT
CCTTCTTCCGCCATG
>Mimulus_guttatus Migut.B00351 | scaffold_2:1625417..1628610 reverse
TTATACTCGAAAGAAATGAGAGATATTTTTTTTATATGTCAAAGAAATTGTGGTCGGTGAACTAAGGAACAGAAAGAACCCAAATGCAGTGCCAAGA
GACTCCCATTCCACTACTTTCATTATAGGCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTATATCTTTCTCTCTCTACAATTTTCTTATCTCCAAGCGTC
GCCACCGCCACAAGCCATTTTTTTGTCGCACAATTCCCATAATAACCCCACCCATTTCCACGCATTCACGTGGAGCACCTTCTGGGACCCACGCTTG
ATAAAATAACTAACACAGGGGCATTTTCGACAAACCACACCAACTGTCAGCATTAGCATAGACAAATGGCCTTTAGCAGTAATCAAGAGCCCAATAT
ACCATTCCCCCCACACTACAAAACTATCACCTCAGTCCATTGTCCCTCTACCACCACCACCATCACCGCCGCCGGAACTAGCTGAAACCTCCACCAC
CACCGCCTAAACATG
>Boechera_stricta Bostr.18351s0035 | Scaffold18351:398503..400823 forward
GTTGTGTGAAACATGTTAAAATGATACTCTTTGTGATTCTGAGGGAGAATACATTTAACGGATTTGAATTCGTGTGTTTATGATATCCATTGAAACT
CTTAAATAAACTTAGGTAACATATTTATTTTGAAATATGATACAAAATCGAACTAACAAAAAATAAACAACATTATTAGCATGCGAATCATCATATA
TATATATATATATCAAGTTAAAAACCTGAATTATTTTTTTTTTATAATTATAATTATGATGCTAAATTGCTAATTAGTATATTTGACAAAAAGTCAAT
GGACAAAGTGAATAATCATATATTAATACAAAAATTTAAATATTTGTTCAATGAATAAAACAAAGAGAGAAATCGGAAGTGTCTGTTTCGTCCTCAC
AACATACGATTCAGAGAAACTTGACCCTTTGGATTACTCAAGAGCTCATCTCTTCCCTCTACAAAAATGGCCGCACGTCTCCAACCTTCTCCTACCT
TCTCCTTGTGCCATG
>Linum_usitatissimum Lus10014059.g | scaffold1247:211033..213455 reverse
ACATATTGGCTTTTAATTTACAGACGCAAAAAATTTCCCCTTGTACGAATTTAAGAAATTAAACACCTTTGGCTGAGATCGAACTTGCGCCGGATTT
GGTGATAGATGGAGTAGTGTACTCTACCGCTTATAGGTTGTTTAAATAATAAACAATAAGTAAAAATTCAATATTCTAAATTTCTCGAAAAAGGGAA
AATAAGAAAGAAAACAGAGAGGGGCCGAGTAGGCTAAAAGAGCCAAAGCAGCAGAGACTACGATATCTATAAAACAATCTCTCATCATGTTGTAATC
TGTCTGTCCATTCCACATCCTCAAACAATGGCGGTGACCTTTGGCTCCTCCCAAAGGTTCCATCTTTGCCCACCACACGCCCCTCTTCCCGCCGCCT
CCTCCTCCATTTCCTCCTCCTCTCTACAAAACCTCCCTTTTCCTCTCTCAATCTTCTTCTACAACCAATCTCTCATCTCCAATTCCAAAATAAACCA
GAGACCACCATCATG
>Capsella_rubella Carubv10013302m.g | scaffold_3:12391763..12394307 forward
AAACTCGATTATTATGTTACTTGTTAGAAAAATTTGTATTATGCGAATTTAAATTCCGATGTTAAATTATTATAAATTTACGGAGAAAAATAATTAT
AAAATTCATATAATGATATTTTGTAACATAGTTCTATTTTACATATAAGATAAAACAAAAATTAAATATTATTAGCAATCGGATCATCAAATAATAT
CTCATGTTGAATCATCATAGTTATAAAAAATTAAAAATATGATGCTAATTTGTATTTGACACAAAGTCAATGGAGGAAGTGAATAATTATACATATT
AATTACAAAATTCAAATTTAAACAAAAGAAAAAGAAAGAAAGAATAAAACAAAGAGAGTAAATCAGAAGTGTCTGTTTCGTCCTCACAACACACACA
TACGGATTCAGAGAAACTTGACCCTTTGGATTACTCAAGAGCTCATCTCTTCCCTCTACAAAAATGGCCGCACGTCTCAAACCTTCTCTTACATCCT
TCTTCCTCCATCATG
>Zea_mays GRMZM2G079381 | 4:178613063..178616618 forward
AAGCAGCGGCCGAAACTGGGGCAGCTCTTTTCCGCCCCTGTCTCCCCCCTGGTAACGGTTTTGGTGGTCTCCGGGTCCACTTGGATAGATGTGAACG
TGTCAGTATGTACTCTCCAGAAAATCGAGCTTCCATTGCTTTAGAGAGGCAGATGCATATATAATTCTTGAAACGCCGCGCTGTTTCTCAACGTAGG
TAGCCTCTGTCAACGCGATACAGCAGTAAACATGAGTACAACGTGAGCTCAGAGCCGTCCATGGCGTACGAGGGCACGTCGGTAAAATCTCAACGTC
CCGTCCCGTGCCTGCCTGCCATTTCGATTAGTTAGAGGAGCTGCGCCTGCGTCTGCCTGCCCCCTTGAACTCCCCAAGAGCCAAGAACCAGCGCAGC
AGCGCGCCCCCTCCCCGCCCCACAATAAAACAGCCACACGCTCGCCACCCATGGCCACCACCACCACCACCACCACAGCACAGGCCCACTGCCACCG
CTACCGCTAACCATG
>Kalanchoe_laxiflora Kalax.0029s0028 | scaffold_29:196845..199513 forward
CATAACATAAATCAGATTTTACAAATTACAAACAATAAAAGGTCACCTTAGTCTCTCCAATTTGCAACCCATTTTCGTTGCCCATTGGTTTAGACTT
GACCAACTTGGCACCATGACTCAAAATTGGCTCAAAGAGCTACACACTTGCACCTATAGAAACCTTTGACATCATTTTAAACACACAAACAAATGTT
TTATAACTATCTAAAAATATCCAATTTTAAAATCCCAAAACACAGCCCATCACATACTTGACCAAATTACCCCTCTAGTAGCCAACCTTCAGAAACC
```

```
TGATTCTCTCTCATTAGCTAAATCAGAGAGAGGGGTATTCTTGATATTTTCGCAAACGGTCGTCGCCTTGTGATGACCCTTTACAGAGCCCAAGAGC
TGATCATCACCATTTTTCCTTCATCCCCTCCATCAGCCACCTCCTCAACCTACATAATGGCCTCTCGCTCATCTCCCAATCTTACCATCACCAGTCA
TCACCGCCCACAATG
>Musa_acuminata GSMUA_Achr9G15320_001 | chr9:10031291..10034045 reverse
ATCCTCGTTCTGTGCACTATCCAGTGGACAAAAACCTTGGGCGGCCAACACCCGGATTCGACTGCTACCAAACCTGCCTCTTATCGGGTCGGATCCA
CTGACAATCCATTTCCAAATCACCGTTGGAATGGCTCCCTCCCTTGCATCGAAACTAGAGGTTCCAAATGGCCCAAGAAGATATGGACGTTCTCGGT
TCACCTCGACCGCTGATCCTTGTTGCTATTTTTCCTTCCCCCATCATCAACCACAAGAATAAATTAAGCATGCCTCCATATATAAACATTGAAGGTG
CTAAATAGAATGAGCTTTGCATAGCATCCAACGAGCAATGACAAGGATACGCGGGTGTACGCTTGCTGCACCGCCTTAAACTTTCGAGTGCAGGAGG
CGGAGTTGGCCTTTTGAAACAGCCAAGAGTCGCATCCCAACCCACTATCCAGCTATAACACGGCCCTTTTGTGTGGCCGAACAAGTCCGAACCACCG
AGTCAGTAGTGAATG
>Oryza_sativa LOC_Os01g25484 | Chr1:14446910..14453474 forward
CTCTCGTCGCTTTCTGCGTCTCCCCGTCTCTCCTTCGCCAACAGCCGAGAAGAGGCAGAGAGAGCGCCGCCCCCCGTCCCTCTCTCTCCCTCTCGTC
CTCGCCCCCATCCCTCTCGTCTTTCCCTTGCCGGCAGCAGAGGAGGCGGCAGCGACGGCTTCAGCTGCTCCCACGGGCCGGATCGGGCAGTGGCGGT
GGCGTCGGCGGCTTCCGCTGGCGAATCCGGCGGGTGGATACAAATCAGTGTTCCGATAGGTAAAACCCTGCTCTCAGCATCTGCCCTTTTGAATTCG
CCAAGAGCCAGCATCTGCCCTTTTGAATTCGCCAAGGGCCAGCATCTGCCCATTTGATTTTGAATTCGCCAAGAGCCAGCAACAGCGCCCCCGCGCC
CCCTCCCTCCTCCGCAATAAACAGCCACACGCGCCGCCCCCATGTCCACCCTCATCGCCACAGCGCACCACCACCACCACCACCACCACCACCACCA
CCGTCTCCAGCCATG
>Setaria_viridis Sevir.1G338000 | Chr_01:38187780..38191361 reverse
GGTACATATTGTTACTGTGCAATGTTGATCAGTTCTAAACCGGGTCTGCAGCTTTTGGAGGTAGCTATGTGAACCATATGGCAGTATGAGTACCACA
TAAAAATTCCAGCCCCGTTGTAATGAAAGAGAATGAAAACGCATGAAAAAAAAAACACGAGCAGAAGAGCAGAACGTGGTGGTTTTTTCTAGAATAC
GTGCAGAACATCGTTTGCTAGAGCTCGTACAAAAAACAGGCAAAACACAAGCGAACGGGGCCGGTTTTACCACCAAAAGTACCAGGCGGGACGTGGG
CACGCCGGTAATTCTGCATCCGGGGGCTGCCATGTCGCGCGGCATCTTGCCCTCTTGAACTCCCCAAGGGCCAACAAGCGCCGCGGCGCGCCCCCTC
CCGCCCCACCATAAACAACCAGACGCTCGCCACCCATGGCCACCACCACCACAGCACACGCACACGCAGCAGCAGTGACACCGCCCCGCCACTGCAG
GCACCAGCAGCCATG
>Ananas_comosus Aco002692 | LG06:10717970..10723337 forward
AATTGATAAGCTATTTAGCCTCCACAGCCAAACATTTGGGCCGCTTCTATCCAACTGTAAATTACCAAAATGCCAATGTGAGTTTGTAACCTCACGC
ACGGCACGATGTGACTTCTTTGATCCGAACCGTCCAATCCAATCCAATCAACACAATGTTATATCGAACGGTGGCCAAATGAGGCCGAATGCAGGAT
CGACCACACGTTACCTACGCTGTAATCCGAGCCGTCCAATTAAGATCGGAGGGTTGCAAATCACGCCGAGGGGGCAGATCAGTAATTTCGCGTCATC
AGCAGCAGCCGCAGCAGCCGCAGCATACTCAGCAGAGAAAAAGAAAAAAAAAACGAAAAAGAAAAAGAAAAAGAAAACGTGGCGGTGGGTGATGATG
ACCTTTTGAGGGCTCCAAGAGACCTCCCCCGATCCTCCACCTCCTCCCATAATACAATACTAAATCCCCGCTCCCATGGCCACAACGAATCCCCGAA
CCGAACCCTCCCATG
>Panicum_virgatum Pavir.J03218 | contig04070:10332..14153 reverse
CAAAAAAAAGAGAGCAAAATGCAGCATATGCCATGTCAAGCAGATTGAAGTATGAGTACTTATCCTTCCGTTTAGGGTAAGTTAATTAGTTAAAAAA
ATCCAAAAAAAGTTAATTAGTTAAAAATAAGTTTTAACTATCTTATGATAAATGTTATATTTCTAATGCAAAAGTACATAAAAGTGCCAGCTCCGTT
GTAAGAATAGAATGTAAATGCGTGAGGAAAAAAAAAACACGAGTAGAACATGGTTTGTTGGAGCCTACTAGAGCCGGACCCAGCGGGACGTGGGCACG
TCGGTAATTCTGCACCCGGTGGCTGCCATTTCGCGGCGCGCGGCATCTGCCCTCTTGAACTCCCCAAGGGCCAAGAAGCGGTGCGGCGCGGCCCCTC
CCGCCCCACAATAAACATCCACACGCTCGCCACCCATGGCCACCACCACCGCCGCAGCACAGGCACAGCAGCCGCGGCAGCAGTGCCACCGCCACCG
CCACCAGCAGCCATG
>Zostera_marina Zosma53g00080 | scaffold_53:47938..49896 forward
GCCTACGAGGCCCAAGGACAGTGAGTGAAGGCTATAAATACCTCCCCTGGTCCACTGGGGGGTTGATAAATCCAGAGAGTAAACTGGTACTAAGTTG
TGGTTGTGAGCTGACTCCGGCGAACCCACACACCAACAGACCATGGACCGAACATGACTATTGGGTCTAGGAGGCCCAGAAAAATATCTTATAAATA
CCTCCTTTGGTGCAACAGGAGAGGGGGTCGGAAATCAACGAGAGTAATACTAGAGAGTTGTATGTGTGAGCATCCTAATACGAACGGACACACCATC
AGGTCTGGTATGGACGAAAAGATAAAATGTCGAGCATGATTAATTAGTGTCCCAACCTTTGGCATCTCCGAAGAGATAAAATGATTTAATTGAAAAG
CCAGAAACGTCCCTTCCGTACTCCACCGGTCTTCATCTTAAAAATAATATTTTTGTATATATATAATCAACAACTTCATGTAAGAATCTTCAGATCA
GAGCTTATCAACATG
>Salix-purpurea SapurV1A.0471s0060 | Scaffold0471:68879..71940 forward
ATCTTACAGTAATTAAAAAAATAAGATTATTAGACACACTAGTTATTTTCCTAAATTAAATAAAAAAACTTAATTTTGTTTTCTAAAAACTCATCCT
CAAGCGAGATGGATGTCAAATTCCGCATATAACACCAAAGAAAGAGAGAGAATTAATAGAGCTCTGTTTTCTTTTTACGTGGCATCTAAAATTTAGT
ACCAAAATTATCCACCACCAAACATGGCCAATCAGAAGCATCTTAACATGATCTCCATATTTGTTTGAAAGAATAAATCAACCACCATCTTCTCAAG
TTCCCACCTCTTCTGAACAAACAAATTAGATCGATGACCCTTGGCACCTCCCAAGAGTCCCCTCCTTCAGAAAAAAACACTCTCCTCTCCTCCGAGG
TCCCATATATCATCTACAAAATGCCCTCCCTCCTCCTCCTCACCACCATTACTTCTCTCTACAGCGCTTGCCAAATTCAACCCTTCTCTAGAAACTA
TCTGCCATCACTATG
>Brachypodium_distachyon Bradi3g57990 | Bd3:57352002..57355167 forward
CGCAAACGGGTGCTCTTTTGGATGTAGAGACCTTAACGTGTAAACAAAAGTAATCCGCCAAGGTCAAAACGACCTTCATTTCTCGCATATCTCAAAA
GGCAGTCATCTAGAAATGAATCGTGAGTTCACAACCATGCACGACTCCTACTCTAGGAGTAGTATAGAAGTGTGACGATTGTTGATCTTTAGCACTT
GTGATGTCGACGTCGGGAGTGACAAAGCTGTTTGTCATCTAGAAAGGTCGTGATACGGGTGAAGGCAGGCAGTATCCCAAACCAGAAACGAACGGTA
CGGTGCCCCTGCGGGACGTGGGCACAAGGGTAAATCTCCACCGCCCGGCTGCCATGGCGCGCGGCATCTGCCCTCTTGAACTCCCCAAGAGCCAAAA
ACTCCACCGCACGCCCCTTCCCTCCCAACCATATACAACCACACGCACCGCACCCATGGCCACCACCGCCAAATCACAGCACCACCACCAGCAGCAT
CAGCATCATCCATG
>Amborella_trichopoda evm_27.TU.AmTr_v1.0_scaffold00132.24 | A1198221..1202417 reverse
TCTTCCAAGTCGATGGAAGGCGTCGCGTGTCCAATTCCCCATTTGTGCGTTTTCTTTCCAAATCCATGTTTTTTTTTACTGTTAAACGCTCGAAGTT
TAAAGTATTTTAACAAAGGCCCACTGCCCGACAGAAAGCACGTGAAAACTTGAAGGCAGCGAGGTTTTAATTATACAAAAACAATGAAACTGGGCGT
CTCCTGCGTCACTCAAGGGCCACCGATCGTGAGAAGAGAGAGAGAGAGAGCGGTTCTTCAAAGTATGGAAGAACAAACCTCTGCCTTCTGCCTCACCCA
AGGGCCACCCCCGCCCGCCCTTCCTCTCTCCACAGGTCAAACACTCAGCATACAAATCCCACCACACTACATGCATATATCACCACGTCCAATTTCC
CACCTGCTATCAGGGGAGGAAGAGAGAGAGACCAATTTGAGGCGTTGCCCCCTTGCAGAGAACCAAGGAATAAGAGAGAGAAGTAAAAGGAGACGAG
TAGAGAGAGAACCATG
>Vitis_vinifera GSVIVG01031737001 | chr3:3930217..3943910 forward
GAAGTTAAAAAGAGGCATTTTAGTAAAATCAGAAAAAAAAAAAAAAACAAAAAACGGTAGTCTCTTGATAATAGTCCTACAGGCAAAACGACCTTTAG
CCTCAGCCAAGAGTCTCTTCTTGCTCTCTCCTCCATCTACAAAAATGGACTCCTACCCCCATCCACCGCTACAATCACCCTCCTTCACATCGTCGCC
ATGGCTTCTATCTCTGTTCCTTTCCTCTCTCAGGCACCCACCCACCTTTCAAACTCCACTTCTCTCCGTCTCAAAACCAGGATCTCTGCCACCCCGA
CTCCGACTCCAACTCCAACCACGGTTGCACCGTCGTCCACGGCGGCGGTGGACGCCTCCAGGATGGAGCCCAGGGTGGAGGAGAGAGGGGGTTACTG
GGTTTTGAAGGAGAAGTTCAGGGAAGGTATAAATCCACAGGAGAAGGTGAAGATTGAGAAGGATCCTATGAAGCTCTTCATAGAAGATGGGTTCAAT
GAGCTGGCCAGCATG
>Spirodela_polyrhiza Spipo15G0020200 | pseudo15:2029327..2032054 forward
CTCTATCTATGTTGTTTAATATGAGTATATGTAAATAATCATTAAAAATTGTTTCATCTCCCCTAAAGTAGGGCCATTTGAAATTTAATCTTAATTC
AGTAATCAAAGTTAGGGCAAATGTTCTAACTTAAGTACTCAAAATAGTAAAAGGATCTAATTGTTTAGTAAATTTAATATGGCCACTAATATGTTAT
CAGATGCAAGCCTTGCTAGAATAGAGTTAATCTGATGGAAGCAACAGTTTTAAGCTCTATTATGTTCAAGCCCCTATGAACCCTAGAATTAATGATC
AGCCCCTTGCCTTTTCACTTTTTTAGATAGGCATCATAACCATGACTTCATTCCTCATTCGAATTGACAAGAATATTATTTTTAACATTATATTTAC
TTTTGGATAGTGGTGCTGATTTCCCTTGAGAAACAATACCATAGATAGAGAGAGACACAAAAATCTTGATTTAGCCATGATTAAAAAATTTTATTGT
TCAGGGAGAGTAATG
>Trifolium_pratense Tp57577_TGAC_v2_gene25435 | Tp57577_TGAC_v2_LG4:15486469..15490440 reverse
```

