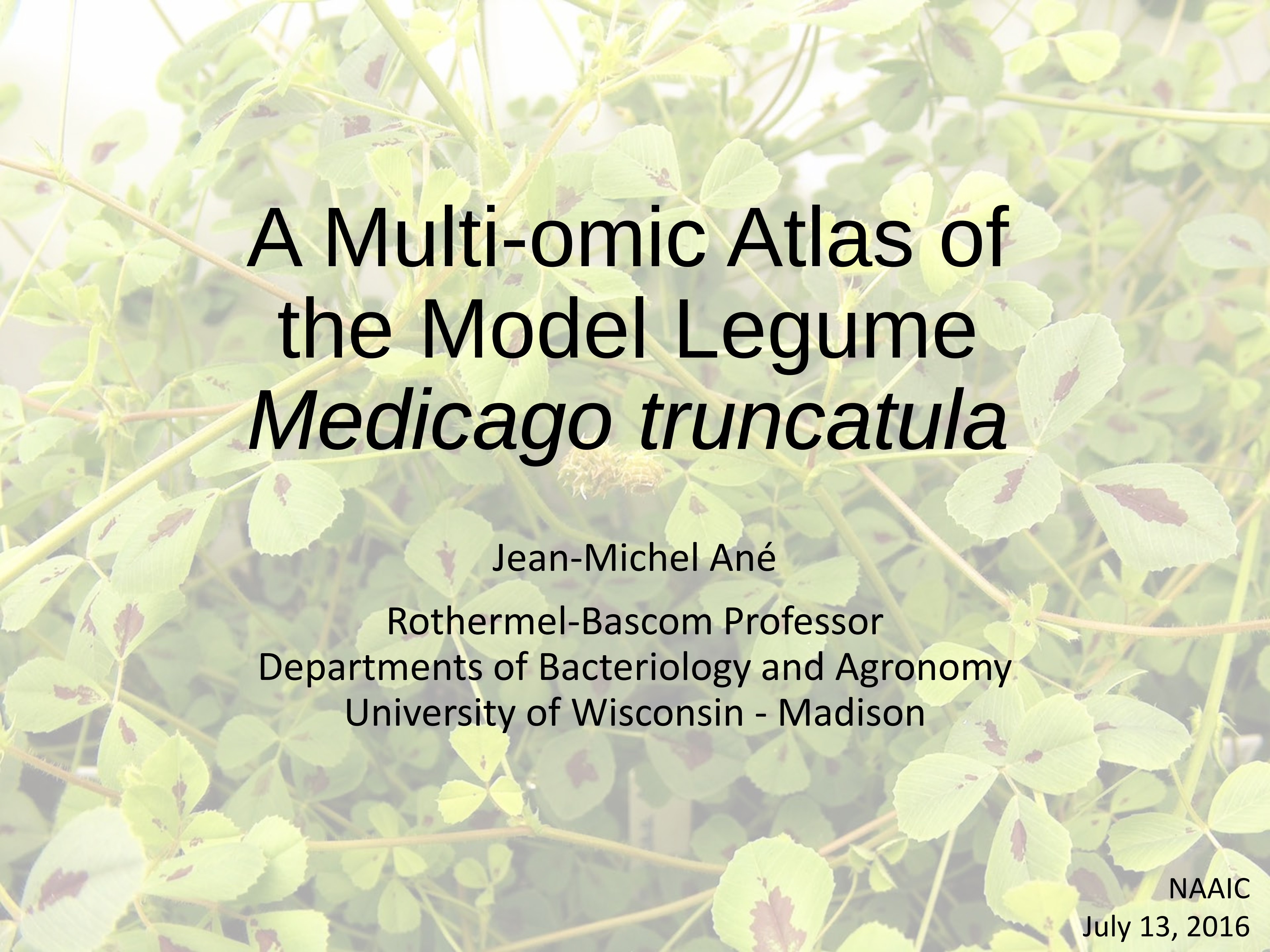


A Multi-omic Atlas of the Model Legume *Medicago truncatula*

Jean-Michel Ané

Rothermel-Bascom Professor
Departments of Bacteriology and Agronomy
University of Wisconsin - Madison

NAAIC
July 13, 2016

Model Legume
Medicago truncatula

Diploid (2n=16)

Autogamous (homozygous lines)

Small genome size (550 Mb / 1C)

Genetic, physical and cytogenetic maps

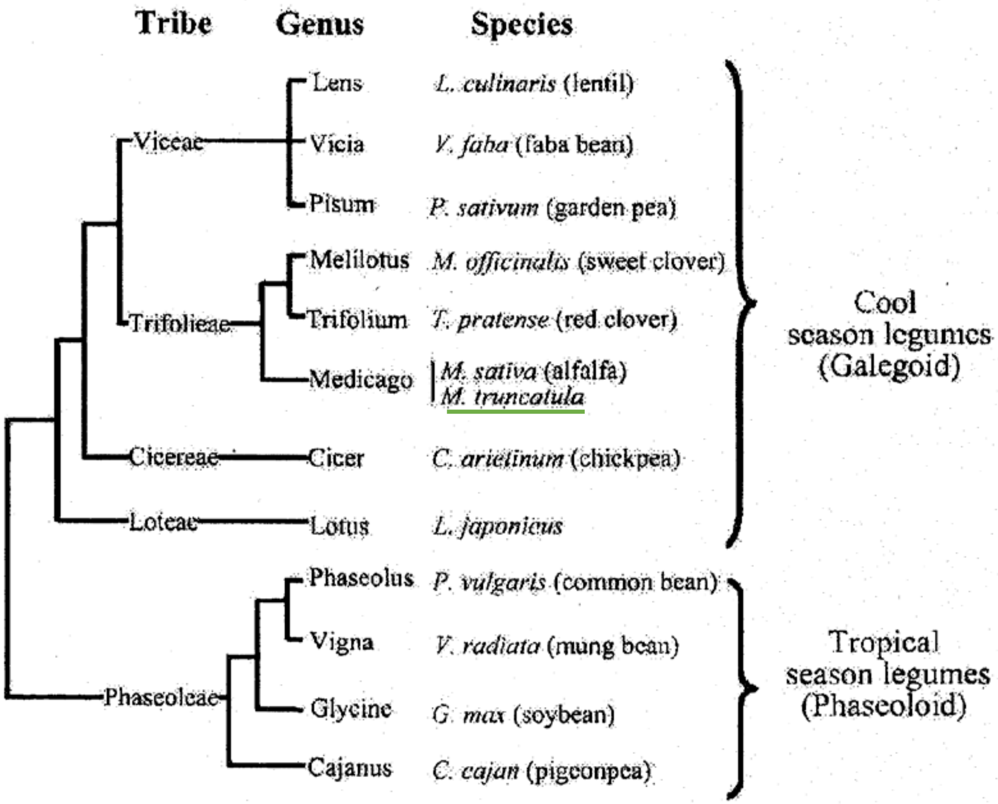
Whole Genome Sequence

Efficient transformation *by Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*

High level of natural polymorphism

Many mutant populations available: Ethyl Methyl Sulfonate (EMS), g-Ray, Fast Neutron Bombardment (FNB)…