```
ACATTTCCGCGTGAACATAACTCAATCCTAATATATTATTCATTTTAAAATGTAAAATAATAACTATTAAACTACTTAACCATTATTTTTTGGCTGG
CTTAAATGCAGTTTTGACCCCCCAAGTTTAGGAAACTAACAAACTTGGGTGAACATTCTGATCTAACTTACCGATAAACTTTTTAGTACAAGATAAA
AATAGAAAGAAATAAAAAATGTATAGTATGATGAATGTAACCAACCTTTGATTTTGACTTTGACACCACTTTGACACCACATGTGCATTTTACAAAT
AGTTTATAGGACAAGTAAGGATAATACTCTTCAATTACTACAAATTATTAAGATGACCCTTAGCTCCAACCAAAGGACGATGAAGGAGTTAAGAAAG
AACCTCTTGCTTACTCGTAAGGTCCCAACTTCACTCTTCTCTCAATCTCATCCTACAAAAATGTCTCACAAACTCATCCTCAACACAAACTCTCTTT
TCATTCCTTCCAATG
>Phaseolus_vulgaris Phvul.003G292100 | Chr03:52952184..52955717 forward
GTTTGGATTAAGAAAGTTATTCATTCGTGACGACATTCTCGATTACATGCATATAAAAAACTTCAATAAAAAACTTCAATAATATTTATCATTTTGA
ACAATTGTAGAAAAAGTTTACATTATTATTTTTTTAAAAAATTACTTTAATAATTGTAAATTAAGAGACACCAATATATATTTTGATTTTAAATACA
ATTTCAAAATTACTTTATTTTAATAAAAAGTATTATTCTTTGTAATTATAATATGGTACAAGTTGTAATATTTCTAAGATTCTATTTTAAAATGTTT
TAGTTAAACTTTGAAAAAGATGAAGAAAAGAGTGTTAGGAGGAGGAAGTGACCCTTGGTAGCACCGAAGAGTCCGAGGAAGATAAGCTCTTCCTTGG
TTGTAAGGTCCTTGTTCACTCTGCTCCTCTCGTCCTACAAAAATGCACACTCTTCTCTCAAACCGCCATTGCAGCAACCTTCAATCATTCTCAACAC
ACCTTCCTTCCAATG
>Daucus_carota DCAR_029058 | DCARv2_Chr9:1291613..1294066 reverse
TGATTTAATATATTAAGAATTTGAAATATTTGTTGGATAATTTGATTTTATGAATGTTTAAAAATTTAAAATAAAAATGAAAATGATTTACATGTAA
CTTATTTGTGAATAATTTCAGAACTTTTTAAAAATTAAGTCACTAAAACATCTAAATTAATTTTAGTTTCTAATTTATTTTTGACTTTTATACAATT
TCTAACTTAATATATAAATAAATAATTTTTAATTTAAAATCGAAAATAACTTAAAACCCGATCTTCAAAGCCAATCAGGCTCTTTTATCTTTTCACC
AATTTCACCCCTTGACAATGAAATTTAACATTTAACACCTTACACCTCAATAACACACAAAAACACCCTGCCCTTTGGCACTCACCAAAAGTCCACC
ACACTCCCCTTCTCATCTACAAAATGTTCATCACCTCCTCTATCAGCCACCACTCACCACCACTCTCCCCGGCAGCACCACCAAAAACTCCGACACC
CTCCTCCACAAAATG
```

### 3. Class I (C1) and Class 2 (C2) non-symbiotic hemoglobin promoters analysed in Fig S13B

```
>papaya_evm.TU.supercontig_62.96_C1
ACGAAGAAAAAGTTTTTGACTTTTCTCAAGTAAATTAATCTCAAAATATTTGCTTGGAAGACACACCACGCAAAGCCAATCTCAAAAATAATTACCGTTA
GTTCATATATATTTACAAAGAGATGAATTTTCGTAATTTAATTATTAAATCATTATATTTGTTGATAATAATCCTAAATAGACTTGTATATTTTAATATT
GGTCAAATGTATTCTGCACAGATCCCGTCTATGGTCAAACCCACCCATTTTCCCCTCTGTCTCCACATGAAGTGAACACCCAAAACCATTGCCTGCTTGC
CACTTACAGCCCCTGGTTTTTGCATATACGGATTTGGCTGCTGAAAAATCTTGATATGCATAAAAAAAGATAACAACAAAAGTGGCTTCTTGACAACATC
AAGAGCCATCCCCCTTGCTATTTATACACCAGTAATTCTGAAACTGCTCCATATCACCAAACTCCTATTTTACCCAAATTTGTTTTCTTAATCATCTCTA
ATG

>castor_bean_29851.t000102_C1
CACAATACAGCAAAAGTGCCAGCAGATTTACATTTTCTCTGAGCCACAGACCCTCTAGTCTCTATTCTGGCAAAATAAAAATCAAAGCTTTGAAAATGAT
TTCAATGTTATCTTTGTTTTATTTTTGTCTAGGGTTTTCAGGTAAAAAAAAGTAAAGTTTAAACTTCAAACACTGTTTAGGTGACTGCTCACTCATCTCT
TGTCATCATTCATTTGCCAAAAACAACAAAAACGTTGGCAGGGAGAGCAAAAGAAAAATAAAAAGAAAATCAAAATCTATCTGCTAAGATTGCACCAGAA
TCACTAAATAGCACAAGAAGCCCAACCTCAAAAACACAACTTTTTGGCGCTTCTATTACATATTTGATATTACAAAACAAACAAACAAAACTCACAAAAC
CCATCACTATGCAGTCACCATTAACCCAAAAGTTCACTGTTTCTCAATGTTTCTAAGGCAACCCACACACAGGAAACAAGAGAGAGAGAAAAAAACATCA
ATG

>cucumber_Cucsa.109820_C1
TGGTTTCTATACTTCAAAAATAAAAAAGTAATTATTTTAGTTACTTGAATTTTATTTTTAAAGGATCAAAATGGTAAATTTTTAAGATTATGGATCAAAA
TTGAAAGAATAAGAACTAAAATAAACATCTAAAAAGTAGAAGAATCCAATTGAAACAAAGTTGAAAAGTAAAAAAAAAAAAAGAAATCAAATAAACACTTA
AAAATACAATGAAAAAATAAGCAAAAGTCCAAAGTAAGAAAAACAAAAACCAAAACCAAAACCAAGACCAAGGAAGAAGATGCAAAATTGGCAGTTTTTG
AAGACAGCCCCCTGTTTTATTTACATGTCTAGTCCTTGCCTTGTGGACCAATGGTTATAGATATATATATTAAAATCAAAGAAGCAAAAAAGCAATTTGG
CATGCCTTTTCATTCCAATGAAGAGGTAAATTTGAACCCAAATTGAAAGACAAAAGGGATGTTGGGTTAGTGCATAAATAGGATCACCCAATATGGATTT
ACTCAAAACAAAACAGAGACAATATG

>grape_vine_GSVIVG01031900001_C1
GTGCTACACACTTCAAATATGGTGTTGTTGATGAACATTTTGAGGTGACAACATAACTCCTGGTAGTCAATGTGCCCAAAATGGGCTCTAGTGCTTGGCA
GTTTGCATAGAGGTTGCTCATCAGCAGTTTTTAATATTTTTTCAATGTTAACAGGTCACTAAGTATGCATTGCTGGAAACTATAAAGGAAGCAGTTCCAG
ATATGTGGTCACCTGAGATGAAGAGTGCTTGGGCAGAAGCTTATGATCAACTGGTTGCTGCTATCAAGAAAGAAATGAAGCCTCCTCAAACTTCTTAAAA
GAAAAAGAGGTCTCCTTCAAGAGAACGGCTCGAGAAAAGTCTAGATTCAATAGACATTTCTTGCTAATAATTTTGACTTTGGGTCGTTGCGAACATTCCA
TTTGCTCTTACCTACAAACTTCATTGATGTATGGTTGAATTAATTGCTTCATAGTTTCATCCAAATTTGTGAAATAAATTATCCTCAATCATGCAATGGA
ATG

>Arabidopsis_AT2G16060_C1
GTCGAAATACCATTTGACTCTGTTGTAGCGAAACGCGATGTAACTTACGGATACGGTTGAGTTTGAAAGAAATTAAACACCCTCCTCTGTAATGTTCATT
GCCTGACATTCAGTGTGTAAAGTTTCAACCTTTTCTTACCTCTTAGATATGAAAGCCACATGGTGCTACATAACTCTCACAAGTTTTGTTCATAGTTGTG
CAAATAAATTGGATGAGGGAGAGATTCATTTGAGTAGTATAATCTCTGTTCTGCCAATTTGATCAATAGATTCAGTGTATCCCAACTTATACCACTTGGT
GTGGTTGGCCAATACATAAATAAGTAAATTAATGAGATAAAGAGATCTAAAAGACTTTTCTTTAGTGTTTTGACTAATAATTGGTCAAGCCTACCATTAC
AAACTATGTTCCATTACCAGTACAGAGCACATGGTTTCACTTTTTAACCAAGCAACTTTATCTTAAACTTGCTTATCAAGAACTGTCTCTTCGGTAACAC
AAAAGGGTCTTTTATCAAAACCTATATATAGGATACTTTTCATAATTGGAGTAAGATCTACAAAACAGAGAGTTGTATACTTTAAATCATTTAGAGGTTG
TGAAATATTATG

>apple_MDP0000273159_C1
TTATAACTACAGATTAGTAAACATGCAATTAATCATACCAATTAGTAAAAAAGAGTATTTGTGCTCCAGTAAAAAAATGCGACACAGTGCTCTAAATTAA
CACATAATTTGTAACAAAAAAAAAATTTAGACTAAATACATACTATTTTGTATTCTTTAGTGGGGAGTTTTGGCATTTCATTGGCAAAAAAAACAAACAA
TAAAAGAGTGAAGGGTCACTAATAATCAAATAAATAAAAAAGAGAAAAATAATGGAAAATGATTCTATCTCATCAAATCTAATCTAGTGGCATGTGAGAG
AGAGAACACCAGAATCACTTAAAACGGTTAAACCCAAAACCAAAACTCCAGTACGGTTCTATTTTGGCGGTACTGATAAACCTTCCTTTCCTCTCCTCTC
CTCTCCCCAATTGAAACAAACCAGCCAAACCCTTTCAATTTCCAGCCCCTTCAACTCTCTCTCTCCCCTCTCTCTCTCTCTCTCTAGAAAGAAAGAAGAG
ATG

>Poplar_Potri.009G110800_C1
AATGCAACCATTGGATTTCACATCAATATATGCTCCCCAAATTTGAACGGAATTTACAGATAGTTCAAATCCTAAATATTGATATACAAATTATTCTAAT
TTGAAAGTCAACTTTCACTATCAACACCCTTCACATATCTATGAAGCTAGCAAACACTGAAGAACCAGGTTCAGTAGTTTTGTGCACTGGATTTACACGG
AATCATAACCCTGTTTATTAAAATCAGCCTTCAAAACTGAAGAAAAAAGAGGGGAAAAAACTGAAAAATCAAGTGGGAAAGCTGTTGAACGAGTGTCTGT
GGAGAATTTACTTCTAGGATCTCCATTAACAATTGAAGAAAATTTCAAATTGAATAGTTGTAGCCTGTCGAAGATATACTAGCCATTGGTCAATCCCATT
ACAATGCCAATCCAAATTGATAATTAAAACCACGTACTTGACGCCACAAACAACCTGGCCTGGCCCCTGGTTTTTGGACATTTGGTTGCTGGTAACTTTT
GTAAATTATTATTATATTCCCAACAAAATGGCTTTCCAAAGCCGCGAAGAGTTAATCTCATCTTCTCTGAATTTGGCTATATAAATG

>strawberry_gene19672-v1.0-hybrid_C1
GTTGAGAGTTTTGACTCACTATATAGTATATAAGTCTATTATTTGATACAGATATATTTTTTTTTTGGTGGATTTTTATCCAATACCATCAAATCACGTT
AGTTGATCACAATATAATTCATTGAATCCTATACACGTAGAAATATGGTCCAGAAATGTAAGCATTGCTAGAATTTTTATACAGAAGTTGGGTCGATACT
ATGTTCTTGCTTGCGTCACTTGAGTTGAACAGAAATAGCCTAAAACCATGTTCCATCGCTTAGCCACAGGCAGCACCTGGTTTCGGGACATGTGGTACAT
ACTACAGAACCCTGCAGATCTTTATGCAGAAATTATTGAAAGATTGGCTCCTCAAAGTCATGAAGAGGTAAATCTGATCCTCAGAGGATATACTATTTAA
ACTCCACCCAAGAATGAATGGTCTCAACATCCTCAGTTCCATCTTTAAACCAAGTTCAAACTCTGAGCGGTAGCTTCAGTTTGTGTGAGAAAATATAACC
ATG

>cacao_Thecc1EG006315_C1
```

```
TTTTATGATCAAGGCTCGATATTGCTAAATGGCATGACATCTTTCATCTTGTGACGTTCTATTTCAATTATTAAAATGCGTATGGCGTCTCTTTGAAGAA
TCGATCAATGTCATTCATTTCATCATTCTATTTATTTTATTTATCAAAGCATTTGATGTGATATTGAAGTCGATTTCATTTATTAAATTGTTTAACATGA
TATCGCATTTAGAAGCAAAAAACATTTTGCAGGATGAGACCTCCCCTGAATGTAGCCGCTACAGTCGGATTGTCAATTTGTCTTGGTATTTAACCACTCC
GCCTTAAGACGTCCACGACGTGAAAAAGTTCACTGTTTCAGGACCTGCATGATGCAACGAATAAAAATACTACCACTATAATAAAACGAAGTAAACAATG
TTGTCGGCTCTTCCTGGCAATGAAAGGTTAGTCTTTTTGTCCAATTGTAAGCCAAAGTTTGTGTTCCCCTCTTTCGTGAGAGCAAAGACAGGAAACCACC
GGGTAATAGCTACATGCCTATATTGTCTTAACTCTTGCTGCTCTATACTGTCCAAACTCCAAGCATCTTCTTCAATTCTTAGAGATGAATCCCATGACAA
TCACCACTTCAGCTTCGAAAGCTACTACCATTCGTTTCTCTGATG

>maize_GRMZM2G168898_C1
TGGGAGAGGAGAGCGGCGCGCCGGCCCGGATCCGCTCGACGCCGGGGAGGAGGAAGAGGGAAGGAGGCGCGGTGGCCGGCCGGAGGGAAGGAGGCTGGGG
AGGTGGAGGCAAGGCGGAGGACGAAGAGCGAGCGGGTGGAGGGGAGGCGAGCGGCTGCTGTCTGCCGCGGGGGAGGCGGAGTGCCGAGCGGGGAGGTAGG
TAGGCAGGGTAGGAGAGTGGGAAAAATAATATAAAACGTGGGAGTTGGTATAGAGTTGAGATGTAGAGTAAACATGAGCGCAGGGAATTGTTGTATAGAG
TAGAGAATATTGCTGATTAGGACGAAATATTCTTTTTAGAGTGGATTTTTAGAGTTGAATGAGTGTGGATAGCCTTATATAAAACCCCAAAAGTAGTGGA
AACGATCTAAAATCTTCCACGACGTGACCAGCACGTGGAGTCATGGACGGGACGATACCCCCGCCCCTCAAAAACAAAAGGTCTCGTGGGAATCAGCGTC
ATG

>Amborella_evm_27.TU.AmTr_v1.0_scaffold00059.174_C1
TCAACCCTAGGACATCACTTTCATGCAGTGCAATTTTAGTTCCTTTTTTTTTTTCTCTTATTCTCTTTTGAAGTTAGTTGCCAGGTTGAAAAGGCTAAAT
CCTCAATTCTGGGTTCCTAGGATCACTCCTAGAGGGGTTTTACTTTTGCCTGAGGGAGTACATTGATATTATGCTCCTTTAGGAGCACAGAATCTATGGG
ATTTTTTTTTTTTCGTTTTTTTAAAGAGGGAGTATGGCAAATAAGAGGTAACCCATGTTGTTGTCTGCATAGATCTCATAATAGGAAGAATGGAGCTCTT
GAGCAGCAGCCCTTCAACCTCTGCAAAACTCATCTCCATGAGTTAGTCTCTCTCTCTCTCTCTCTCTCTCTTATAGAAAATGACTTTTTTTTTCCCCTTG
TTAATAAATGTAGTGAAGGTGTCATTTTCAGCTTGTTTTTCTGGATTTGCTTTCTGAGATATTGCTGCTGAAAATTTCAGGAAATGGAGCAGTGGCAGAG
ATG

>Setaria_V_Sevir.9G487900_C1
GGATTCTCTCTCTCCCGCTCACTTGGATGGAACGCAGGAATTTCTGAAGAGATTTTACAGAGGTTTCACAGAAATTAATTCGACCTCGTACCGCCGATCG
ATCATAAACACTTCGACGTTCCAATCAAGAACAGGGCCTGCTCACAGCGCCTTTTTTTTTATGAGGAGCTCACAGCACCCTTGATAGTGGACACTGTGAC
TTTAATTGAGGTAGTACGAGAACAAAGTTAGGCCCGCCTGACATGACCAACGAGAAACCGGCCCAAGAAAAAGAAAGGCCGGACGCTTTTGCCAAGTTTG
TGGAATTGGCCTGGCCCTTTGGCGCGTACGAAACGAACGCTTCTTGGAGAGCAAGAAAGAACGGAACGAGACTTTGACTTGAACTCCCGAAACCCGCCGC
TGTTTTCCGCGCGCCCAGCCGCCCACGTTATTGACGGCGCCCGAGCTGCAGCTGCTCTACGCTGCGTTTGCACCCCATGACCCCAACCCGCCCTCCCTTT
TGTCGCAGCGGCACATGGTTTTGAGCCTACAACCCCAGTCCCATCCGGCATCCCCCGACTCCGCCGTGCTGCATGCGTAATCGCGTGCCACTACCAGAGG
ATCAACGCAACATGGCCACCACCGGTGCCCGGTCCTCGCGTCCTGGCTCGTCGGGTGCCCCAAGGGCCACCCAAAGCCTGGGCAGATCGACCTATATAAA
CTCGCGCCCCAGTTCGGCTCGAATCACAGCCGATTTCCTCAGCGATACGTGCATCGATCCACTACTACTGTCCAGTGCTCGCTTACGCACTCAACCAACT
TCACATCAAGCTAGCTAATCGACAAGCAGGGAGAGGCATG

>Setaria_I_Seita.3G177600_C1
GCGGCGGACTCCTCCTCGAGCCAGTTGCGGATGTGGCGCACGTTGCGGCGGAACATGGTGGCCGACTGCTTGACGTCTTTCCGCAGCAGCATCACCGCCG
CGCCCACCCCCACCACCGTCAGGATGAAGTTCGTGAGCGCCATCCCACGCCGCTCTCCCTCGCGCGCAGCTCCCCGCTCGCCTCCCAGCCACCCACGATC
CAAAACCCTAGCGCAGAGCCGGAGACAGAGCAGCGGCGAGGAGGCGGCCAAAAGCAAATGGTACGGCACCAGTCGACCGGAAGAGCTTGCTAGTCACGAT
GAACTTACAGGCCGTGGGCTGAGAAATGGGCTTCTGGGAAAGGCCCAGCCGGCCAGTCACCGTGCTTGGCCCATCTATGTTCAGGCCCATCTACTTGCAC
TCTGTGCTGCTGACATGCGTCCATACCGCCGGCCATGTCGATCTCAAATTCTCGATCTGGTCCCTAGTCGGTGGCTGGCTGCTGCTTGTGGCCAACGGCC
AACCGATCGCAGCGCTCGAGGCAGCCACACGCCTTCTTCAGTCTTCAGACTGGGATAGACCGCTCCACTAGCCTGACGACGGATAGGACACACGAGCTGC
AGCAGCCGCGCCGCAGGCTGCGAGAGCAATG

>Brachypodium_Bradi2g19690_C1
TCCTCCAATAAAAAAAATCTAACGTGCGCCCGTCCGTGTCCACCAACGGTCGTGCCATATCTGAAAGTTCCTTCTGGACACAGGGAAGCGTGCAGCCCAC
CACACAAGTCCGGCCGCCCGGGCAAACGGACAGCAATGCCATGCGGCGTGCCGTCCACACTGCCCCCGCTGTACCTGAAGATATTTCTTTGCCGCCTCTT
CAAGCACCATCCCTCTCTGAACTTACTCATCTCATCTCACTCTCGATCACCCCCGGGCGCCCACCGGATGGAGCAATTAACTAAGAAAGAATCAACCAAT
GTGCAAATTGCCAGTCTCTCTAGTATCTTCTGTCTGATAAAAACGCTGTTCAAGTAATCAAGCCCCCCTTTAGAATCCGGCTAGAAATGAAGATAAAAGC
TGTTGAGATGTGTTAGGATACGGGCAAGGCAAACACGTGGGGCGATCAGTCAAAAGTGCCCACGGAATTATCAGGACATGGACAAGCTTAGTTTTTCTTC
CCAGAAGTCCCTGCCCCGCTCACGCACTCAAGGGCCAGCTACAAAAAATATAATATCACGAGCTCTCTGTTTTCGCTCAAACCAGAGAGCCACCGGGCTC
TGCCTATATAAACGATCAGAAATTCATCAGAACCCACACTCCTAGACGATCACACGCACCAAGAAAACAAAATAAGGGTTAAGAACAGCAGAGAGAAGAT
TTGATG