RNAi and CRISPR/Cas systems



genome

doi:10.1038/nature10625

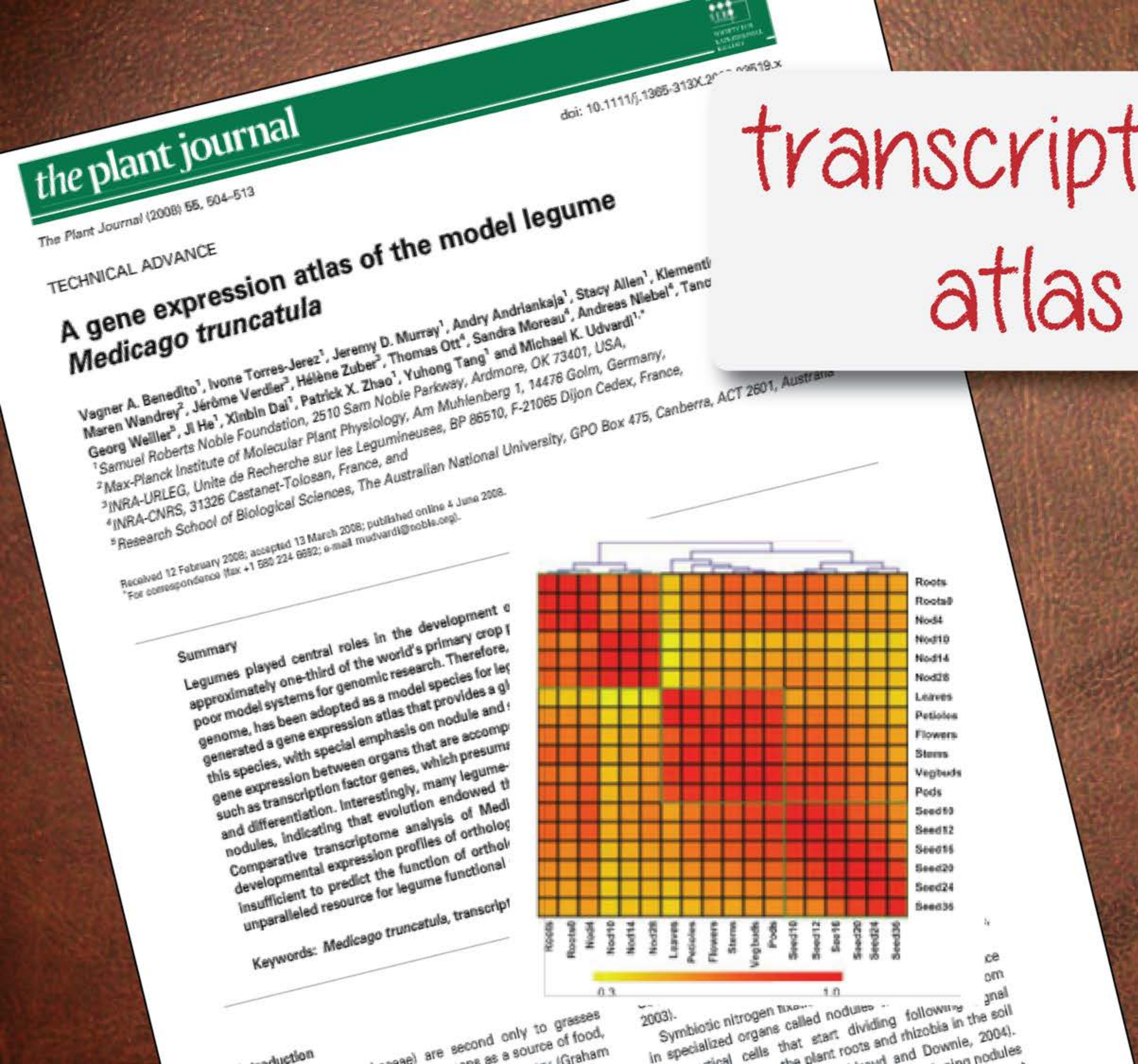
LETTER

The *Medicago* genome provides insight into the evolution of rhizobial symbioses

Nevin D. Young[1]*, Frédéric Debellé[2,3]*, Giles E. D. Oldroyd[4]*, Rene Geurts[5], Steven B. Cannon[6,7], Michael K. Udvardi[8],
Vagner A. Benedito[9], Klaus F. X. Mayer[10], Jérôme Gouzy[2,3], Heiko Schoof[11], Yves Van de Peer[12], Sebastian Proost[12],
Douglas R. Cook[13], Blake C. Meyers[14], Manuel Spannagl[10], Foo Cheung[15], Stéphane De Mita[5], Vivek Krishnakumar[15],
Heidrun Gundlach[10], Shiguo Zhou[16], Joann Mudge[17], Arvind K. Bharti[17], Jeremy D. Murray[4,8], Marina A. Naoumkina[8],
Benjamin Rosen[13], Kevin A. T. Silverstein[18], Haibao Tang[15], Stephane Rombauts[12], Patrick X. Zhao[8], Peng Zhou[1], Valérie Barbe[19],
Philippe Bardou[2,3], Michael Bechner[16], Arnaud Bellec[20], Anne Berger[19], Hélène Bergès[20], Shelby Bidwell[15], Ton Bisseling[5,21],
Nathalie Choisne[19], Arnaud Couloux[19], Roxanne Denny[1], Shweta Deshpande[22], Xinbin Dai[8], Jeff J. Doyle[23], Anne-Marie Dudez[2,3],
Andrew D. Farmer[17], Stéphanie Fouteau[19], Carolien Franken[5], Chrystel Gibelin[2,3], John Gish[13], Steven Goldstein[16],
Alvaro J. González[24], Pamela J. Green[14], Asis Hallab[25], Marijke Hartog[5], Axin Hua[22], Sean J. Humphray[26], Dong-Hoon Jeong[14],
Yi Jing[22], Anika Jöcker[25], Steve M. Kenton[22], Dong-Jin Kim[13,27], Kathrin Klee[25], Hongshing Lai[22], Chunting Lang[5], Shaoping Lin[22],
Simone L. Macmil[22], Ghislaine Magdelenat[19], Lucy Matthews[26], Jamison McCorrison[15], Erin L. Monaghan[15], Jeong-Hwan Mun[13,28],
Fares Z. Najar[22], Christine Nicholson[26], Céline Noirot[29], Majesta O'Bleness[22], Charles R. Paule[1], Julie Poulain[19], Florent Prion[2,3],
Baifang Qin[22], Chunmei Qu[22], Ernest F. Retzel[17], Claire Riddle[26], Erika Sallet[2,3], Sylvie Samain[19], Nicolas Samson[2,3],
Iryna Sanders[22], Olivier Saurat[2,3], Claude Scarpelli[19], Thomas Schiex[29], Béatrice Segurens[19], Andrew J. Severin[7],
D. Janine Sherrier[14], Ruihua Shi[22], Sarah Sims[26], Susan R. Singer[30], Senjuti Sinharoy[8], Lieven Sterck[12], Agnès Viollet[19],
Bing-Bing Wang[1], Keqin Wang[22], Mingyi Wang[8], Xiaohong Wang[1], J mann[25], Jean Weissenbach[19], Doug D. White[22],
Jim D. White[22], Graham B. Wiley[22], Patrick Wincker[19], Yanbo Xing Yao[22], Fu Ying[22], Jixian Zhai[14],
Liping Zhou[22], Antoine Zuber[2,3], Jean Dénarié[2,3], Richard A. Dixon Schwartz[16], Jane Rogers[31],
Francis Quétier[19], Christopher D. Town[15] & Bruce A. Roe[22]

Legumes (Fabaceae or Leguminosae) are unique among cultivated plants for their ability to carry out endosymbiotic nitrogen fixation with rhizobial bacteria, a process that takes place in a specialized structure known as the nodule. Legumes belong to one of the tw main groups of eurosids, the Fabidae, which includes most speci capable of endosymbiotic nitrogen fixation[1]. Legumes compr several evolutionary lineages derived from a common ance 60 million years ago (Myr ago). Papilionoids are the largest cl dating nearly to the origin of legumes and containing most culti species[2]. *Medicago truncatula* is a long-established model f study of legume biology. Here we describe the draft sequence *M. truncatula* euchromatin based on a recently complete assembly supplemented with Illumina shotgun sequence, capturing ~94% of all *M. truncatula* genes. A whole-genome tion (WGD) approximately 58 Myr ago had a major role i the *M. truncatula* genome and thereby contributed to the of endosymbiotic nitrogen fixation. Subsequent to the *M. truncatula* genome experienced higher levels of re than two other sequenced legumes, *Glycine max* and *Lo*

transcriptome atlas



the plant journal

The Plant Journal (2008) 55, 504–513

doi: 10.1111/j.1365-313X.2

TECHNICAL ADVANCE

A gene expression atlas of the model legume *Medicago truncatula*

Vagner A. Benedito[1], Ivone Torres-Jerez[1], Jeremy D. Murray[1], Andry Andriankaja[1], Stacy Allen[1], Klementi
Maren Wandrey[2], Jérôme Verdier[3], Hélène Zuber[3], Thomas Ott[4], Sandra Moreau[4], Andreas Niebel[4], Tanc
Georg Weiller[5], Ji He[1], Xinbin Dai[1], Patrick X. Zhao[1], Yuhong Tang[1] and Michael K. Udvardi[1,*]
[1]*Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA,*
[2]*Max-Planck Institute of Molecular Plant Physiology, Am Muhlenberg 1, 14476 Golm, Germany,*
[3]*INRA-URLEG, Unite de Recherche sur les Legumineuses, BP 86510, F-21065 Dijon Cedex, France,*
[4]*INRA-CNRS, 31326 Castanet-Tolosan, France, and*
[5]*Research School of Biological Sciences, The Australian National University, GPO Box 475, Canberra, ACT 2601, Austral*

Received 12 February 2008; accepted 13 March 2008; published online 4 June 2008.
*For correspondence (fax +1 580 224 6692; e-mail mudvardi@noble.org).

Summary

Legumes played central roles in the development c
approximately one-third of the world's primary crop p
poor model systems for genomic research. Therefore,
genome, has been adopted as a model species for leg
generated a gene expression atlas that provides a gl
this species, with special emphasis on nodule and s
gene expression between organs that are accomp
such as transcription factor genes, which presuma
and differentiation. Interestingly, many legume-
nodules, indicating that evolution endowed th
Comparative transcriptome analysis of Medi
developmental expression profiles of ortholog
insufficient to predict the function of orthol
unparalleled resource for legume functional

Keywords: *Medicago truncatula*, transcript

troduction

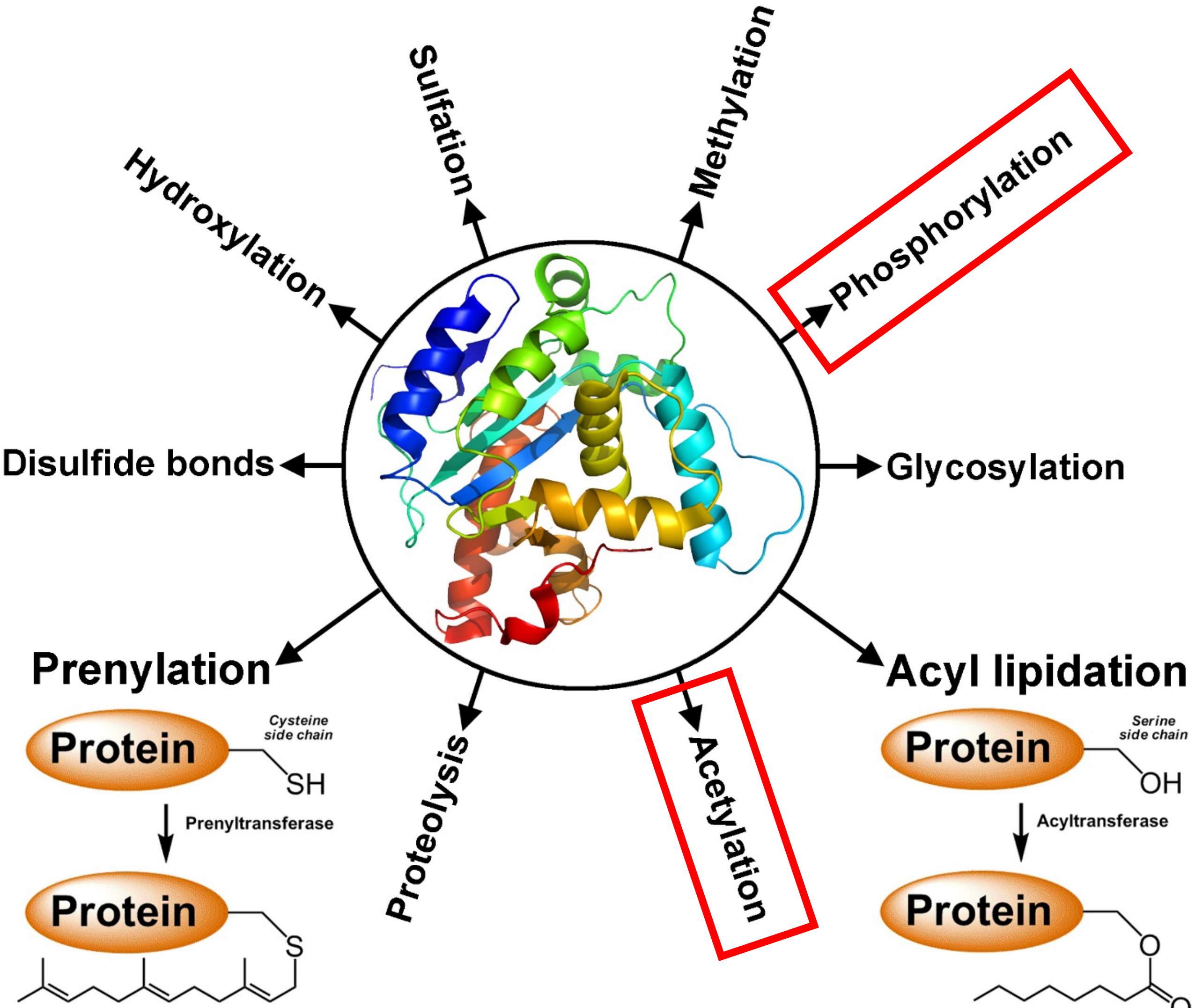
) are second only to grasses
as a source of food,
(Graham

2003).