>rice_LOC_Os03g12510_C1
GACTTTCGTTTGTGGAACGTTCCACACTACTGGACTTTTGGGCTGCACGAACGACCGAGCGAACACGACCGTTTTGCGTCTTGGTGGACAAGAAAGAAAC
GACCCGAGACCGAGAGGAGAGGCGAGAGAGCGTGGAGTATTTGACTTGGTTTCTTAAAGGTTAAAGCCGCCGCTGGTTTTCGCGCGCCCGGCCGCCCACG
TTATTGAGGCGAGACGAGAGCGCACACTGCAGAAGCATGCGGCATGCCTGCCTTGCGCAGGCACAGCTACCCGCTCTCGCAGCTGCATCCGGCACACAGA
TTTTCGGCTCGAATGGATATCCGCCGTAGTCTCCAAGTCGATCTGTTGTTCCTTCCCCCCTGTACGTCTCACCGTAGAGTTGGAGTGGCATCATGGCATG
GCTATCCCGTCTCCATGCTTCTTGGCTCCTCGCGAGCGCGAAAGGCCGCTCCAGCCTGGGACCGATCCACCTATATAAATCGGCCTCGCTCTGCTCCAGT
TCGGACTCAGATCGCAACCAACAGTTGAGTTGAATTGAGCTCGAATTGTACTCGATTCACCACACAGAGGAATCAAATCGAAGCAGCCATG

>carrot_DCAR_028135_C1
CACACACCCGTTTTTAACAATATTTGAATACTCGAAATCTGGTGAGGGAGCGAGGGAGATCTATTAAACTAAGAGTCCGTTGGTTCATGTTTATGATATC
ATCAAATACTGTAAGTTGTACTAACTAACCTGATATATTATCCTCTGGAACCATTGAAGTATACAATTTGTATTTTGTATCATTTACTTGCTGGCATATT
CTATATATAATTACCACCATATGGTACTCTCAAGTTGCCAGCAATCAAGATTTCTAAGATGAAGCAGGATACTTCCAAGTTCCATTTAGTACAGGGTTTG
GAAAATCATCAAATCCAGGGTGGTAAAGCCCACTGTATTGATATGTCAAAGTTTACTACAATGCATGCATTTGTAAGTTGTCTCTTGGGTGACCTAAAAG
ATCATCCTTGTTACTCTATATATTGTCTCACACTGTTTGTATCTTTATAACATTCTAGTATAGTAATATACTTTGAAGCCCCTTCATTTGTACTATGAGT
ATG

>tomato_Solyc07g008240.2_C1
ACAAAAAAAAAACCATGTTCTAGCCTCCACAGGTAGAAGAGCCCATGGTTTTCATTTACATAGTAAAACAAAGGCATATTACATATATTAAACTTTAAAT
TTGGATTTAAATTTTAATTTTGACCTTCAATTTTTATAATGCACAAATAGACATTTTAACTATTCAATTTATAAATAAATAAATACATGAGTCTTTCATG
GTAAAATACACGTTAGACACCACTGTAGGACAAAAAATGTATTATAACATGTGTGTCTATTTGTTCAACTTTATACAAATTTAAGTGCCTATTTGTTCAC
ATCCAAAATTAAAGAGCATAAATGTGATTTGAAGCCAAGTTAAAGAACACATTTATGTATAATGCCGTAAAACAACTAGTGGCTTTCAAAATTTTGATTT
TTTTTTCAACTTGTTCAAAACTAATTTAAATCTTGTTTGGACCACAAGAAATGTCTTTTCGTTAACATGAAAGGCTAAGCCTTTAATTACTATAGCTAGT
AGTATATATACACATACACATAATTGTTAACTTCTATAAGTAACAAATTAATTGAATTGAGTTCCCTTATATAATTTTATTGTAAAGTTTGATCATCAAT
CATCAATCATG
>Araha.4954s0014_C2
AACCCTAATTTTTGTCGATCTGGCCAAGTTTTGAAACGAAATTTGATTGTTTGGATTTTGGTAAAGGAGATTAGGATTCATCGCCGTTACCATGTTGGTT
AACTGGCTCGCCGGGAGACCCAGCAATAAGTTTTATGCTCTCCAAAAGGCAATAGACGAGCTACACAAGGTATTATCATATATATTGACAACTGATATTG
TTAAGTTACTATTTATTTTTCTGTGACATTGAATCTAATTTCAAAAGCATCCAATAATGTCGTTTTCATTCTTCTTTCTGATCACTAATGTGCTCACCTC
AAGACGCTTTACTCTTAACTACTTATTTGTATTAATCTAGCTCATAGTATTATTCTACGTCTATTTAGACTGATCGGACGGACATAATGTCTATCCATAT
TTCTTGGCTACGCGTGTCCACCTTTTAGAGACTATGCCTCTAGTTAATAGATGTTTTACGATAAAATATGTTTATTTAGTCATCAAAAAAATAAAAATTG
AGAGAATCTTCTACAAAAGTAGCTCACAGACCCAACCAAAGGACCATTGAATACCATATTACCATATATATAGATACACAGACATATAAACACAGAAATA
TTCGCGTCTTCGAACTGTGAGAGATAAAGAAAGAGAGAAAGAG

>AL3G21900_C2
GAAGCTGTGAGTTTTTTCTTTTTTTCTTTTTTGTGTTTTGCTAAATCATTTTTTTTATAACTACTTTTTAAGCGATAAATTGAATTTATTATCACTAGTT
TTCTTTCTTCTATACAGCTGATTTTGTGTGAAAGTAGAAAAGAACAAATGATTGAAGCTATGCGGGATGGAGATTTTACTACGCAGAAGACAACGCAAGT
AACAATTTATCTTTGTTTGTTCCTTTTACTCTTAACTTATATGTATTAATCTAGCTCATAGTATTATTCTACGTCTATTTAGACTGATCGGACGGACATA
ATGTGTATCCATATTTCTTGGCTACGCGTGTCCACCTTTTAGAGACTATGCCTCTAGTTAACATATGTTTTACTATAAAATAAACATATTTAGTCATCAA
```

```
AAAAATAAAAATTGAGAGAATCTTCTACAAAAGTAGCTCACAGACCCAACCAAAGGACCATTGAATACCGTATATATAGATACACAGACATATAAACACA
CAAATATTCGCGTCTTCAAATTGTGAGAGATAAAGAAAGAGAGAAAGAG

>AT3G10520_C2
CGAACAGCTCAATATAATAGGAATTGTCAGCTTCTCTTACATGAAGCTGTGAGTTTTTTCTTTTTCTTTTTGTGTCTTACTAAATTATTTTTTTATTAAC
TACTTTTTAAGCGATGAATTGAATTTATTGTCACTATTTTTCTTTCTTCTTAATTATACAGCTGATTTTGTGTGAAAGTAGAAAAGAACAAATGATTGAA
GCTATGCGGGATGGAGATTTTACTACGCAGAAGACAATGCAAGTTTTTATTTATCTTTGTTTGTTCCTTTTAATCTTAACTTATTTGTATCAATCTAACT
CATGTATTATCCTACGTCTATCTAGACTGATCGGACGGACATAATGTGTATCCATATTTCTTGGCTACGCGTGTCCACCTTTTAGAGACTATGCCTTTAG
GTAGTAGATGTTTTACTACAAAATAAACATATTTAGTCAAATAAAATAAAATTTGAGAGAATCTTCTACATAAGTAGCTCACAGACCCAACCAAAGGACC
ATTGAATACCATATATATATAGATACACAGACATATAAACACACAAATATTCGTGTTTTTTTCAAACTGTGAGAGAAAAAGAAAGAGAGAAAGAG

>Bol005842_C2
TAGAGAGAGAGAGCTGGTGGCTGGAGAGAGAGAAACCCTAGTTTCCTTTTTCTCTTTTTCCTTATTAGTTTATGTTTTATACTCTTTCCATATTTTGTAT
TTTCTTTTTTTAGTCTTTCCATTATTCTATGACGGTGTAATTCCTTATATATAAAAAACTCCTTATGTTTATGAATAGATATATAAACATACAGATTTTA
TTGTCTTGTTTTACTACATTTTCTGATTTTAGTTTACAGTTCGGTTCTGTTTGTTGGAGAAGTTAACACCGAGTTATTCAGATTTCATTTTTGTTTGGGT
TGGTTTATTGGACGTACTATATAGAAATGTGTATTGGGCTGGACATCTTTTAGGATCGTATGAGGGCCATATGGCCCATGATAAATAATTAATACTGGGA
AAATATTAGAAAATAGTGAAAGTTTTAACAGAAATAAAAATGCATACAAACAAAGAATTAGTAGGTCTCAAAGAGTGAGAGAGAGAGAAAGAGGAGAGAG

>Brara.A03477_C2
TATGAATGCATGCTTTCGAGTTAAAAAAAAGGCATAACGGTATCATCATGTCGCATGTAAGATGAGCTATATTTTCCTGACTAATTAATGTATTTATTTA
CTTTCTCTTGTAAGAAAAAGTATTTCAAAAATCAAACAAAATTACAAATATATATAACTTCGGCTTTGTTTTGATTCACTAAGAGGCATCTGTTTGGTTT
TAAACAAATTTGACTAAAACGTTTCTTATGAATCAAAGGTTTAGTTTAGTTTACACACTTTGGTTCTGTCAGTTGCCCATATGAATATATGATTATATCA
AATTTTGATCGGTAGAAAAAATAATTCTAAAGAACAACTCTATCCAACCGGTTTACATGGGAAAACAACAGTTCCATTTTTGTTTGGTTTGGTTTATTGG
CGATGAAGTTCTATAAAGATGTGTAGATCCAATGTGTATCGGACCGGACATCCCTTGGGATCTTGTAAGGGCCATCAGGCCCATGATAATAGTTAATATT
CGAGCGAATGTAGTGAAAGTTTTAGCAGCATATTTAATATAGGTCTCAAGTTTGAGAGAGAAAGAGAGAGCTCAGAAGAGAGAG

>Bostr.22252s0009_C2
CTGAATAAAATTGATGGTAATGGGGTTGAAGATCCAAATGTGGATTGGGCCGGACATATTTTTTAGATCATCTAAGGGCCATATGGCCCATGAGAAATCA
TTTTTATTCAGAATGTACATTTCAAATAAGAAATATTGAAAGTCGTAAGTCGTAACATAAGTAAAAACGCATTCGTTATAAGAAAAATATATTATAAATT
TGATCTAATTTTGAAAAATGTAACTATATATAACCTTAAATAGTTAAATGTTACGTTTATACAAAAATGGACGTAAACTTCAATTTATGTAAGAGTGAAT
TTATTCATTTTAATAACACTAATTATTTCCTTTTTTTTTAATCATTTCCCTTAAAAAGATTTTAGAATTGAGTACGTCTCTTTAGACTGGTCGGACGGAC
ATATGTATTCATAGAGATAATGCCTCTAAATAATAGGGTGTTTTACTATAAAATAAACAATTTTAATGTAAAAATAATTGAGAGAATCTTCTACAGTAAG
AGCTCACAGACCCAACCAAAGGACCATTGAATACCATATATTTATATAGATACACATAAATTTAGCCATACAAATATTAGTGTCTTCAAACTGTCAGAGA
TAAAGAAGAGAGAGAAAGA

>Cagra.0834s0026_C2
ACACATGGTGGATTGGGCTGGACATATTTTATCAGGTCATAATGGGCCATATGGCCCAAATATTAGCATATCACGAACGTATCAGAAATATTACGCTAAC
TACTTGTAGGAACTAGTAAATGTTTTAACTTTTAAGATTAGGAATCGAACTAGGGAATTTATTATCATTAATATTTTTCGTACACAGGATTTTGATAAAT
GGCTGAAGACTCGAAGCAATGCGGCGTGGATATTTTACTACGAAGAAGACAACGCAAGTTGCTGTTTAACTTATTTGTATAAATATGGACGTAAACATTT
TAAGTAGGTGGATCTTGGAAGTGGTGTACAACGTCTCTTTAGAGTTTAGAATAGTCGGACGGACATATCACATTCACATATATATGTATCCATATTTCTT
GGCTACGCCTACGCGTGTCCACCACATTTAGGAGATAATGCCTCTAGTTATAAAATAAACAATTTTAATGAAAAGATAACTGAGAGAATCTTCTACAGTA
AGAGCTCACAGGCCCAACCAAAGGACCATTGAATCCCATATATATTGAAGATATACAGGATACATATATATAAACATACAAATATTTACTGTCAGAGATA
AAGAGAGAGAG

>Cucsa.308830_C2
TTCAATTTTTTTAAAAATGCAACATGAAAATATGTTTTAATAAAATGAAAAATGTTCAACACACCATATGTAATGGATTTTGACATTTTACGAGATGTAA
ATAAATATTTTCTTTATTTTTGAAAAAACTCTTTTAAAAATGAAAAAATTGAAACGATGACAAAATAAGCTAATAGTAAAGCCAAAGTGAAGTTAACTAA
GTCATAATTGGTATAGTATTCGTATTTTGTTCGAGATAAAAGATTGGACAATGTTAATATATATATATATATATATATATATATATATATATATATATAT
AATATATANNNNNNNNNNNNNNNNNNNNNTATATATATATATAATATGATCTAAATTAAAAATGGATAAAATTACATAAAATATTAAAATAAAATGGGCAG
TTTAATATTTTAATACTAATTTTTATTTAGAAAGAGAAGAAAGAAAAAAGCATTATAAAATGAGAGTAATGAAAGTGGGTATTAAAGAGAGGTTCCAAAG
CCCCTTATCCTTCAAACATTTAGAGAGACAAAGAATAGGAGAAAGAGAAAGAGAAAGAGAGAGAGAG

>orange1.1g037487m.g_C2
AAAGTTAAACAATAAATATATCATGTTTCTTAGAAAGTCTTAGATGTTATAAGAGTATTGCCTTTTTTTCTTTTTTTTTATTCACTCATTATCATATAAT
TAATTGATAATATTCGTACATATGTAGTATGCGACAAGGAGCAATGCACCTAAAGTACCCAAGACCCTCTCAACATTCCTTACAACAATTTCTCATTAAA
GATAAACACTTTTTTTTTAATTCTTCTTTTAAATGTTAGACCTTAAAACCTGCCCCGCCGGCTGCATAAGCCTTACAACAGACAAACCAACACGATTTAT
ATAAATATTATTTAAATTTAATGGTCGAAGGTCTCCTATAGCTAGATAGAAGTCACATTTTTTAACGACTTGTTCCCTTAGAATATGGGACGTGGCTGCC
ACACGTGCAAATTTTAGGCAGCTCTATCTCTTTGCTTATATATGTTTATCAGACTTCAGAAACAGATCAGAGAGTAAGGAAAACCGGAAGCAAAATTAAG

>Eucgr.G02733_C2
AGTCAAGCTAGGACCAAAGAAAAAAAAATATTACATTTACAGTGATTATATAAAGCTATAATCAAGTTTAGCTAATTGGCTCATAGGATATTTAACAATT
ATCAAGAGTTCGATTTGGGCAACGTTGGAGGTCGGGGAGATACGACACAAGGACGGGCTACCTGAGTGGTGTAAAAATATCAATACGAGCTTGCTCGGCA
TACAAGTCCGATCAAATCTTGACCATGTCAAAATTCACAGTGCCGACTAATCAAAGGATAAAGCTCTCTTCTTGTTGGGCAAGTGTATGACTGGTCATAG
CGGGCATAAACCATCCAACGGGCCGACGACGCTTACGTCTGAGCACGTACTCTCGTCTGACGGTCAAAACTTTCTTCCTTGCTGGGCAAGTACATAGCTC
GTTTGAACGGGCATAAACCGTCGAAAGCACGGAAGCAAGACGTGACCCCTCGGCACGTGGCGGCCTTTATATACGAGCTACTTCACTCGCACTCAGATTC
ACAGAGCGAGCGAGCGAGCGAGAGGGAGAGAGAGAGAGAGAGAG

>Thhalv10021685m.g _C2
CAAATGCTTCATCGTAATTTTTAAGTCAGTAGCAATTGTATTTTTTTATGTATAGTAATTTTTTTTGGACAAGTATATTATATATATAATAATTTTCTTA
TTTCCATTGTTTCTAAGGAAAAAGCCCCACTTATCTTTTTCTTACTAGGTATGTATGTGGTATGTCGACTTGCAAGTAATCTTTTGAGCTAGCAGCCTAG
CACTCGTTCAAATATATCTAGTTATATTAATATTAATATATTACGTACTGTATCTAGTTTTTTTCTCTTACGTACGCGTGTATACCATGAAGAGATAGTG
CTTTTAGTTAATGGATCTTTTACTACATATAAATACATTTTTAATCATTTAATAATTGAGAGAATCTTCATCAAAAAGAACTCACAGTCACAGATGCAGC
CAAAGGACCATTGAGTTACTTACAAGGCCAAAACGTTGACAAAAAACACACACACACATATATAGGGATACACATACAAATATTTGTGTCTCAAATTTAG
AGAGAGAGAAAAAAAAGAGAGAGAGAAAGA

>Kaladp0011s0573_C2
TTTCATCTGATGAACATACCAATTCATAAAGTGATCGTTTATTTTTTCCTTAAGCATTAAAACCTCAACTAACGTTTTATTTTTTGGTTTAAATGTAACC
TGTTTATTCTACCTTATTTCCTTATTATATTCATGATTAATTGATGAATATAATAATCTATCTCATGAATATCATGCTCATATGATCAACATGCATATTC
ATTTAAATATATATATATATATATTATCTAGTTATAAAATTATCGAAATTAATAATTTTGGTGAATATTTGTTAAAAATTGATAAATAAAAAATATTTAA
ATCACTGATTCTTAATGATTGTAATTTTTTTATTTAAAATTTTTCAGAATTAAATTTGATTTTTTCTGTTTTTTTTAATTGTTCATGTGAGCCTAAAAAC
ATTGTATAAGCCGAACAGGAGGCTCGGAACGATGTCGTTTGCAAAGAATTTACGGGATAATACGGAAGGATGAGAGTGTGAGCATGTGACTCTTGGCTGC
TGAGATAGTCTCAACTCTCAAATGTCCCATTAAGTTCAGTTCTTATAAAAATAGCCCGACCTTATCGTCCCTAGAGACAAATGTGGTGCTCGCATTTATG
AGCGATTAAACTGTAGAAGACAACATATGATGTACAAGATCATTCTTAAATTGAGAAGAAGACAAAGGGATCTACTTGTTTTTTGGTGGCTGTTTCTAAT
ATCCCATCTTGAAGGATTTGTGCGCTCCACACTTCTCCCTGGTCTCTAAAACTTATCCCAATACATATATTCTTGCTGTGGGAGCATCCGGGAGTTTGAA
TATATCTGAACAA

>Kalax.0199s0018_C2
TTAATTGATGAATATAATAATCTATCTCATGAATATCATGCTCATATGATCAACATGCATATTCATTTAATAATATATATATATATATATATATATATTA
TCTGGTTATAAAATTATCGAAATTAATAATTTTGGTGAATATTTGTTAAAAATTGATAAATAAAAAATATTTAATTCACTGATTCTTAATACTTGTAATT
TTTTTATTTAAAATTTTTCAGAATTAAATTTGATTTTTTATGTTTTTTTTAATTGTTCATGTGAGCCTAAAAACATTGTATAAGCCGAACAGGAGGCTCG
GAACGATGTCGTTTGCAAAGAATTTACGGGATAATACGGAAGGATGAGAGTGTGAGCATGTGACTCTTGGCTGCTGAGATAGTCTCAACTCTCAAATGTC
CCATTAAGTTCAGTTCTTATAAAAATAGCCCGACCTTATCGTCCCTAGAGACAAATGTGGTGCTCGCATTTATGAGCGATTAAACTATAGAAGACAACAT
```

```
>Manes.07G093600_C2
AAAGCCAAGGGATCAATCGACTCAATTGCGGCTGACTGAATTCAGAAACACCTCATCTTGGGGAACAATCATCAGTTATTGCCACAATGTTTGTCCCTTG
AATCCTTAAAGATAAGGTGATGTCTTGACCCTTGACGATGTTTATTATTATTATTTTATTTTTAAATTTTTTTATTTAAAATTTAAATGTAAACTCAATT
TAATTTAAAGATATATATTTTTAATCCTTAATTCTTGTATAAGATTTTTTTAATTTTTGAATTTCTAAACAATTTTGATATATTAATATTAAAATATTT
TTAATAAAATTTTAAGATTTTAAATTAAGAAAAATAAATATTTTAAAAAATATTTAAATAAATAATTAATAAATTTTTTCATATCACAAAAATTAAATAA
CAATTTTTTTCCATAATTTTTATCATCACATTTCCCAATATATATTTTTTTTTTCGCATCTTTCTATATACGTAGCTCCCTTATCTATCCATATGTATAC
TGAATAGCAACAAACCAGAGACTAAGCTATACCCACAGGATCAGAGAGAAAGAGGGAGAGAGAGAGGGG

>Migut.F01595_C2
TTAATATAACACCTCTGTATTAAACGAATTCGGGACCATAAATTCTTTCAGAAATCATTTGTTTTAAAATGATTAGTGATAATTTGGATTTCCTTTTCTT
CTTTTCAATGTCGCGCACCAAAAATATAATTGCATATCCCCTTTTCTCCACCAAAATCAAAAGCACCACTAGGGGTGAAAGTTGAAACGAATTTTCCAGT
TTCTGATGGTTCCGATCCGAAAAGTTCTGATTTTAGAACAACTATTTTCATTCCGATTTTCAAATTTCCGATATCTGATTTTTTTTTTTTTTTTTTTTTT
TTTTTTTTTTTGGGTTGGAATTGCAATTATATAAAAAAAAATATTTGGGATTTCCGATTTTCCATCTGATTTAATTTCCGAAAAATTCCATTTCACCTGAA
AAAAAAAAAAGAAAAATAATTTTGTAAGATGAGCACACACACCCCCACTTGGCATATATTCCCATAACGATGATGAAATCTCTTACAGGATAAAATAAAT
AATAAATAAAAAATCTAAAACATGGCCTATCTTTATCGTCTAATATTATTGCTACCATTTTGTGTATAAAACCACCCACACACACACACACACACACACA
CACTAAAGATGGGTAGTTTCAGTGAGAAGCAAGAAGCATTGGTTAAAGAATC

>Solyc03g071690.2 _C2
TAGGTATCCCTTTCACAAAGACAGCTTTGTGCTATGTTCCACCTAATAAGTGTGGTAATGATATTGCGAGGATCATCACATACAACATATAAATAAGCTA
TGTGTAATAATTTTAGAAATAATAGTTTTGCCTTTGATTGTTCTCCGGCCTCTTAATACTTTTACATAAATAAAAATTAACATCGTATTATTTGAATGTT
GGAAGTCAAATAATTTTTTTATAGACCTTCACTTTATATATATTTGTAATTATATAAATTTTATTTTAAAAAAATTAAAAAATTCAATATTCAACTTCAC
AATCAAAATTAAAAAATTAACTCCCCCACCCCCACCTCACCCAATGTGAATGGAGACAAAGAAAATGATATAATTGTATATCTCTATCTTTGACAAATGT
TCAAAGAATCATATCTTATGAGCAAAGGTGTGACGTAGTGGTTCTTATATTTGTTAATTGTTTGTATATTTCCTTTTCTTTTTGCCTTTTTTTTTTTTAT
AAAAAGATAACTAAACAGTTGTACAAGTATAAATAGTTAAGGATGCTTTACTAAAGTATAGAAATAAGAGAATAATCATACAAG

>PGSC0003DMG400016915_C2
ATATAAATAAACAATGTACATATAATTTTTAGAAATGATAGTTTTGTCCGATATCGTCCTCCTCTTCACGCTTCTACATAAATAGAAATTAACATCCTCC
AACATGGTATTGTTTGAATGTCGGAAGTCAATTTTTAAAAATATATTTTAACAATATTTAACTTATTTTTAAATATATAAATTTTATTTTAGAAAATTAA
AAAATTCAATATCCAACTTCACAATTAAAATAAAAAAATTAACTCTCCACCTCACCCCAATGTGAATAGAGAGTAGAGACAAAGAAAATGATATAATTAT
ATGTATCTATCTTGGACAAATGTTCAAAGAATCATATCTTATGAGCAAAGGTGTGACGTAGTGGTTCTTATTTTTGTTATATAATTGTTTGTACATTTCC
TTTTTCTTTTTTTTGCCTATTTTTAAAGATAACTAAACAGTTGTACAAGTATAAATAGTTAAGGATGCTTTGGTAAAGGATAGAAATAAGAGAATAGAAG
```

## 4. Leghemoglobin promoters of medicago, soybean and common bean, analyzed in Fig S13B.

```
>Phvul007G141900
TATTTATAACTAATAAAATATTATGTAATAAAAAAAATACAAAATTTTACTTTATAAATAGAAGAAGAGAAAATATTAGTATAAAAAATAAATATAAAAT
GCTAAAAGTTAGAAACTACTATTTTAAACACTACAAGCCATTGAAAAAAAAAAACTTATCACTTCAAATATGTTTAATGAGTATATTTTGAGACTTAAAA
AGGATAGTAAAAATCAGATTTGTTTTATTAAGATTAAATTTAAGAACGCAGTATTTTACAGATATATTTTAGGTTAATAGTTATGTTTACATGAAACTGA
AAAGACATTAAATATTTTAATGTTTAGAATACTTTGATTAATTAAAAGATTATTTGATATTAAGATTTTAAAAAGATAATTGTCTCTTCACCATGTGTAT
TAGCAGACTCCATTAGCCAAGAGACGTACGTAACACTTAGATTTTTTTCTGACTCCAATTAGGTTTCTATATAAAGAAGTATAGGATGTGAAGTTGTTGC
ATAACTTGAATTGAAGAGATAGAAATAAGAGAAAAGTAGAAAAGAAAA

>Phvul.007G141900
TATTTATAACTAATAAAATATTATGTAATAAAAAAAATACAAAATTTTACTTTATAAATAGAAGAAGAGAAAATATTAGTATAAAAAATAAATATAAAAT
GCTAAAAGTTAGAAACTACTATTTTAAACACTACAAGCCATTGAAAAAAAAAAACTTATCACTTCAAATATGTTTAATGAGTATATTTTGAGACTTAAAA
AGGATAGTAAAAATCAGATTTGTTTTATTAAGATTAAATTTAAGAACGCAGTATTTTACAGATATATTTTAGGTTAATAGTTATGTTTACATGAAACTGA
AAAGACATTAAATATTTTAATGTTTAGAATACTTTGATTAATTAAAAGATTATTTGATATTAAGATTTTAAAAAGATAATTGTCTCTTCACCATGTGTAT
TAGCAGACTCCATTAGCCAAGAGACGTACGTAACACTTAGATTTTTTTCTGACTCCAATTAGGTTTCTATATAAAGAAGTATAGGATGTGAAGTTGTTGC
ATAACTTGAATTGAAGAGATAGAAATAAGAGAAAAGTAGAAAAGAAAA