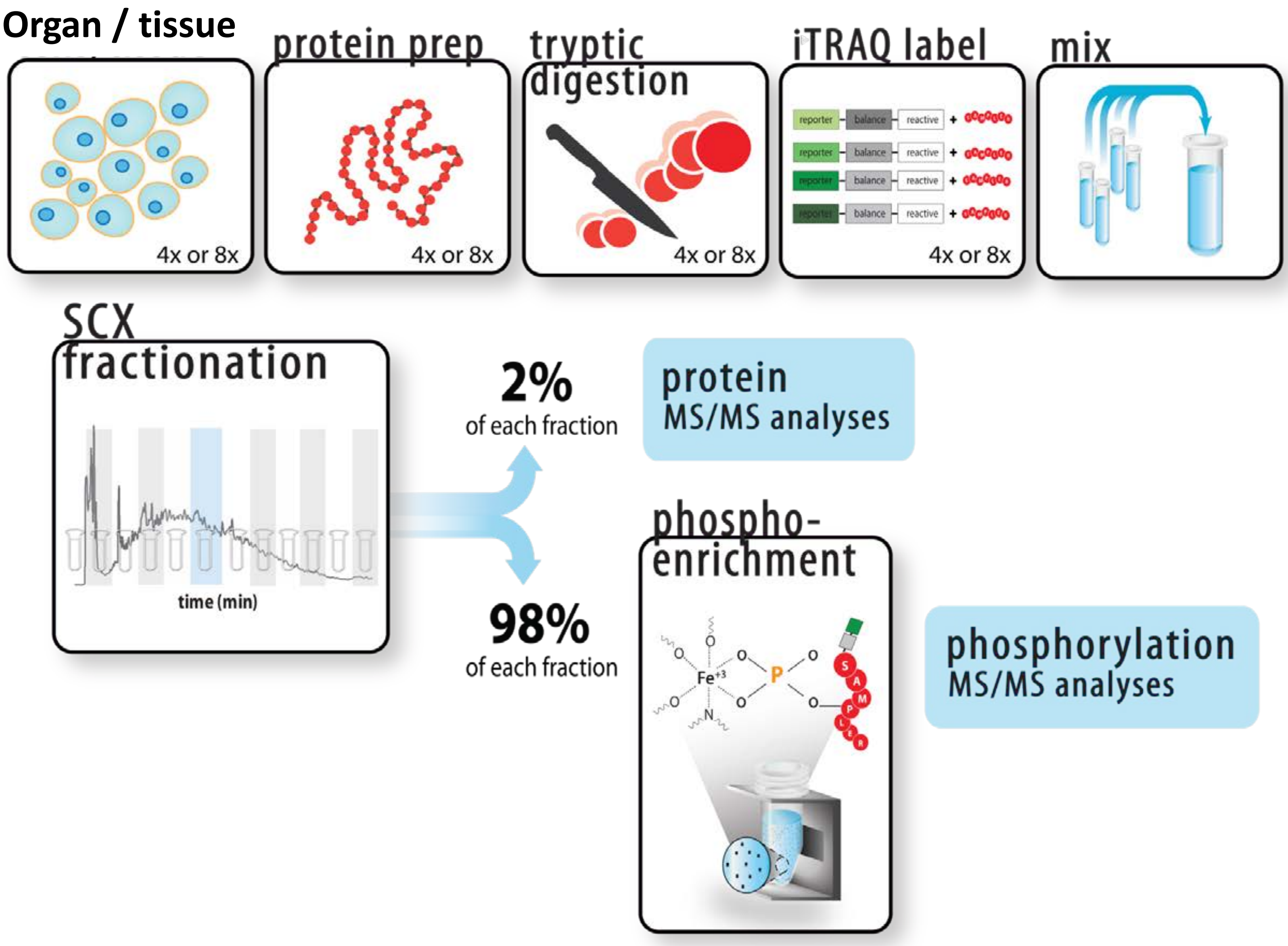
Symbiotic nitrogen fixa nodules following
in specialized organs called nodules following
cells that start dividing
the plant roots and rhizobia in the soil
and Downie, 2004).

proteome atlas?

Post-Translational Modifications



Quantitative proteomics workflow



Organ / tissue

protein prep

tryptic digestion

iTRAQ label

mix

SCX fractionation

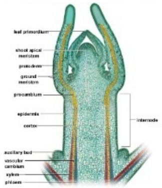
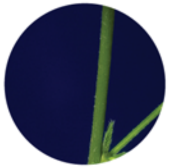
2% of each fraction

protein MS/MS analyses

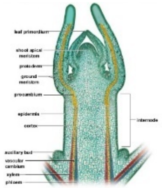
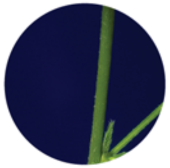
98% of each fraction

phospho-enrichment

phosphorylation MS/MS analyses

Plant Organ	Proteins identified
 *Flower*	13,783
 *Meristem (Bud)*	11,700
 *Leaf*	11,658
 *Stem*	9,919
 *Seed*	9,941

Plant Organ	Proteins identified
 *Root*	11,051
 *Nodule (10 dpi)*	11,099
  *Nodule (14 dpi)*	11,244
 *Nodule (28 dpi)*	9,388

Plant Organ	Proteins identified
 *Flow*	
 *Meris*	
 *Leaf*	
 *Stem*	
 *Seed*	9,941

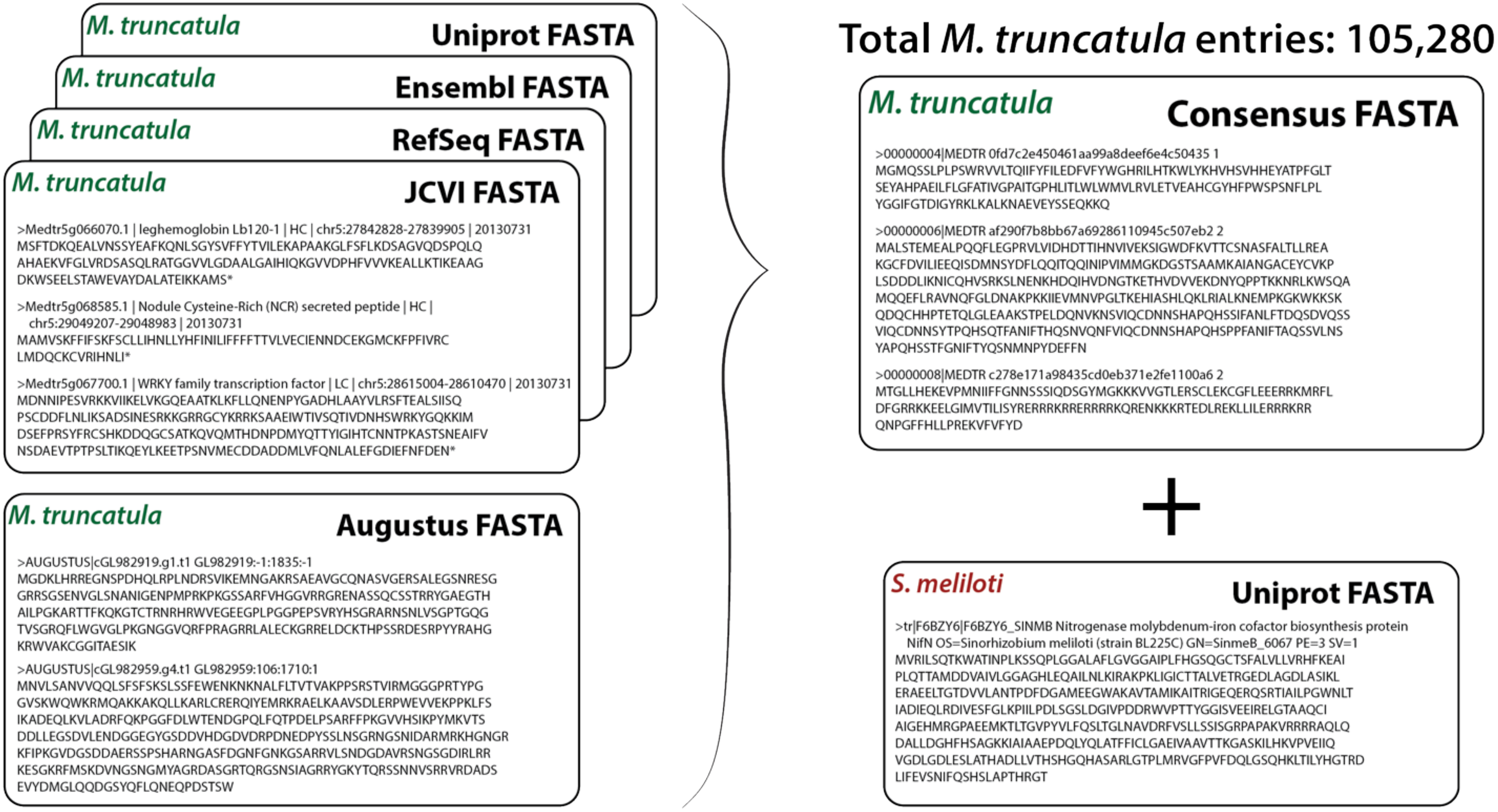
Plant Organ	Proteins identified
	11,051
	11,099
	11,244
	9,388

proteins: 18,798

phosphorylation sites: 21,452

acetylation sites: 674

M. truncatula consensus database



M. truncatula **Uniprot FASTA**

M. truncatula **Ensembl FASTA**

M. truncatula **RefSeq FASTA**

M. truncatula **JCVI FASTA**

```
>Medtr5g066070.1 | leghemoglobin Lb120-1 | HC | chr5:27842828-27839905 | 20130731
MSFTDKQEALVNSSYEAFKQNLSGYSVFFYTVILEKAPAAKGLFSFLKDSAGVQDSPQLQ
AHAEKVFGLVRDSASQLRATGGVVLGDAALGAIHIQKGVVDPHFVVVKEALLKTIKEAAG
DKWSEELSTAWEVAYDALATEIKKAMS*

>Medtr5g068585.1 | Nodule Cysteine-Rich (NCR) secreted peptide | HC |
   chr5:29049207-29048983 | 20130731
MAMVSKFFIFSKFSCLLIHNLLYHFINILIFFFFTTVLVECIENNDCEKGMCKFPFIVRC
LMDQCKCVRIHNLI*

>Medtr5g067700.1 | WRKY family transcription factor | LC | chr5:28615004-28610470 | 20130731
MDNNIPESVRKKVIIKELVKGQEAATKLKFLLQNENPYGADHLAAYVLRSFTEALSIISQ
PSCDDFLNLIKSADSINESRKKGRRGCYKRRKSAAEIWTIVSQTIVDNHSWRKYGQKKIM
DSEFPRSYFRCSHKDDQGCSATKQVQMTHDNPDMYQTTYIGIHTCNNTPKASTSNEAIFV
NSDAEVTPTPSLTIKQEYLKEETPSNVMECDDADDMLVFQNLALEFGDIEFNFDEN*
```

M. truncatula **Augustus FASTA**

```
>AUGUSTUS|cGL982919.g1.t1 GL982919:-1:1835:-1
MGDKLHRREGNSPDHQLRPLNDRSVIKEMNGAKRSAEAVGCQNASVGERSALEGSNRESG
GRRSGSENVGLSNANIGENPMPRKPKGSSARFVHGGVRRGRENASSQCSSTRRYGAEGTH
AILPGKARTTFKQKGTCTRNRHRWVEGEEGPLPGGPEPSVRYHSGRARNSNLVSGPTGQG
TVSGRQFLWGVGLPKGNGGVQRFPRAGRRLALECKGRRELDCKTHPSSRDESRPYYRAHG
KRWVAKCGGITAESIK

>AUGUSTUS|cGL982959.g4.t1 GL982959:106:1710:1
MNVLSANVVQQLSFSFSKSLSSFEWENKNKNALFLTVTVAKPPSRSTVIRMGGGPRTYPG
GVSKWQWKRMQAKKAKQLLKARLCRERQIYEMRKRAELKAAVSDLERPWEVVEKPPKLFS
IKADEQLKVLADRFQKPGGFDLWTENDGPQLFQTPDELPSARFFPKGVVHSIKPYMKVTS
DDLLEGSDVLENDGGEGYGSDDVHDGDVDRPDNEDPYSSLNSGRNGSNIDARMRKHGNGR
KFIPKGVDGSDDAERSSPSHARNGASFDGNFGNKGSARRVLSNDGDAVRSNGSGDIRLRR
KESGKRFMSKDVNGSNGMYAGRDASGRTQRGSNSIAGRRYGKYTQRSSNNVSRRVRDADS
EVYDMGLQQDGSYQFLQNEQPDSTSW
```