>Phvul007G142600
CTTATCATTAAGAATATTGAAAGTAATAAAAAATAAAATAACAAAGTTTTACTTCATAAATGGAAGAAGAGAAACTATAAATACAAGAAAAATTATCTCA
TATAGAAAATATTAAAAGATAAATTCAATAATAGTATAAAAAATAAATATAAAAAGCTAAAAGTTTAGAAACTACTGTTTTAAAAGCCATTGAAAAAAAA
TCAGATTTGTTTTATTAAGATTAAATTTAAGAACGCAGTATTTTACAGATATATTTTAGGTTAATAGTTATGTTTACATGAAACTGAAAAGACATTAAAT
ATTTTAATGTTTAGAATACTTTGATTAATTAAAAGATTATTTGATATTAAGATTTTAAAAAGATAATTGTCTCTTCACCATGTGTATTAGCAGACTCCAT
TAGCCAAGAGACGTACGTAACACTTAGATTTTTTTCTGACTCCAATTAGGTTTCTATATAAAGAAGTATAGGATGTGAAGTTGTTGCATAACTTGAATTG
AAGAGATAGAAATAAGAGAAAAGTAGAAAAGAAAA

>Phvul007G142500
TCGTAATAAGGTATTTTAAATTAATAAAATTATTGAATTATTGGTCTAAATAAGTCTATTAATAATTATATAATGACAATATTTGATGTTTGATTCTAGG
TGCTCACATGCAGATTTAGAGATGTAATATTTCATTTCTAATTTTTTTTATTCTATATTTTGAAAACATGTGGTTAAAAACAATGTATAAAATATATATT
TATATTAGTTTTGATTTTTACATATACTGGACAAGTATAGATTATTTTTGGGTGTTACATTTGTTGTATAATTTGTAATACAAAAAATGCAGAATTTTTT
TTCTTTTAATTTCTTAATTAATGTCCTTAGGGCATGTGTTAACCACACTTGTAATTAAATTACGAATTTACTTTGATTAACTGAATTCTTTTATTATTTG
AAAAGATTATTGGCTTTTCGTCATGCCAATTGACACCCTCAACTCAACAAGCCAATTACAAGAGACATATGATTTAGTTTTTCTGACTCCAATTAGGTTT
CTATATAAAGAAGTAGAGGATGTGAAGTTGTTGCATAACTTGAATTGAAGAGATAGAAATAAGAGAAAAGTAGAAAAGAAAA

>Phvul007G142200
AAATAAGTCTATTAATAATTATATAATGACAATATTTGATGTTTGATTCTAGGTGCTCACATGCAGATTTAGAGATGTAATATTTCATTTCTAATTTCTT
TTATTCTATATTTTGAAAACATGTGGTTAAAAACAATGTATAAAATATATATTTATATTAGTTTTGATTTTTACATATACTGGACAAGTATAGATTATTT
TTGGGTGTTGCATTTGTTGTATAATTTGTAATACAAAAAATGCAGAATCTTTTTTTCTTTTAATTTCTTAATTAATGTCCTTAGGGCATGTGTTAACCAC
ACTTGTAATTAAATTACGAATTTACTTTGATTAACTGAATTCTTTTATTATTTGAAAAGATTATTGGCTTTTCGTCATGCCAATTGACACCCTCAACTCA
ACAAGCCAATTACAAGAGACATATGATTTAGTTTTTCTGACTCCAATTAGGTTTCTATATAAAGAAGTAGAGGATGTGAAGTTGTTGCATAACTTGAATT
GAAGAGATAGAAATAAGAGAAAAGTAGAAAAGAAAA

>Phvul011G048700
GTCTTCTCTGTTATATATTCACATTTACCCTATAATAAAATGATACAAATCACATTGATTGTTTCATTAAAAGAATCAGTACAAGAAAAGACAAGTGTAG
TAAGCAGTGATTATAAGATGATCATAACAGGAAAAAGATAAATATATAGTAATCTGATCTCATCATGGCTTACGTAGGGTGTGCTCAACCAAACTAACTC
AAGTTCAAATTCAACTACACCACCATTTTGACTTCTCCAAACCAAGAAGCAGCCCCTGGTTTTTGAGCTCCTAGTATGCCACATTCCTACTGTTTCCCTT
CTGATATACAACAAATGGTGGCATGAACCAAATCTATTCAGATATAATCCAAAAAAAAAGTGGTTTTTTGGATATATTTTACTATTCTAGAAAGCAAATTA
ATGGACTAAGAAAAACCTGAGTGGTATGACTCACCGTTGCAGTGAAAGGCCAACTTCATTTCGATGAATCCCACTATAAAAACTTCACCAAACCATGCCA
CAACATCAACCATTTCCTCATCCAACTTCCACTGCAATCACAATCCCAAAACAAGTACTCAAC

>Phvul011G048600
GTTATTATTTTCAAAAATATGTTACTATCACTGATTAAGTTACATAATTTTCTTAATCTTAAGTTAAGTTACATAACTTCTCATACACATGGTTCATCTC
AGGTTTAATTACTTTCTCCACGTTTTATATAATATTAGTTTTTTATCTATCATTTTTAGAATTTCAAGATGATATAAATTTTTTTATTATACTTTTATTT
CCATTTTTATTATATTAGATAACTTATTAAAAATAAAATAATTAATTTTTTTTTTCTTAATTTAGGTAATCCAATACCTTAAACACCTAATATAATAATTA
CATAGTTTTTCCAAATCCTAAAAAAATAGTAATTTTATATTCCTAAGATAATGGTGATTATTTTAATATTTTAATTGATATTTTTATTTTCTTGGAAATA
TCAGAAAGAACGTGGGAGTGAAACAGGAATGCCTACCCATTGCAATGAAGGGGCAACCTCATCTTGAAGTAATCTCAGTATAAAAGCAAGCCACAACACC
AACCACTTCTTTCATATCAGTCACAAAC
```

```
>Glyma11G121700
CATTCTATGGACGTCACTAATTTGGCATTTGGCCAATTCATGAATCAAGTAACGGTGGGGACCGGAAAGGTGTAAATGAAATCAAACGTACTTATTTTAA
GATCCTAATTAATTACTTAATTAAACTTTATAATAAAACCAGAGAGAATATAAAGTTTCATTGACATACATTTTTTATACCGTCAATTAATAAAATATCA
TTATTATTATAATTTTAAAAATAATTATCATAAAATTTAACAAACTTGTTATATAATCAGAAAAAAAAATTGTAATATGAGAGTTTAAAATAATTAGAAG
ACATTTTCAGTGCATATATTTATCTTCAGTCAAAATAAAAAGTGCATATATTTATTTATTGTTATTATATACGCATTTTTATTTCATTTACCTCATTATA
TACAATTTAATTGGTGGCAGATAATGTGATGCTATTTTCCATCTTATTGATATGGTGATTATATTAATATTTTAATTGATATTGTTTGTTTCTTGAAAGC
TCAATCAAAGAACGTGGCAACCAAAAGCATGGGAATGCCTTCCCATTACATTGAAGGGGCAACCTCATCTTGAAGGGGAAATATCTCACTACAAAACCAA
GCTACAACATCAACCACTGCCATCATATGACAATCATATATACCAATACTAGCTAGTAGCTAGAGAGTAGAGAAC

>Glyma11G121800
AGGTGTTCAAGTAAAATTACGTTGTCTGATTATAACAATTACAGTCTACTGCATTTGGCTGTCTAGGAACAAGCTGATTTTTGAAGATTATCAAAAATAG
AAGTTATTAGCAAGATTAAGTTTCTTATATATAGACAAGCGCATCTATTGCATTTGTTTTAGCATCTTGATATAGTTATTCTTGTATTAGGGAGACTTTT
GATTTTCTTTAACTGCGGGGTATGCCCCATTTATTATTGTATCATTTTGGGTTTAATATAATTTACATTTTTTCCCAAAAAATAATTTGATTTGGTCATG
GCTTACCCAAGGAGTGGTCAACTAAACTAATTGAAATTCAATGGCACGGTCATGTTGACTGGCCAAAACCATTTTTTTCTCGCCGCAAGCAGCCCCTGGT
TTTGGGGCACGAGCACCTAGTATGGCACATTGCAAGTGTACCTTTGTATTTGGATATATAACTAAGGTGGCATGACTCAAATTGATTCAGATATAATCGA
AAAAAATTACTGTTTGGATATATTCCACTATTTAGAAAGCAAAATGGACTACGAAAACTTGAGTAACAAGGTAAGCCACACAAATGGGAATGACTCCCCA
TTACAATGAAGGGCCAACTTCATTTTCAATGAATCCCACTATAAAAACTTTAGCAATGCAAAAGCTAAAACATCAACCATTTCCTCATCCACTTTCACTG
GAATCACAATCCTGAAACAAAAACATCTTAGCATTTAACATACTACTAGACAAC

>Glyma10G198900
TTTTCCCATCATATTATATATGTGAAAATTAATCTTCGGTGTGTGTAACTTATAACTTTTACGTCTTAATTAACCACATTTTTTTTATATCTACAAGGTT
TTACTCCAAGATCTTACTTAAAGAATTTAAATTTAATTTTACTCAAATCAATGATATATATTTTGGTAACTTTTTTTCTTTTACTTATAATTTTGTTTAC
GTTAAAACTGAAAAAAGAATACATTAAAAAATTAAAAATTCACCGAACAACTTAAATTATTTATTTACTTTGACTAAGTGAAAAATTACTTGATTAAGTT
TTTGAAAAGGTCGTTGTGTCTTCATAATGCCGATTGATACGCTCCACATTCAATAAGCCAAGAGAGACATATTCAATAACAATCGCAACAAATTTTTTTT
TTCAGTCTCCAAACCATCTATATAAACAAGTATTGGATGTGAACTTATAACTGGATTGAAAATAGAAATTAAATAACAGAAAATTACAAAAGATCGAAAT
>Glyma10G199000
ATTTATTTTTACAAAGGAGACTTCAGAAAAGTAATTACATAAAGATAGTGAACATCATTTTATTTATTATAATAAACTTTAAAATCAAACTTTTTTATAT
TTTTTGTTACCCTTTTCATTATTGGGTGAAATCTCATAGTGAAGCCATTAAATAATTTGGGCTCAAGTTTTATTAGTAAAGTCTGCATGAAATTTAACTT
AACAATAGAGAGAGTTTTCGAAAGGGAGCGAATGTTAAAAAGTGTGATATTATATTTTATTTCGATTAATAATTATGTTTACATGAAAACATACAAAAAA
AATACTTTTAAATTCAGAATAATACTTAAAATATTTATTTGCTTAATTGATTAACTGAAAATTATTTGATTAGGATTTTGAAAAGATCATTGGCTCTTCG
TCATGCCGATTGACACCCTCCACAAGCCAAGAGAAACTTAAGTTGTAAACTTTCTCACTCCAAGCCTTCTATATAAACATGTATTGGATGTGAAGTTATT
GCATAACTTGCATTGAACAATAGAAAATAACAAAAAAAAGTAAAAAAGTAGAAAAGAAAT

>Glyma10G199100
TATAAATAAATCTCTCTAGTGTCTATTTATTATTTTATCTGGTGATATATACCTTCTCGTATACTGTTATTTTTTCAATCTTGTAGATTTACTTCTTTTA
TTTTTATAAAAAAGACTTTATTTTTTTAAAAAAAATAAAGTGAATTTTGAAAACATGCTTTTTGACAATTTTCTGTTTCCTTTTTCATCATTGGGTTAAA
TCTCATAGTGCCTCTATTCAATAATTTGGGCTCAATTTAATTAGTAGAGTCTACATAAAATTTACCTTAATAGTAGAGAATAGAGAGTCTTGGAAAGCTG
GTTTTTCTCGAGGAAGAAAGGAAATGTTAAAAACTGTGATATTTTTTTTTTGGATTAATAGTTATGTTTATATGAAAACTGAAAATAAATAAACTAACCA
TATTAAATTTAGAACAACACTTCAATTATTTTTTTTAATTTGATTAATTAAAAAATTATTTGATTAAATTTTTTAAAAGATCGTTGTTTCTTCTTCATCAT
GCTGATTGACACCCTCCACAAGCCAAGAGAAACACATAAGCTTTGGTTTTCTCACTCTCCAAGCCCTCTATATAAACAAATATTGGAGTGAAGTTGTTGC
ATAACTTGCATCGAACAATTAATAGAAATAACAGAAAATTAAAAAAGAAAT

>Glyma10G198800
CTAAATCGTATGTATCTTGTCTTAGAGCCATTTTTGTTTAAATTGGATAAGATCACACTATAAAGTTCTTCCTCCGAGTTTGATATAAAAAAAATTGTTT
CCCTTTTGATTATTGGATAAAATCTCGTAGTGACATTATATTAAAAAAATTAGGGCTCAATTTTTATTAGTATAGTTTGCATAAATTTTAACTTAAAAAT
AGAGAAAATCTGGAAAAGGGACTGTTAAAAAGTGTGATATTAGAAATTTGTCGGATATATTAATATTTTATTTTATATGGAAACTAAAAAAATATATATT
AAAATTTTAAATTCAGAATAATACTTAAATTATTTATTTACTGAAAATGAGTTGATTTAAGTTTTTGAAAAGATGATTGTCTCTTCACCATACCAATTGA
TCACCCTCCTCCAACAAGCCAAGAGAGACATAAGTTTTATTAGTTATTCTGATCACTCTTCAAGCCTTCTATATAAATAAGTATTGGATGTGAAGTTGTT
GCATAACTTGCATTGAACAATTAATAGAAATAACAGAAAAGTAGAAAAGAAA