Total *M. truncatula* entries: 105,280

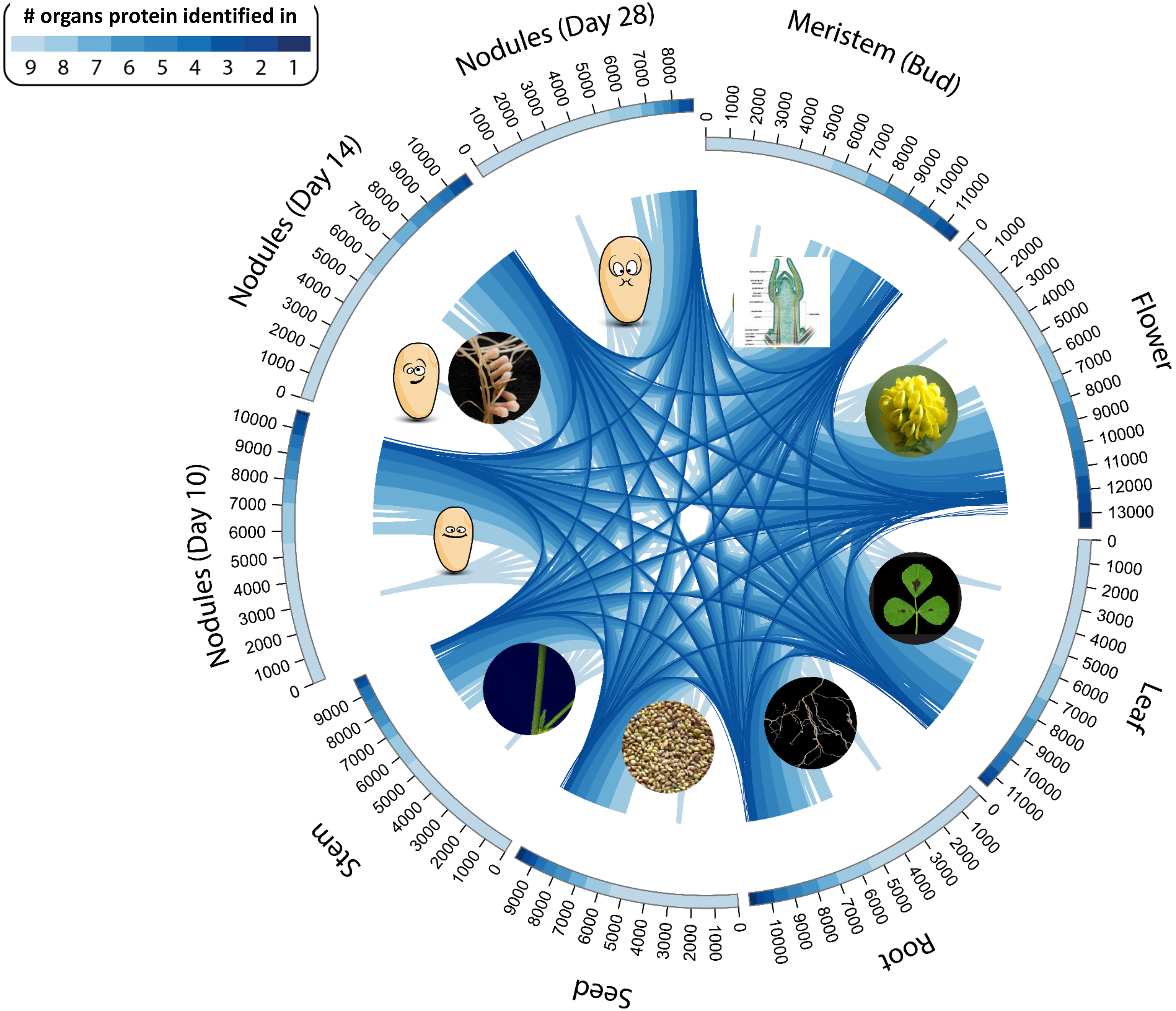
M. truncatula **Consensus FASTA**

```
>00000004|MEDTR 0fd7c2e450461aa99a8deef6e4c50435 1
MGMQSSLPLPSWRVVLTQIIFYFILEDFVFYWGHRILHTKWLYKHVHSVHHEYATPFGLT
SEYAHPAEILFLGFATIVGPAITGPHLITLWLWMVLRVLETVEAHCGYHFPWSPSNFLPL
YGGIFGTDIGYRKLKALKNAEVEYSSEQKKQ

>00000006|MEDTR af290f7b8bb67a69286110945c507eb2 2
MALSTEMEALPQQFLEGPRVLVIDHDTTIHNVIVEKSIGWDFKVTTCSNASFALTLLREA
KGCFDVILIEEQISDMNSYDFLQQITQQINIPVIMMGKDGSTSAAMKAIANGACEYCVKP
LSDDDLIKNICQHVSRKSLNENKHDQIHVDNGTKETHVDVVEKDNYQPPTKKNRLKWSQA
MQQEFLRAVNQFGLDNAKPKKIIEVMNVPGLTKEHIASHLQKLRIALKNEMPKGKWKKSK
QDQCHHPTETQLGLEAAKSTPELDQNVKNSVIQCDNNSHAPQHSSIFANLFTDQSDVQSS
VIQCDNNSYTPQHSQTFANIFTHQSNVQNFVIQCDNNSHAPQHSPPFANIFTAQSSVLNS
YAPQHSSTFGNIFTYQSNMNPYDEFFN

>00000008|MEDTR c278e171a98435cd0eb371e2fe1100a6 2
MTGLLHEKEVPMNIIFFGNNSSSIQDSGYMGKKKVVGTLERSCLEKCGFLEEERRKMRFL
DFGRRKKEELGIMVTILISYRERRRKRRERRRRKQRENKKKRTEDLREKLLILERRRKRR
QNPGFFHLLPREKVFVFYD
```