>Glyma20G191200
ATCTTTCCAGTACATTTATTTCTTTTATTTTTACAAAGGAAACTTCACGAAAGTAATTACAAAAAAGATAGTGAACATCATTTTTTTAGTTAAGATGAAT
TTTAAAATCACACTTTTTTATATTTTTTTGTTACCCTTTTCATTATTGGGTGAAATCTCATAGTGAAACTATTAAATAGTTTGGGCTCAAGTTTTATTAG
TAAAGTCTGCATGAAATTTAACTTAATAATAGAGAGAGTTTTGGAAAGGTAACGAATGTTAGAAAGTGTGATATTATTATAGTTTTATTTAGATTAATAA
TTATGTTTACATGAAAATTGACAATTTATTTTTAAAATTCAGAGTAATACTTAAATTACTTATTTACTTTAAGATTTTGAAAAGATCATTTGGCTCTTCA
TCATGCCGATTGACACCCTCCACAAGCCAAGAGAAACTTAAGTTGTAATTTTTCTAACTCCAAGCCTTCTATATAAACACGTATTGGATGTGAAGTTGTT
GCATAACTTGCATTGAACAATAGAAATAACAACAAAGAAAATAAGTGAAAAAAGAAAT

>Medtr1g090820
ATGGTGCAATGTTCAACGACTTCATCCAAGGATAAAGACATACGTAATGCATATGTATGTATGTATATATTAAAAAATATGATTTATTTTAGATATTTCT
TAAAATGAATGTTAACACAACCAATTTTGATTATATTTTTAAAAATTATTAATTAAATAATTTATATGGTCTCATTGAACAACTCGTGCGTACTCTTTAA
AAATAGGTTCTAAAAAGATTTTACTAAATAAAAAAGTATTTTAAATAATATTTGCTATTTGCTAGTAGTTTTTAACAGATTAATAATTACATTAGATATA
TGAATATTGGAACTTCCATAAAAAAATTTAAAATGGAAGAGACCATTAAATGTTTATATATTTTTTAAAATTATATTTGCTTTAATTAATTAAAATTTATG
TGATGGTGTTTTTGAAAAGACAATTGTCTCTTAATATTGCCAATGGCCACCTCTAAATTAAAGACCAATAGATTACCTAAACCAATAATTAAAAAAAATA
AAACAAATTACCACCTCTTTATTATTCATTAACCAAGCCTTCTATATAAACAGTTGATGGTTTGTAAAATTGTAGTATTCAAAAACTCCAATTCAAATTA
AAAGAAGGAGAAAAAAGAGAAA

>Medtr5g081030
GTGGTTCACGAAGATGTAGAAACCCTTATTTTTCACTGCATTCAGGCGGCGATGCAAATTCAAAAATAGTAATTGTTTTGTCAATATTACACTATATTAT
TTATTTCAAAGAGAGAAATAGATAGAATGAGATAACAATTTAACATTTGTTATTTATAAGAAAAAGATAAAAAAAATAATAATAATAAAAAATAAAAACA
ATTAAAAGTAATAAAATTTATTGGTTTCCTTGATGTAAGTATAAAGCCCCAAAATGAAAACTAAAAAAAAAATAGAAGGAGCATTACCTTTATTTGATTT
GTTTTTTTGGTGAGAGGTAGGGGGCAAAGCCCCCGCCATATTTGATTAAATTTTGTTAGTTTTAAATGTTAAAAAAGATTCTGGTCTCTTCATAATTTCA
ATGGTCATTTCCACAAGCCAATAGATTCTTTGAGAAAACAATTACAATGAACTAACTGCATCTCTTTACACCTATAAGTTGTCTATATAAACAAGTATTG
AATGCAAAATTATTACATTGTCAATTCAAATATAACAAAACTAAACTAAAAAGAATAAGAGAAAT

>Medtr0026s0210
CTACCACCACAATTTAATAAACAAATATTACACTACCTTATAACACATAGGCTTAATTACTCTTTTGATCCCTTAACTTATTTTTTGGTTTCGTTTTGGT
TCCTTAACTATTAAAAGTTTTGTTTCGGTCCCATAACTTATTTTTTGGTTTTATTTTGGTCCCTTAACTATTAAAATTTTCGTTTTGGTCCTTTAACTTA
CTTTTTGGTTTCGTTTTGGTCCCTTAACTTATTTTTTAGTTTAATTTTGGTCCCTTAACTCTAATACATTCATCCTAACATAAAGGACCAAAGTGGTTGA
CAGAGGATGAGTTAAGGGACCAAAACAAAACCAAAAAGTAAGTTAAAGGACCAAAACGAAACTTTTAATAGTTAAGGGACCAAAACGAAACCAAAAAGTA
AGTTAAGGGACCAAAACGAAACTTTTAATAGTTAAGGTACCAAAACGAAACCAAAAAATAAGTTAAGGGACCAAAAGAGTAATTAAGCCTAACACATAAC
ATG

>Medtr4g068870
AATTCAGGTCCCTAAAAGACTACCTATCATTCAGTTAGTTTACATATTTGAGAGAAAGCAAAAATGAGAAGAGAAAATTTGAGTGAAGATAAGTTTTTAA
AAATTTTACTTTTTTCTACGTGAAGAAATCAATGTTAAAGTATGATACCATTATTCTATTCTTTCTCTTCAATCCAATTGTGTAATGTAACCTACTGATT
ATTAGTATTACACCAAGTAATTTAATTTATCGTTATATTTATATATTGTGAATATTAAAAGTAGTTTTTTAGAAAGAAAACAAATTAATTGTTTTTTTTA
AAAAAAAAAAAAAAAAACTTTTTCTTGGAAAAATCATAGCCTGACAAATTTCCAATTAAATAATTCGTCCGGATTGACACGAAAGGGCAAACATCATCTT
AAATTTCAAACAATGAATCTCACTATAAAAACCGTAACAAAGCAAACTACAACATCAACTATTTCATCATCCTTTGCCACTACCACCACAATTTAATAAA
TAAATATTATACTCTCTTATAACACATAAC

>Medtr4g068860
```

```
AATTTATAGAACATAAAATTTGTAATCTACTTATCTGAAGATTTTAATTTCATACCAACGGTACGGTGTGTATAATAATTAAATTCTTTATCTTCTATTG
GAAATAAGAAACTGTGTTTTGCCCCTAAAAAAAATAAAAAAATAAAAAACTTTGTTTTGCTCTAAACACACGTATTGTTATGAAACATGACATCCATAAA
TAAAGTCAAGAGAAAGTTAAAAAACTCGTGTATTGATAGAAAGATAACCTTCTAATGAGGTGATTACATGTGTATATCTTCCACACTTATCATCATATTA
TTTATTTAAACAACAAAACATAAAGCACTAGCAATAAATGAATCAATTATAATACCACTCTTCAATAAAAAATACAAATTATACTACTAGTCAAAATAAT
TTGGTCATGGCTTACGCATAGTGGTAAACTAGTTGGAAGAAAAAGTTCAACCTTTGCTTACGTGATTAGCAGAAACCTTTATTTTATTTTATTTTGGTTA
AATGGTAGAAACCATTTTGTGTTCACCACGTATAGCACCTAGTTTTTAGGCCCCCGGTATAGTACTTTACTATTAGCAAAAACATACATGTCAACATGGC
TGCCAAATTATGGAATGAATGAAAGTATGAATGACTCCCCATTGGTATGAAGGGTCAATTTCATCTTCAATTAATCTCACTATAAAACCTTTATCCTACA
AAATTGTAAACATCATCAACCATTTTCTCACTCCCATCACAATAATCCACAAATCACATACCTTAATTAGCACGTAACCCTTACCCTTCAGAAT

>Medtr1g011540
GAGAGAGAGAGGATGAGGGTTCTTTGGAGGTGTTTACCTAATTTGGGATCTTCAAATTTTCCACTTATGCTCTCCACGCTCTTAGTTCACAAAAATAAAA
CTATATAACTCCACAAAAATATAATTAATAAAATTATTTTAGTAAATGAAACTCATTTTGTTATTAAAATCAAAATAACTCACATTATTAAAAATTGGGC
ACAACTCAAGTGAAATATTCACCATGAGACGAGGAGAATAATTATTTTTATTAGATTAATTAATAATTACATTGATTAAATCAAGTTAAACAATACTTTC
ATTAATGATTAGTTTTTTATTTAAAATATTTATTAACTTCAATGAAAACTTGTTAGATTAAGTTTTTTAAATGATTATTGTCTCTTTAATAATGTCAACA
GCCATTTCCACAAGCCAATAGATTCTTTAAAAAAAAAATCATAAAAAATCAATTGCATCTCTTTTCACCTCCAAGTTTTCTATATAAACAAGTATTGAAT
GTACAATTAATACATAGCAAATACAAATAGCAAAAATAAAAACAAAAAGAAGAAGAAAAGAGAT

>Medtr1g049330
CAAAGCAGCACAACTGTCGTAAACCCCATCCCTAGGTCGACGCAAATCAACGTGTCGCAAACGCGAAACCAAGATTAGACCATAGAACAGTACCGCAAAA
ACGCGGCGGCAACAAGAAGCCACCAAAAGAACTCGATTGCACCACCAAAAGCCATGGTCTACATCGTGGCGGCCAGAATTGAGGTGCGGCAAAGGGTTGC
GGCTGTGGAGGTTTGTTAGAGAGAGGGGAGCGGAAACGAGAAGAAAAGAAAGGCGTCTATGACTATAGCTTTATTAGGAAAAGGTTAAAAAATCAAATTA
TGAGCGGGTAAAAAATTAAAAGTCGATAAAATTTAAATTAGATGTTTACTTCTTTTTGTTTTTAGAATTATTCATTAGTTTAATTAATTAATTAATTGAT
TATGATTTTGAAAAGATGATTGTCTCTTAATAATACCAATAGCCATCTCATAATTAAAGGCCAATAAATACTCTAATTGAATAACCACGAGAAAGCAACT
CTCTTCTCTTAATTTTGGTTCACCATGCCTTCTATATAAACTTTTAATGGATGTAAATTTGTACCATTGAAAAATTCTAAATTAAATTAAAAAAAAAAAA
AGTTGAAAAAAAGGAGAGAACT

>Medtr1g090810
CAACTATTAACATAAAATTCAATATTAATATTTTATTTATAAATGCTTTACCACAAACAGATAGAATTTAACATATGAGAATAAGACATGTTTTTGAGTT
GGATTGAGTCTTTGCTAGCCTCTAAATAATTAAATAATGTGCTCTTAGCGTACTTTTTTAAGTTTTGATAGCTACCATTAGTTACATGAAAAACTGAAAA
GAATCATATTCAAGTTTCACAACCACAGTCTAACGCAAGTGATTAATGTGCGGTGTCTGTTGTAAATCTTAGAGTATTGTGTTTGATTCTTGATTAATAT
TTAAAGTTTTAAATGTTTTGAATTATTTAATGAGTTTAATCTTTTAAATAAATTATTTAAGTATATATTATCTCTTAATAAAGGTTGCCTCCTCTCAAAG
GCCAAGAAATACTTCAAATAGCAATCACAATAAAACCAACATTAGCTCTTCGTTGTCTCAAAGCCTTTTTTATGAACAAGTGTTAGATGTTGCGTAAATT
GTTGCATTATAGAATACAAAAGAAAAGCAAAAAACACAAACAC

>Medtr5g041610
GTAATTTTTATGTCTTTCTTCTAGTGTTGTATTCTACATTTTTAATATACCTATGAAATTTAACAAAACTGCACAACCAGTTATGATATTGTTTTATTCA
ACATTATTGTTGTAAGAATTTCTACAATGGCTCGGGCAACATGAAATTCTAATAACATGGTAATAATAGGATTGTGTGAAACAGTAAAAATGTTACTGTT
TCTTAACATTTTCTATTCGACTAAATTAGTGTTTATATTGTTCAAATGAAAGGCTACTTATATCAATGTTTAGATTTTTTATTTTTAAAAATAAATAGTA
AGATTAATTAAAAAAAATTGAGGAAGATTAATTAAAAATTGTTTTGTTAAGTTCTTGAAAAGATTCTTGTCTCTTAATAATGTCAATAGCCACCTCCACA
AGCCAACAAATTCTTTGAAAAATAAATTAGAATAAATCATCTACATCTCTTACACCTTCAAGTTTTCTATATAAACAAGTCATGAATGTAAATTTATTAC
ATTGCAAATATATATTGCAATATCAAAACTTAAAACAAGAAGTAGAGAAAT

>Medtr5g066070
AATGATCTAACACATTAGCGGGCATGACCGCAGATACACCTACCCACCGATATATATAGTTCCAAACAGAGCCAAATAAATTACAATGTAAAAATAAGTG
GGAAGATGATTCCACCTCGCCACATCCTCCTATACAAACTCTTCATCATACAAACCTAGAAAATCAAATTTCATTGGAAAAATCTCATAAAAAATGATAA
AATATTGTCATTTTATAACATAAAAGACTAAATCGTTACAAAAAAAAGATAAAGGACGAAAATGTTACCTAAAATTAAGTTAAGGGACTAATTATGTAAT
TTTGCCAAATTTTAAAGTTACTATCTTCGAGCTCCCCTCTATTGTAGCCGAAAAAGGGTATTATTTTATTTAATTGTAATTTATTTGCTTAAGTTTTTGA
AAAGTTTTTTGTCTCTTAATAACTACAATGGTCACCTCCACAAGCCATTATATTCTTTAAAAATAGAATTACAATAAAATAATACACCTCTTTGCACCTC
AAACTTTCTATATAAACAAAGGATGCATGTAACTTTATTGCATTCAAAATACAAATAATAAAACAAAAACAAGAAAAGAGAGAAAT

>Medtr5g080400
TTTATCTTATATATTGTTTTACTGCAATATATTTTTGTTGAGTTTGAAATTTTTATGTATTTCTTATGGTGTAGAATTCTTTACATTTTATACAAGTAGG
AAATTTTAAAAAACCTACACAGTCTTTGGTTATATTTTTTCAACTTTACTTTTTTTTTTCACCTCGTTTTTCAACATTATTGTGATTAGCATTTTTTTTA
TTAGATTAAACAATGATTACGTTGTAATTAAAAAAAACTAAGGATTACATTGATTGCATGGAATAATATTTATATTAATGTTCAGTTTATTTTATTTTAA
TTTAATTTTTAAAAATTTATTACGTCTATTTAATTAAAGTTTGTTTGAATAATTTTTTTAAAAGATTCTTGTCTCTTAATAATTTCAATGGTCACCCCCA
GAAGCCAATAGATTCTTTGAAAAAACAAGCACATGAATTTACTTCATTTCTTTCCACCTCAAAGTTGTCTATATAAACAAGTATTTAATGTAAAATTATA
ATATCGACAACACAAACAACAACAAAATTAGAAAAGAAGAAGAGAAAA