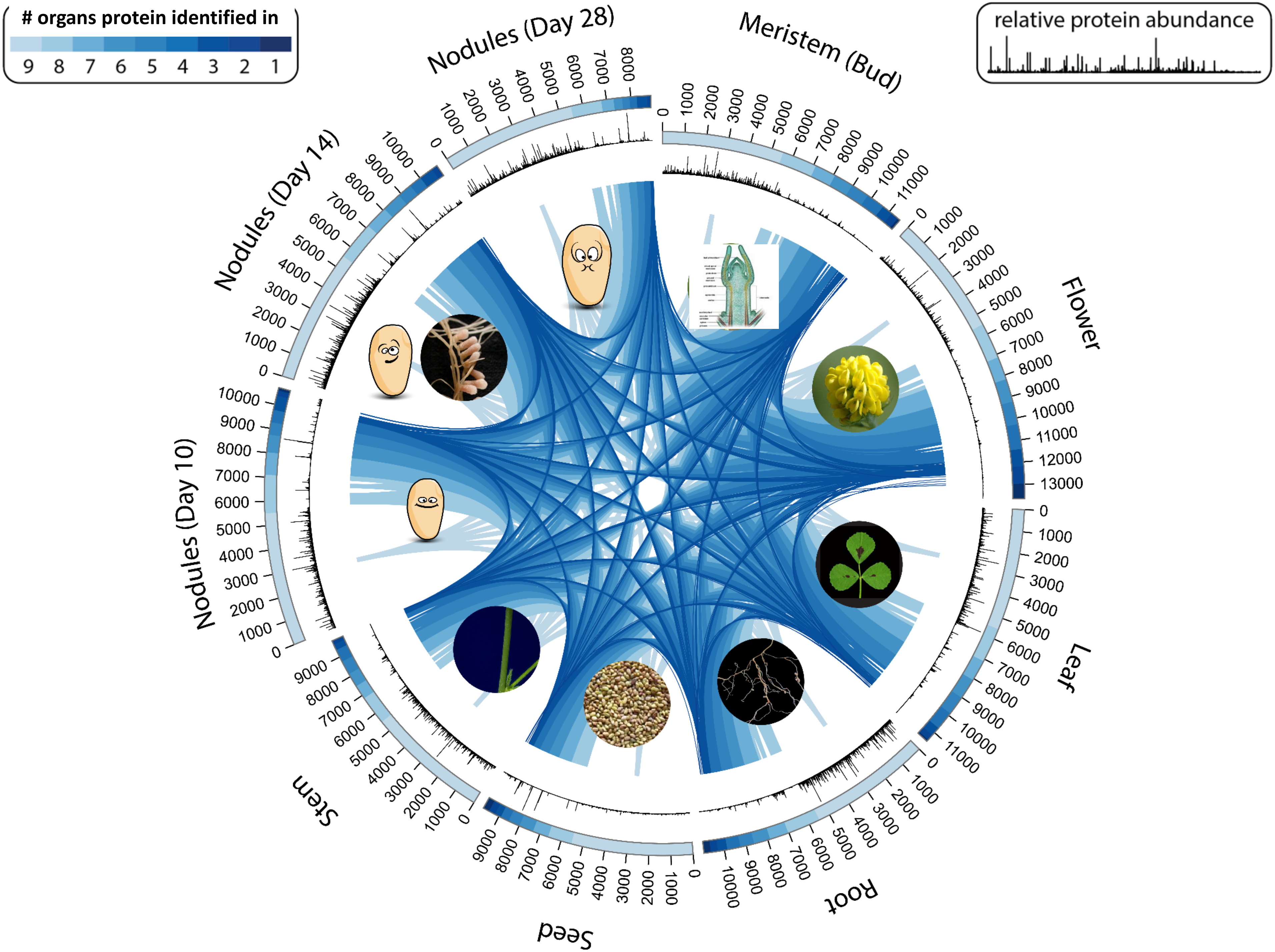
+

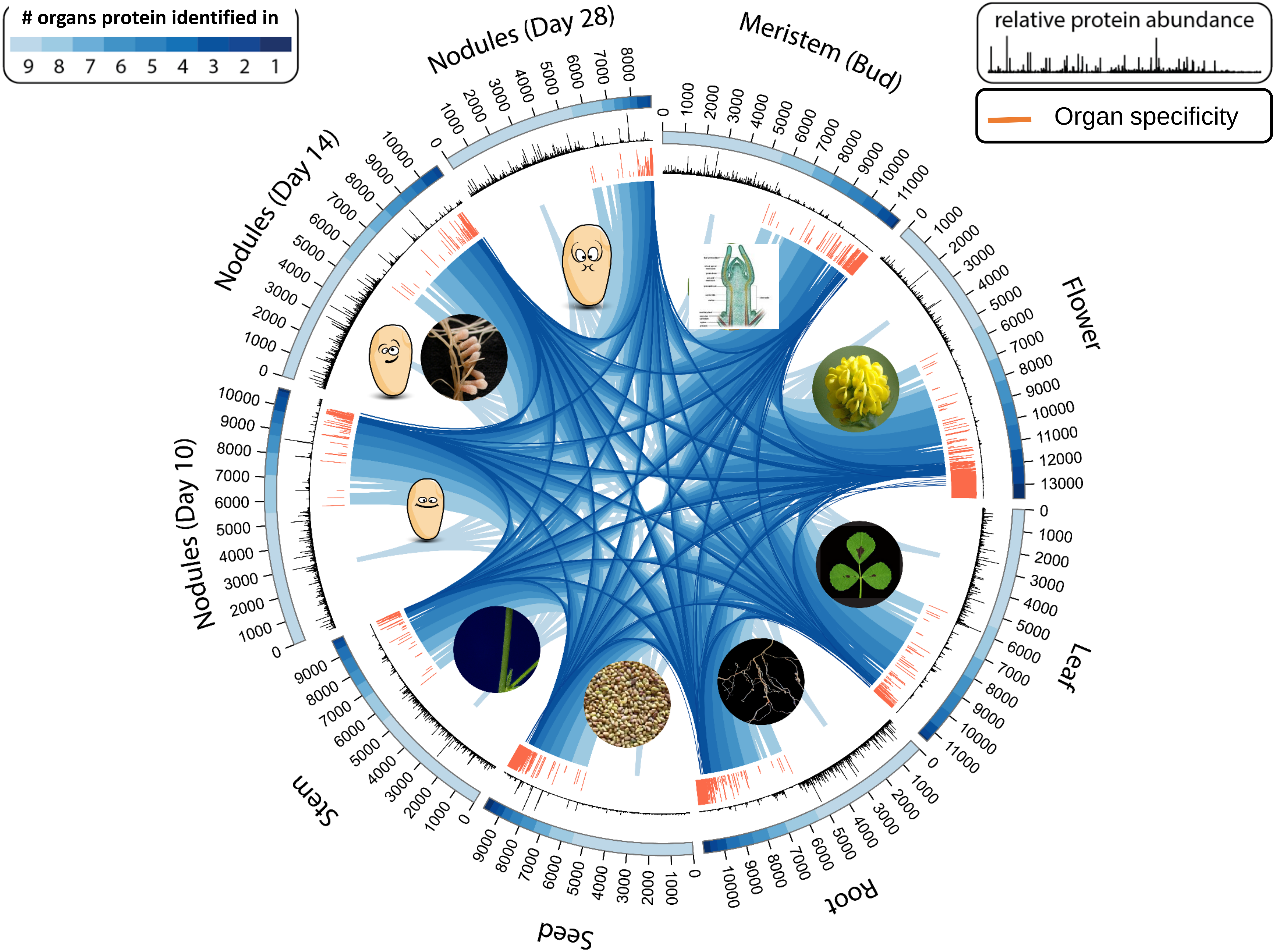
S. meliloti **Uniprot FASTA**

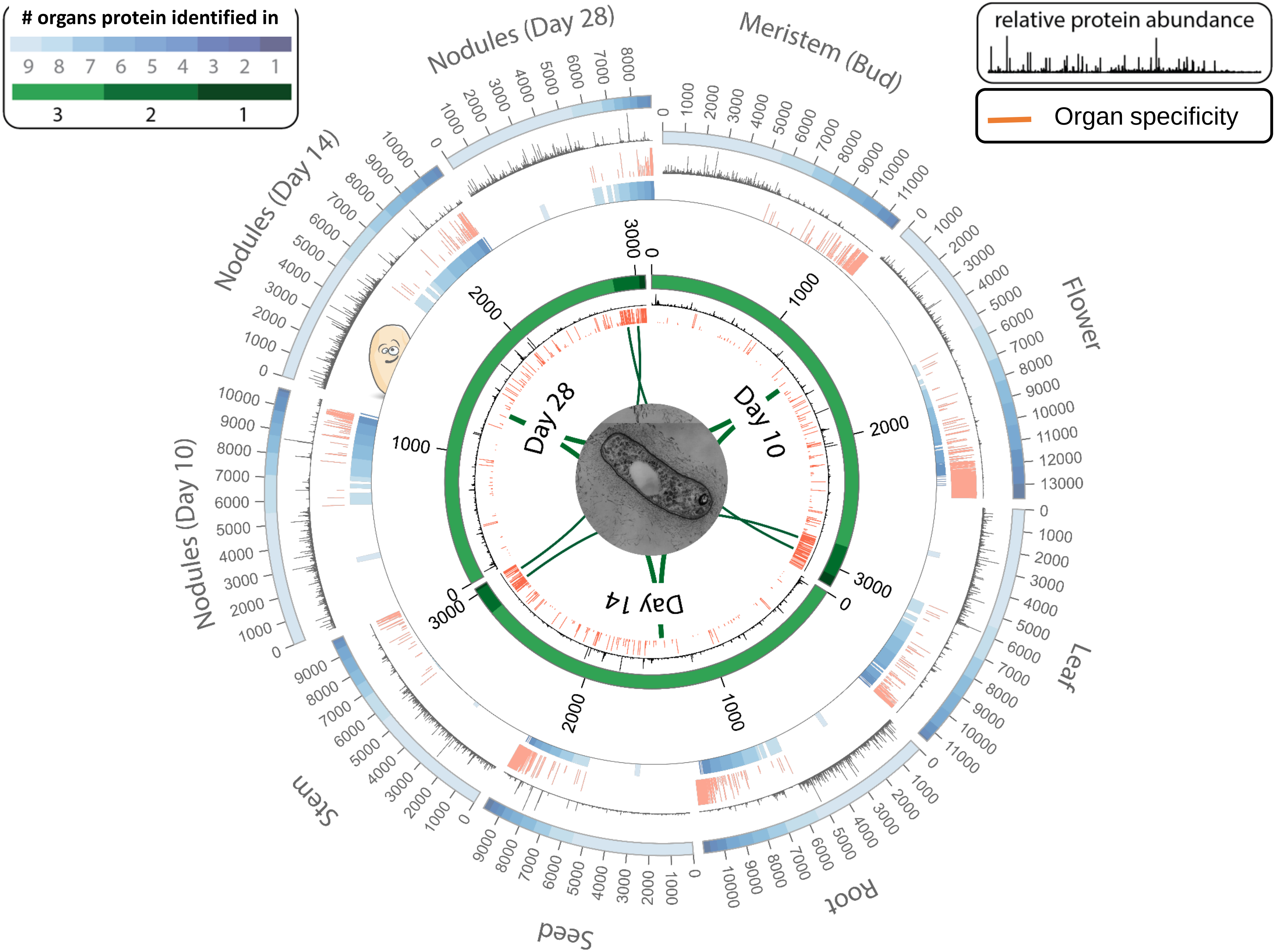
```
>tr|F6BZY6|F6BZY6_SINMB Nitrogenase molybdenum-iron cofactor biosynthesis protein
   NifN OS=Sinorhizobium meliloti (strain BL225C) GN=SinmeB_6067 PE=3 SV=1
MVRILSQTKWATINPLKSSQPLGGALAFLGVGGAIPLFHGSQGCTSFALVLLVRHFKEAI
PLQTTAMDDVAIVLGGAGHLEQAILNLKIRAKPKLIGICTTALVETRGEDLAGDLASIKL
ERAEELTGTDVVLANTPDFDGAMEEGWAKAVTAMIKAITRIGEQERQSRTIAILPGWNLT
IADIEQLRDIVESFGLKPIILPDLSGSLDGIVPDDRWVPTTYGGISVEEIRELGTAAQCI
AIGEHMRGPAEEMKTLTGVPYVLFQSLTGLNAVDRFVSLLSSISGRPAPAKVRRRRAQLQ
DALLDGHFHSAGKKIAIAAEPDQLYQLATFFICLGAEIVAAVTTKGASKILHKVPVEIIQ
VGDLGDLESLATHADLLVTHSHGQHASARLGTPLMRVGFPVFDQLGSQHKLTILYHGTRD
LIFEVSNIFQSHSLAPTHRGT
```



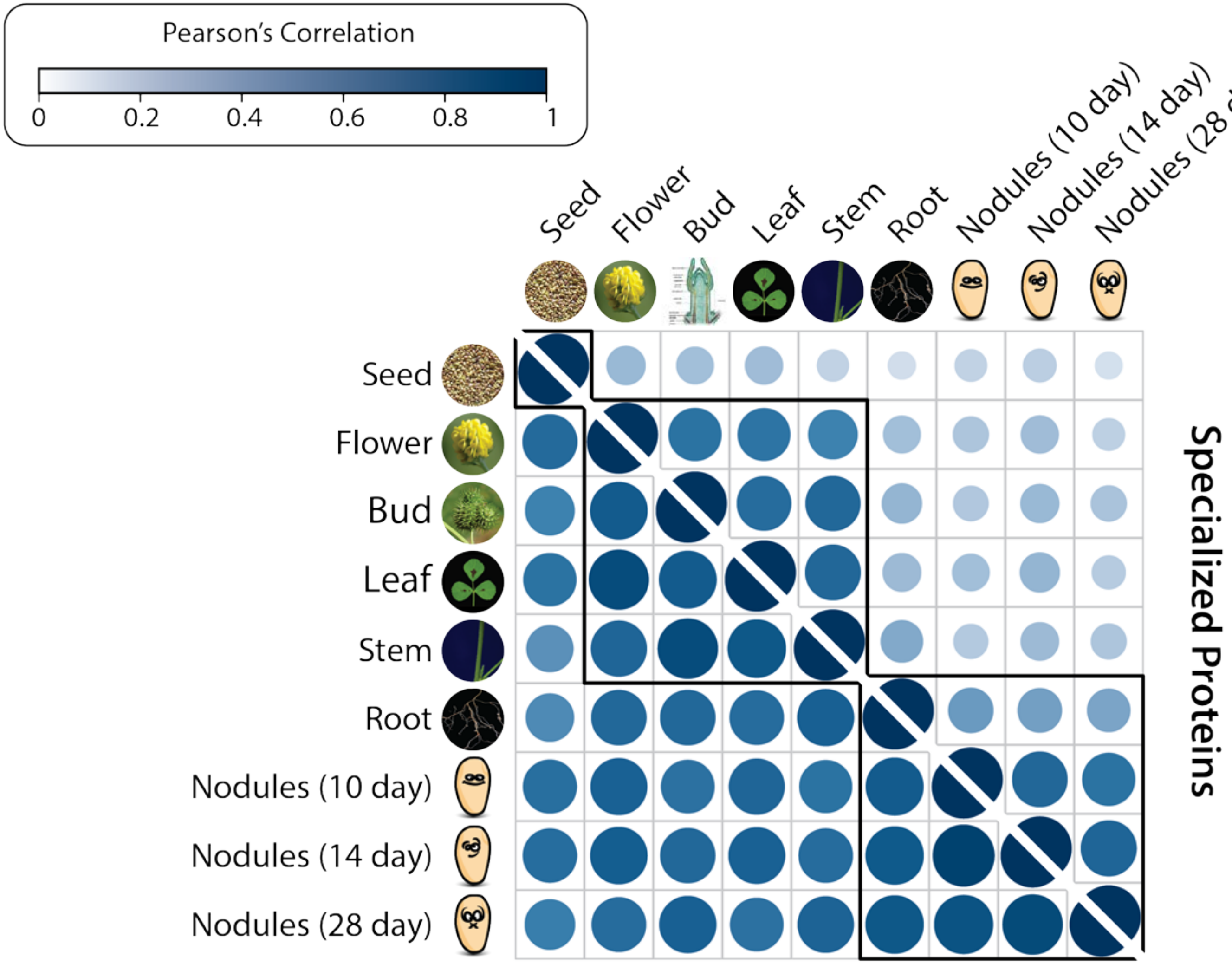
organs protein identified in
9 8 7 6 5 4 3 2 1
Nodules (Day 28)
Meristem (Bud)
Flower
Leaf
Root
Seed
Stem
Nodules (Day 10)
Nodules (Day 14)
0 1000 2000 3000 4000 5000 6000 7000 8000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000 12000 13000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000


organs protein identified in
9 8 7 6 5 4 3 2 1
relative protein abundance
Nodules (Day 28)
0 1000 2000 3000 4000 5000 6000 7000 8000
Meristem (Bud)
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000
Flower
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000 12000 13000
Leaf
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000
Root
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
Seed
0 1000 2000 3000 4000 5000 6000 7000 8000 9000
Stem
0 1000 2000 3000 4000 5000 6000 7000 8000 9000
Nodules (Day 10)
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
Nodules (Day 14)
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000


organs protein identified in
9 8 7 6 5 4 3 2 1
relative protein abundance
Organ specificity
Nodules (Day 28)
Meristem (Bud)
Flower
Leaf
Root
Seed
Stem
Nodules (Day 10)
Nodules (Day 14)