>Medtr5g080440
ATTGATGTCTCTGGAATACTCATGTCTCCAGATCTGTAACGTTGACGACGACAAGTCATTGTCATTGATCGGATCTGCACTTGCTCAAAGTTAATCTCAA
CCTGAATCAAATAAACACCGCACTAAGACAGGAAATAAAAAACACACCGCACAAAGACGGGAAATCAAAACAAACAGAGTCGTTCAAAGACGACGAAATC
ACAAAATCAAACGAAAGAAAACAGAAAATATGTGAAATCACTTATTCAATTAAAATAGAAGAAAAAACAATTCGAATGGGTGATTTTTGGGTCAAAATTG
ACCCTAAATCATGAAGAAGAAAAGTTTATTTAATAAAAAATATTATTAAGTTTTTGGAAAGATTCTTGTCTCTTAATAATTTTAATGGTCACCGCCACAA
GCCAATAGATTCTTTAAAAAAAAAACAAGCACATTAATATACTACATGTGTTTCACCCTTCAAGTTCTCTATATAAACAAGTGTCATATGTAAAATTATC
ACAACGACAATACAAACACAAAAAATACAAAAAAAATTAAAATTAAAAAAGAAGAAGGGAAAT

>Medtr5g080900
AAAGCTGAATATGTTTTTAATATCTACAAAATTCACTTTAGAACTTATTTTACATTCAAGGATTTAAAACACTAATTTTTTGTACAAAAAATATCACTTA
TTTTTGCAAGACGTTTTGAAATAGAGACCAAAATTGGATATAAGGAATAAACCGGAAATGACCAAAAAACTTCAAATTACAAGGACGAAATCCAAAAAAA
AAAAAAATTTATAGGAGGAAAAACCGGAAATGACTTATATTACAGGGGATAAAACACTATTAACCCATAAAATATTTATAAACTTTAGTGAAAACATGTT
TGAATAAGTTTTTAAATACATTATTGTTTCTTTGATAATGTCAATGACCATTTTCACAAGCCAATAGATTCTTTTAAATATAGTCACAACAAAGCAACTA
CATCTGTTTACACCAACAAGTTTTCTATATAAACAAGTATTGAATGAACAATTAATACATTGCAAAAATAAATAGCAAAACAAAATACAAGAAGAGAAAT

>Medtr5g081000
TATTTGTTGATTCCTTTGTCTCATTATAAGTGTATCATTTGCAATATTTTATTCTTTTAAATTAATTATCACTTTAAAATATCAATCAGCATTAATTATT
TTTTGCACTAATATATCTTTATTTTTTCTATCTTAACCCACTGAACTATAATTAATAAGGATATTTTAGAAAATCAAACTTATTTATCATTAAAATTAAA
TAAAATAATCATATTTATCATTAAAATGAGATATTTACAACGAGACAAATAAAACTTATTTTTATTAGATTAATTAATTATTACATTGACTAAATGAAAT
AATATTCATATTGATGATTATATTTGATTTTTTCTAAAATATTTATTAACTTTAATGAAAACTCGTTTGATTAAGTTTTTGAAAAGATTATTGTCTCTAT
AATATTGTCAATAGCCATTTCCACAAGCCAATAGATTCATTAAAAAAATAAATCGCACGAAATACTACATCTCTTTTCACCTCTAATTTTTCTATATAAA
CATGCATTGAATGTACAATTATTACATTGCAAATACAAATAGCAAAACAAAATAAAAAATAAGAAGAGAAAT

>Medtr7g110180
ATATTACAAGGGGTAAAGACCTATTAACCTTAAAATAAATGGACATCTACGTTACAATATTTTTTGTGTATAAATAACACACACCTCTTTGCCATCAAAT
TAAGCTCCCTTCTCATCCGAAAACTAATCTCTCTCTCTCTCTTCCCTCTACTTTCCACCAATCAATCTCAAAATTAGAAGATGTCGCTCTTTATTTAAAT
TTTTTCAATAGATTAATGAATGGGTTGATGAAATGATTTTTTTAATAAACGTGTTTATATTATTATTAATTTATTAATTTTTTTTTATTAATTTATTGAA
AAGATTCACGACTCTTTATAATAGTAATGGACGTTTCCACAAGCCAACAAATTTAAAATAAAATAAAAAAATGACAATAAATAAACTCCTTTTTTGTACA
CCTCCTGATTTTCTATATAAATAAACAAGCATTAGAAGTAAATTCTTATACCACAAGATACAAACACAAATAAGAAAAGGAGGGAAAAAGAAAGAGAAAT
```

### 5. Hemoglobin promoters of early diverged legumes and nodulating non-legumes presented in Fig S13C.

```
>Dg_nsHB1 isolate NF-2018-6 scaffold1056_cov82
AATCGAGTTAGATATTTTTACATGAGATACTAATCTTGCTTTCAATGGTCTGACTCTACTCACACGTTTC
TTAGTACTGCAAATTAAAAAAGAAAAAGAAAGAATTAATTGGCATTGAAAACAAAATCAATGACTCATAT
CTAGAGAAATTTTTATATAGAGCATCTTCTAAAAGAAAAAATCATTTTGAGATCTTTTTTATTCACTAGA
TCATGATCTAGATTTAACGAACTGGTATTTTTAATCTCTCTCATATTAAAAGACATGATTTCATAAAAGA
ATTTAAAAAGTCTTCAAAACTATTAAATATCTCATTTTAGAAGAATTGGAGAGGATCTCAATAGTCTCAT
ATTGATTATTCTATTCATTTACGAACACACTCGGTTAGACGTGATGGGTACAGATCTAATGTCACATAAA
AGGATGGCATATTCGAAATGTTGCTACTAGTCCCTATTAAATGAACGTATTAAATACGGTCAAAGGCTCT
CATTATAAGTCTACTTCAATCCCAAATACCCAAAACCAGGTTCGCTTTGCCGCAAACAGCCCCTTGTTTA
TGGACATCTGGTGGTTAATTAGACACACAACAACTTTCCCAACGATAAGGTTAAACTAAATCATTCAAAC
ACTCAGAAATGGCTCTTCGTTGTTATGAAGAGGTAATTTAAGTCTTGTTTGGGTTTCCTTTCCTATATAA
ATGAATTCCACTCAACATTTTTCAAAACACCATCCATTTCCATACCATCTTTACCAAATCAATTTTTTTA
AAACTCTTTTTAGCTTTTAAGACATTTGAAAACTAGACTCCAATG

>Dg_nsHB2 isolate NF-2018-6 scaffold27403_cov80
ATTTATTACCTAAAAAACACATATAAATGTCAGATTGCACGAATAAATGACACTTCAGTTTGTCAAAAAA
CATTTTTTTGTTTCTCTTTTGAACGTGAACTTAAATATTATAATAATAGTAAAAATGATAATAATTATTG
AAATAATTACAATGATAATTATTAATAATTATTACTATATAAAATTGGTCATGGGACTAATTAATAGATA
TCGTTGATGATGATGAAGGATACTAGAGATAGCTCCAGCTGCAAAGAACATGTTTTCTTTCCTAAAGGAT
TCGGATGAGATACCACAGAACAATCCTAAGCTCAAAGCTCACGCTGTTAAGGTTTTTAAGGTGGTATGTC
TTTTCTTTGCAAAACAAAAACTAATTTTTTATTATTATTTTCCAAGTTTATTGAGATTTCCAAATTTTTA
TATAAATAATTTGAAAATTTTATGACTTGTGACTAGTGATCTGACAGACTTCTATATTTTATAATAGGGG
TTGTTCTTGTTATATTTATAATTTTTTTAATTTTAATTTTAGAAAACAATAAATAAAAACTAATTGATCT
AAATTAATAATGTGATG

>LupAng_Hb1_XP_019465301
TTTTATCTAGTGTGGTGCAGGGGAAAAATCACACTATGAAATAGAAAAATATTATGTGAACAAAATATGA
TATGCCAACATGACAACTCTATGATCCCATTGACTCACTAAACTCTAAATTGTAATAATATGATCCACTA
TCTTACCATTATTTCTTATTTGTCCTGCTCTGTATATCCCGAGTAATGACCGTCTCCATTAATGAAAAGA
CAAGTGCTACAAAATAATATGAAATGACACTTTTATTATTATCATATTACATTTTTGTTATTGAAAATAT
AAGAAAAAACAATCACATGTCTCTCTATTAATAGAAAAGATACTTGTTGATTGAAATGTGCATGTTGGGT
AGCAAAAACCATTTTGTTCTCAGCACCTAGTTTTAATGATACTTTTAAGGATCTTCCTGACGCAATTCAG
CATCACAGCCCAACCAAACTAGAATGGCTCCCTATTGTCATGAAGGGACAATTTCATTTCCTATAAGAAC
TTGTTCATAGCAAACTCCAAAATTTCACTCATCCACTCCCATCACAATTCACAAACCAACACCTTAGCAG
TTAGCATATCTGAGATCATGAGCACCCTGGAAGTAAGAG

>LupAng_Hb2_XP_019450687
TTAGAAATTCATGTTGAATGCACTTACCAATTTGAAAACCTTTCCACCATGGACTTGTAGATCAAGGTTA
TTCTGGGGTATTTCATTAGAGCCCCTTAGAAATGAAAACAAGTCTTTTGCTGCAGGTGCAATCTCCACTA
CACTATTTTCAATCAAACCATAACCAAACAGTTAACTCATACATTTATCTGTTGATTGTAATATCTATGC
AATGACAATATAAAGATTTATACACAATCAATATATCATTTGCCATCCATTTGTATTATTTTTTGAAGAA
TCAATAATGATAAAAGTGAATAATAGTGTATAAAACTCATAAATATCAGTGCATATTTTTTCTCTCTCTT
ATCTCTTTCATAAATACAACAAAATGGAAAATACATAGTAAAAGTAGTGAACGTACAAGGTGAAGAAACG
ATGAGTATTTTGAGGTATGTTTGAGTTAAATACTTCCCATGAACTCTTCACCAAAGCCACTTGCTTCTCA
GTAAAAGCAAC

>LupAng_Hb2_XP_019433600
TTTCATACCATGACCAAAGAGTAGAGACATTTGACCTCACATTAGCTCTTTCCAAAACATAATATATAAA
AGAAGAATATGTTCTCCCAAATAGTATTTTGAACCATGTCTCGCACTAATAATATTAAAATAAAACAGAA
CAAATAATGAATATTTAATTTTTTAATTGACATGATATATTTTTATTGATGCAATACATGATTCGGGATA
CCATTTGGGATAACGAATCTCTGTCATATATATAAGGAAAAGATACAATGAAAAAGAAAACTTACAAGGT
GAAGAAACGATGGGTGTGTTTAGGAATATTTGCGTTAAATTCTTCAAATGAACTCTTCACCAAAGCCACT
TGCACATCAGTTAAAATAGCCATATTCCTTTTACTCTTTCTTTTATTTGTGATCTGTATGATTTTGCGCC
CCAAGACTCATATATATATAGGAGGCTAAAAGTGAATTTCTTGCAAGTGTTATTCACATTGGCTTCTTGTGT
TTTCTAATGGC

>Pro_alb_nsHB1_NW_021636642_rc
GTTGCTTTTACTGAGAAGCAAGTGGCTTTGGTGAAGAGTTCATGGGAAGTATTTAACTCAAACATACCTCAAAATACTCATCGTTTCTTCACCTTGT
ACGTTCACTACTTTTACTATGTATTTTCCATTTTGTTGTATTTATGAAAGAGATAAGAGAGAGAAAAAATATGCACTGATATTTATGAGTTTTATAC
ACTATTATTCACTTTTATCATTATTGATTCTTCAAAAAATAATACAAATGGATGGCAAATGATATATTGATTGTGTATAAATCTTTATATTGTCATT
GCATAGATATTACAATCAACAGATAAATGTATGAGTTAACTGTTTGGTTATGGTTTGATTGAAAATAGTGTAGTGGAGATTGCACCTGCAGCAAAAG
ACTTGTTTTCATTTCTAAGGGGCTCTAATGAAATACCCCAGAATAACCTTGATCTACAAGTCCATGGTGGAAAGGTTTTCAAATTGGTAAGTGCATT
CAACATGAATTTCTAA

>Pro_alb_nsHB2_XP_028763937
CGCGAAAAAGATGGACGGAAGAAAGAATCACGTTTATAATTTTAATAGAAGACATATTTGGCCAAGTGAA
GGAGAGGCCATATATTTATGTACGTGTGTGTGTTTATATATATTAGGAATCAATGAGCCGAACCACTTCT
CTTCTTTTTTTAATCTCCCCCCCCCCCCCAACTTTTTTAGTTTTTTAATTAGTCTTGTAAGAAAAATAGT
GTAATTTATTCGAGCGGGCTAGGAGCCTAAGAGGCGTTCCAGTGGGTTCTAATAGTGAAGTTTAACCGTT
GGAGGCACTCAAAGCTAATCTTGGCTCTTGAGATGCGTTTATTGGACACCAAAAGAATTCCATCTACTAT
CTATTTTGAAAATTCCTTCCACATCAACTTGAATTCAGGTAATTTGATCTCTCTTGCTTTCATATCAACA
GCTTCTATATAGAGTCGCTAATCTAAGTTTAGTCTTTTGTGAATATCAGAAAGTAAATTAAAGATAGATT
ATCAATGGGTG
```