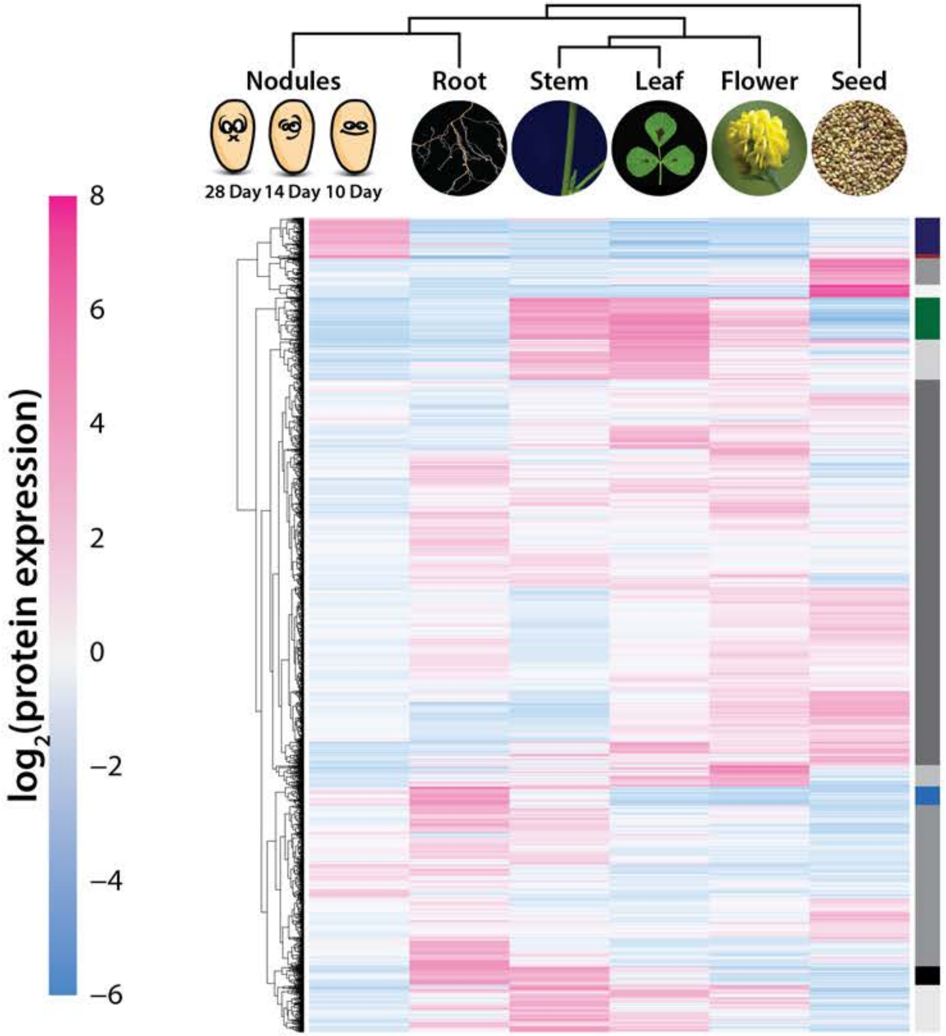

organs protein identified in
9
8
7
6
5
4
3
2
1
3
2
1
relative protein abundance
Organ specificity
Nodules (Day 28)
Meristem (Bud)
Flower
Leaf
Root
Seed
Stem
Nodules (Day 10)
Nodules (Day 14)
Day 28
Day 10
Day 14

Proteome comparison of the major organs of the *M. truncatula* plant

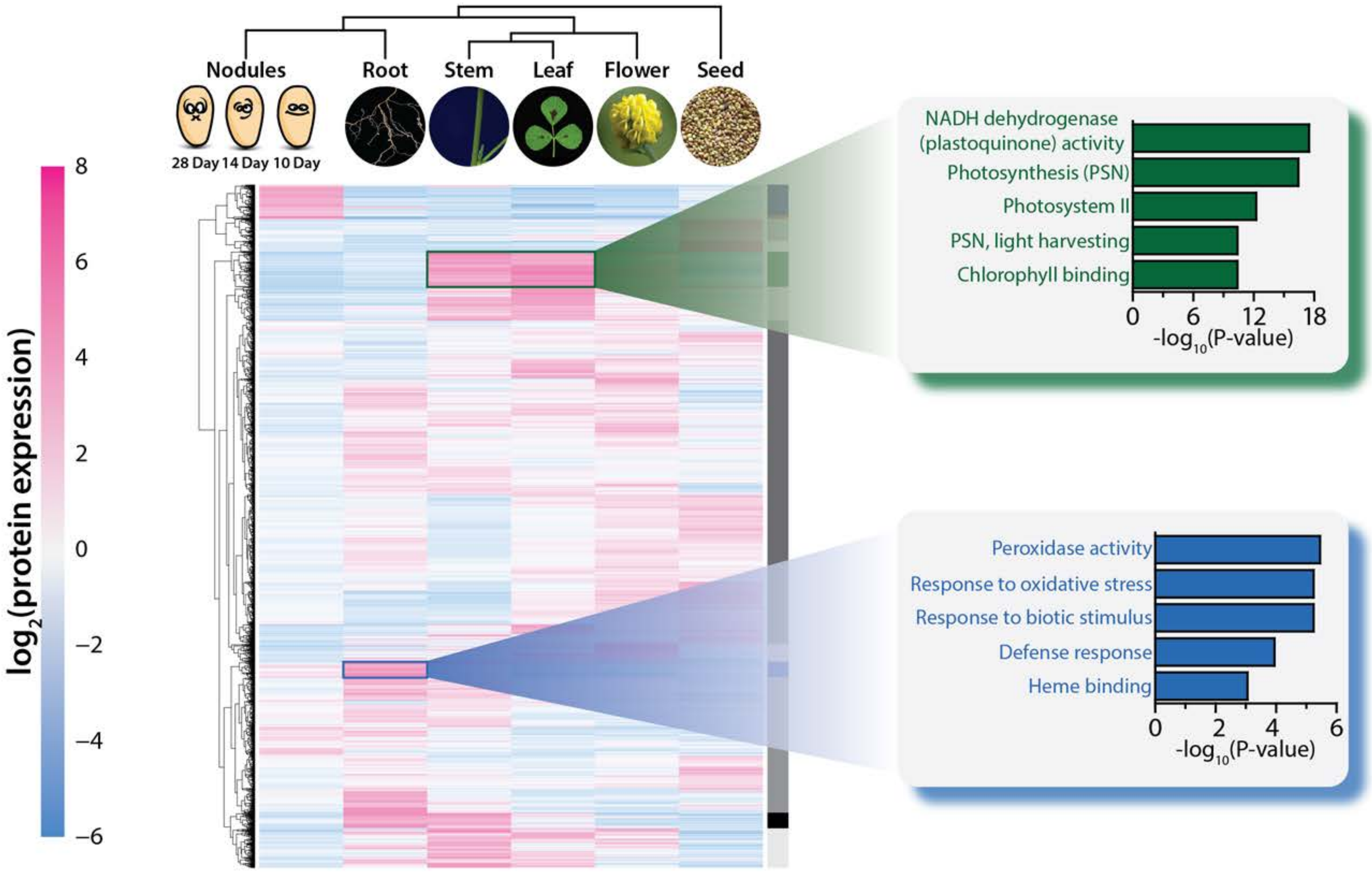


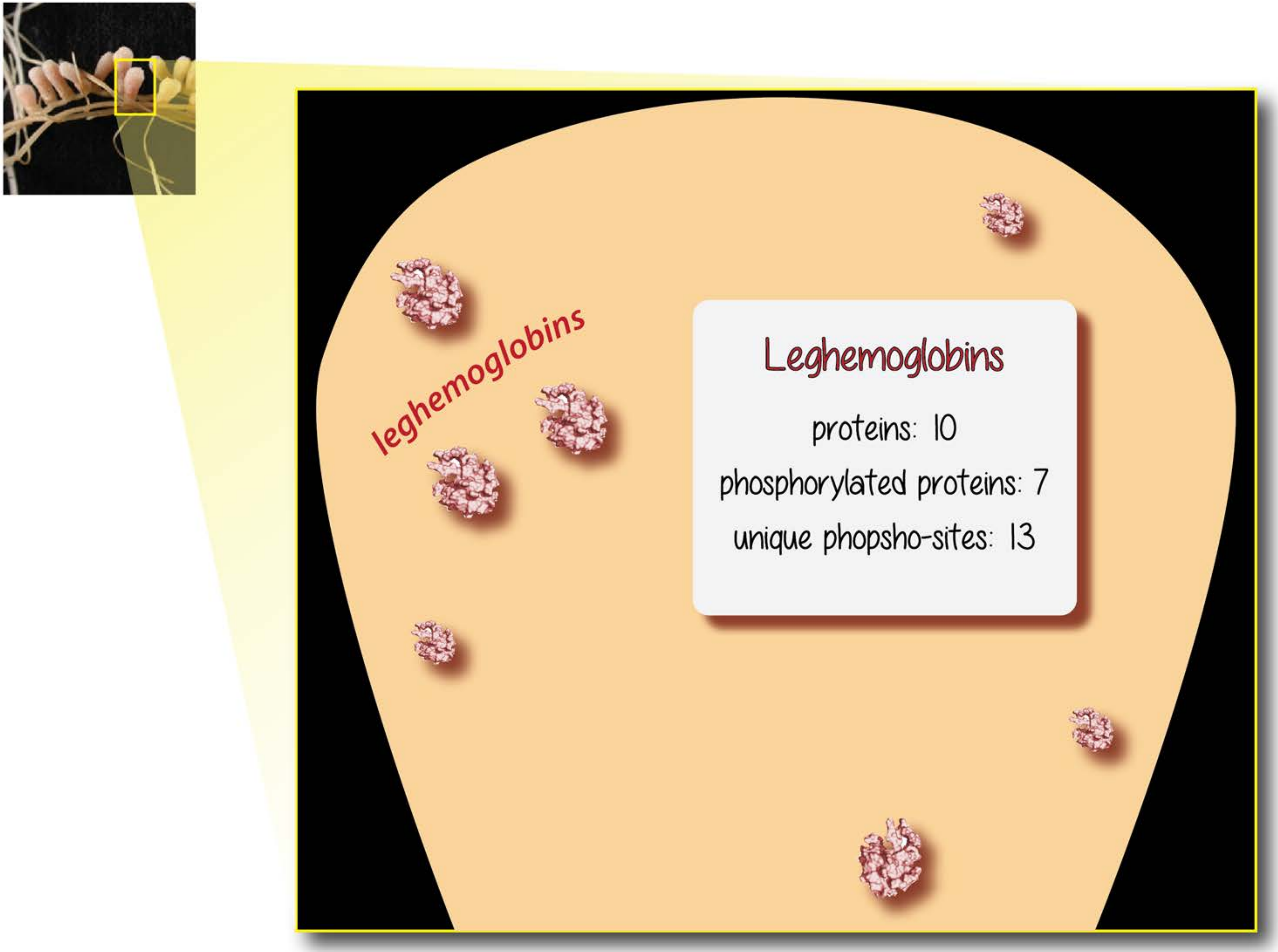
Core Proteins

Proteome comparison of the major organs of the *M. truncatula* plant



Proteome comparison of the major organs of the *M. truncatula* plant



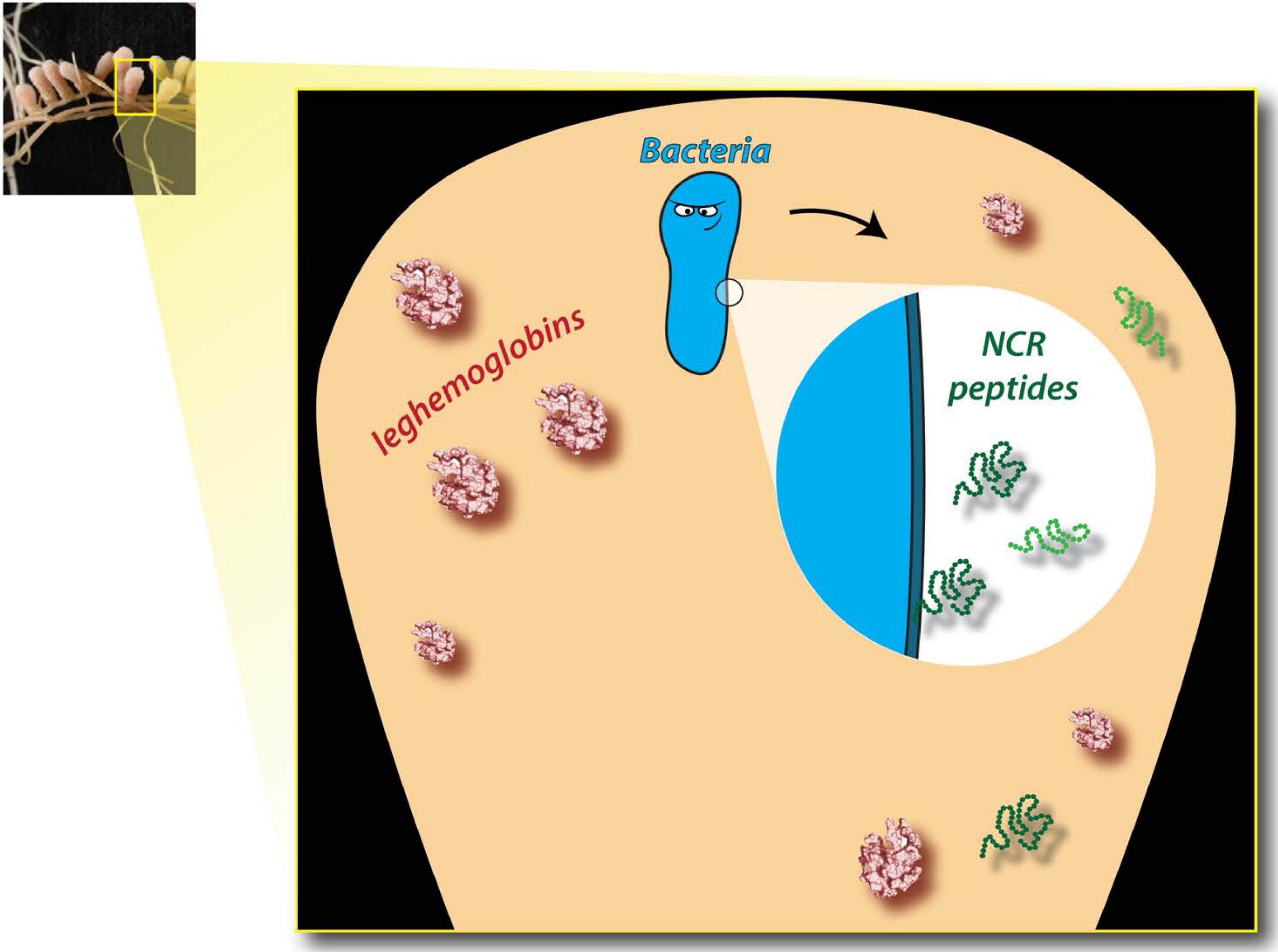


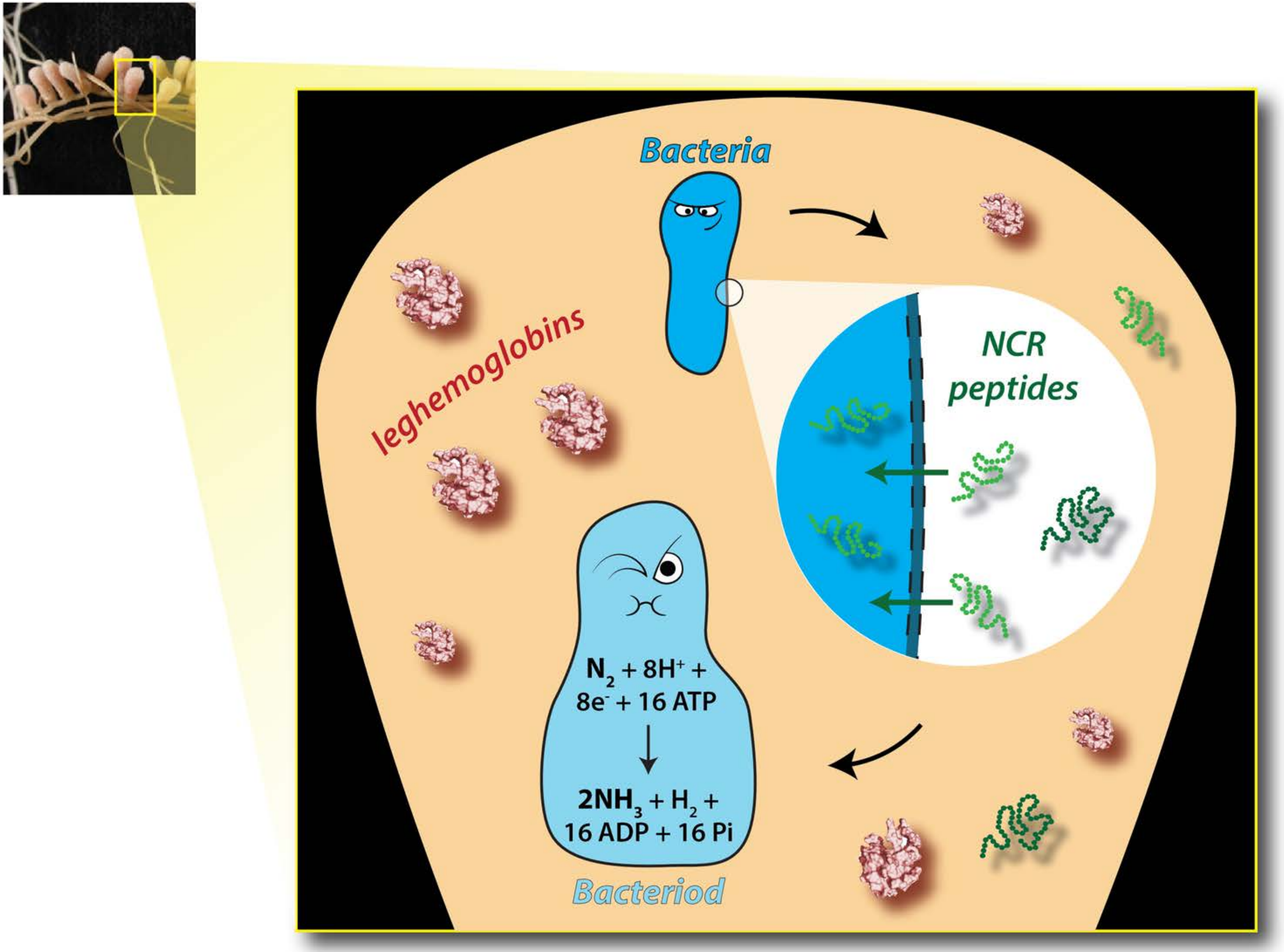
Leghemoglobins

proteins: 10

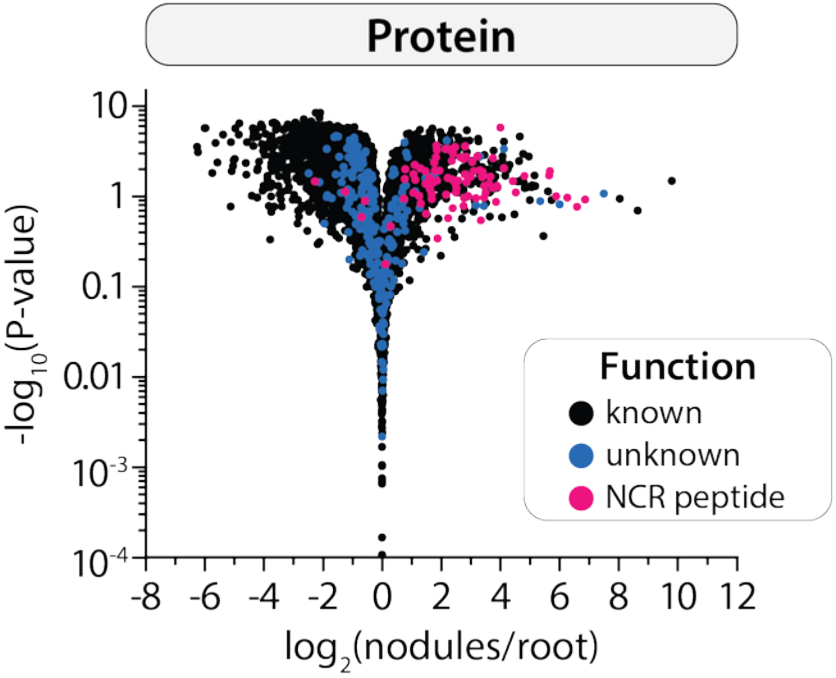
phosphorylated proteins: 7

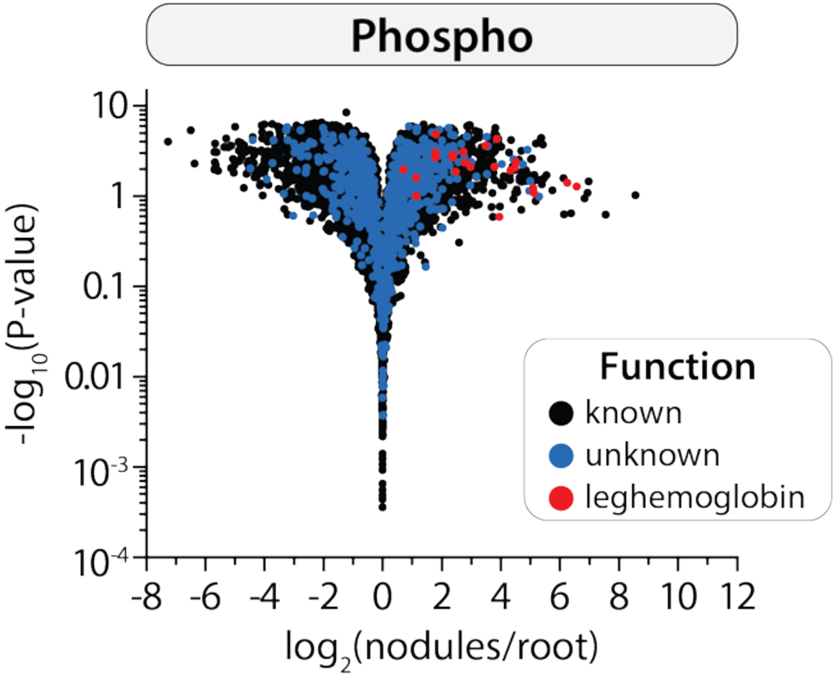
unique phopsho-sites: 13


Bacteria
leghemoglobins
NCR peptides

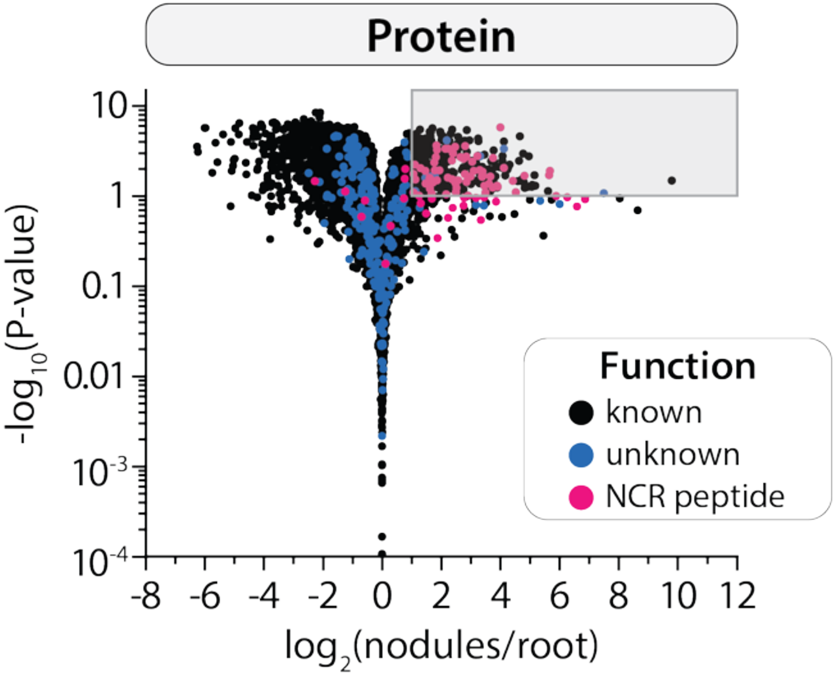

Bacteria
leghemoglobins
NCR peptides
N$_2$ + 8H$^+$ +
8e$^-$ + 16 ATP
2NH$_3$ + H$_2$ +
16 ADP + 16 Pi
Bacteriod

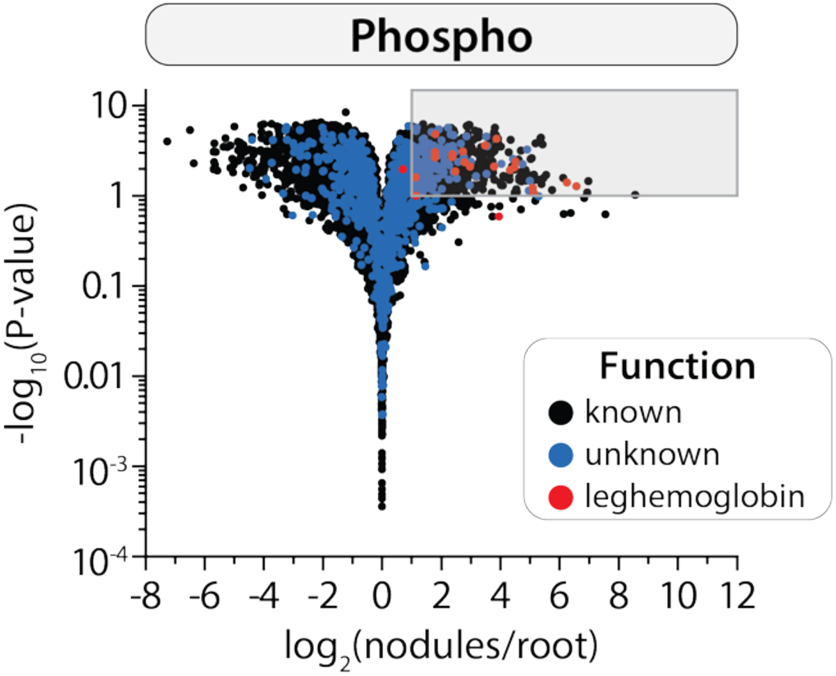
Nodule proteome gives insight into mechanisms behind symbiotic nitrogen fixation



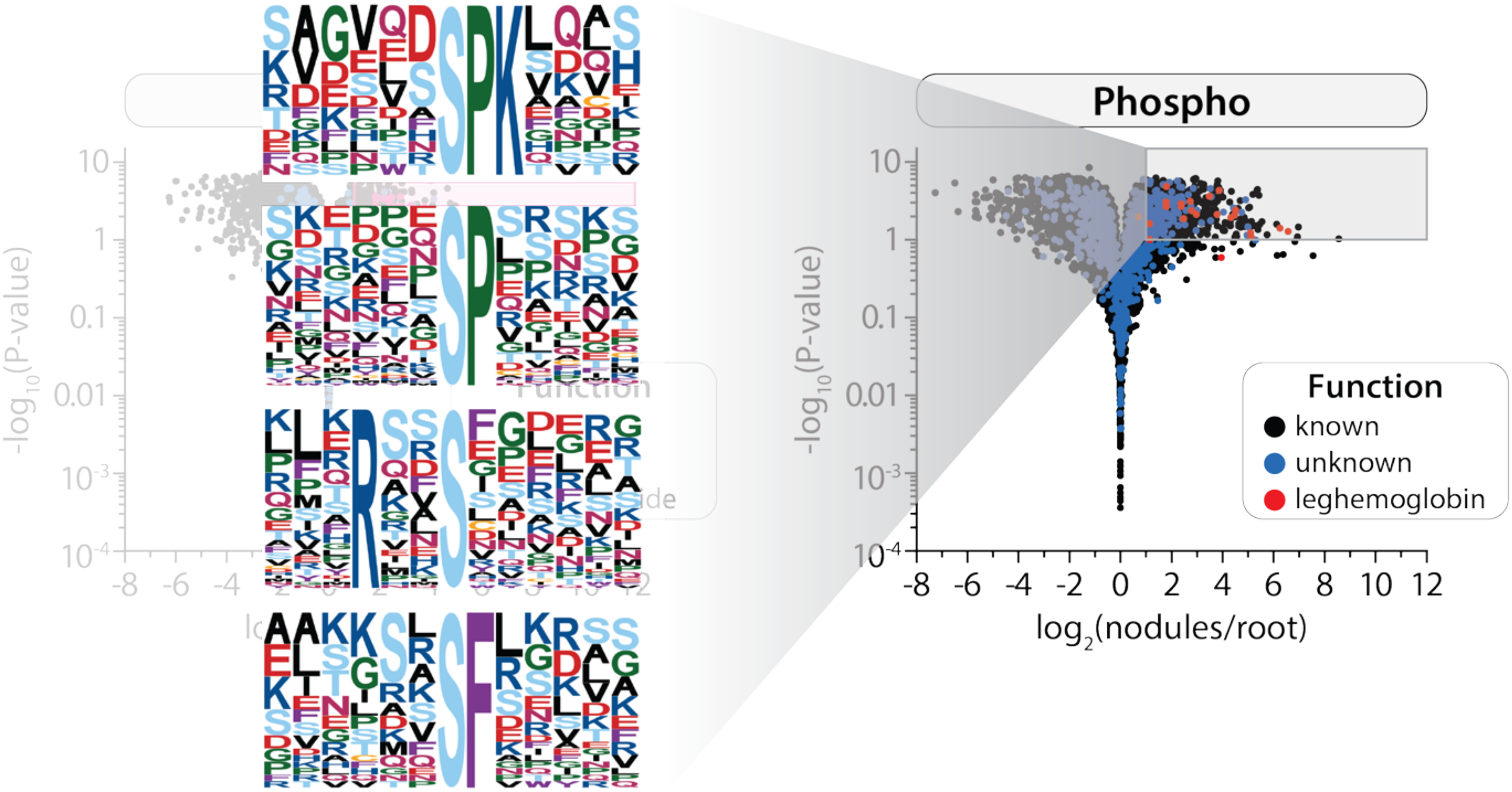


Nodule proteome gives insight into mechanisms behind symbiotic nitrogen fixation

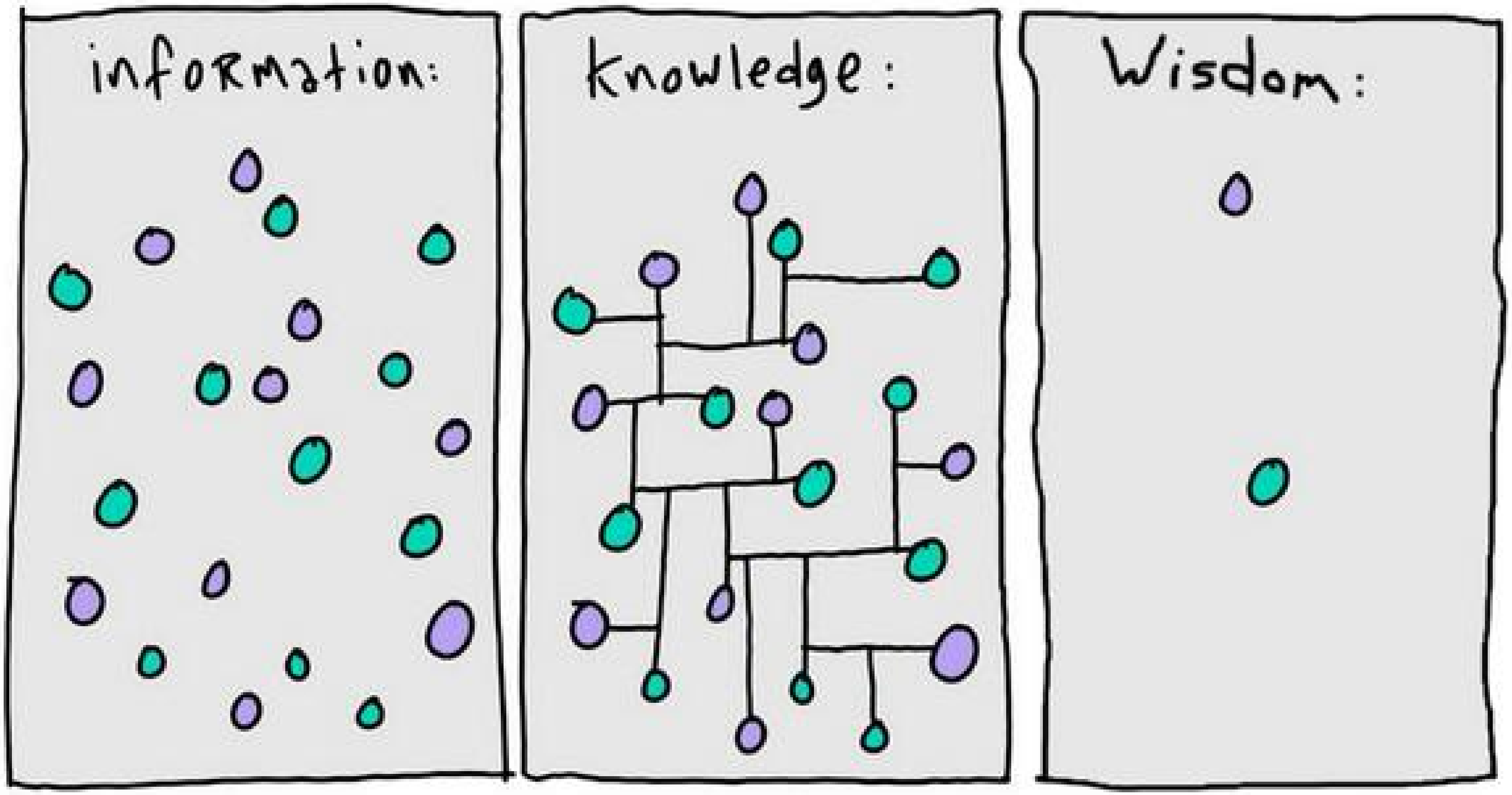




Nodule proteome gives insight into mechanisms behind symbiotic nitrogen fixation

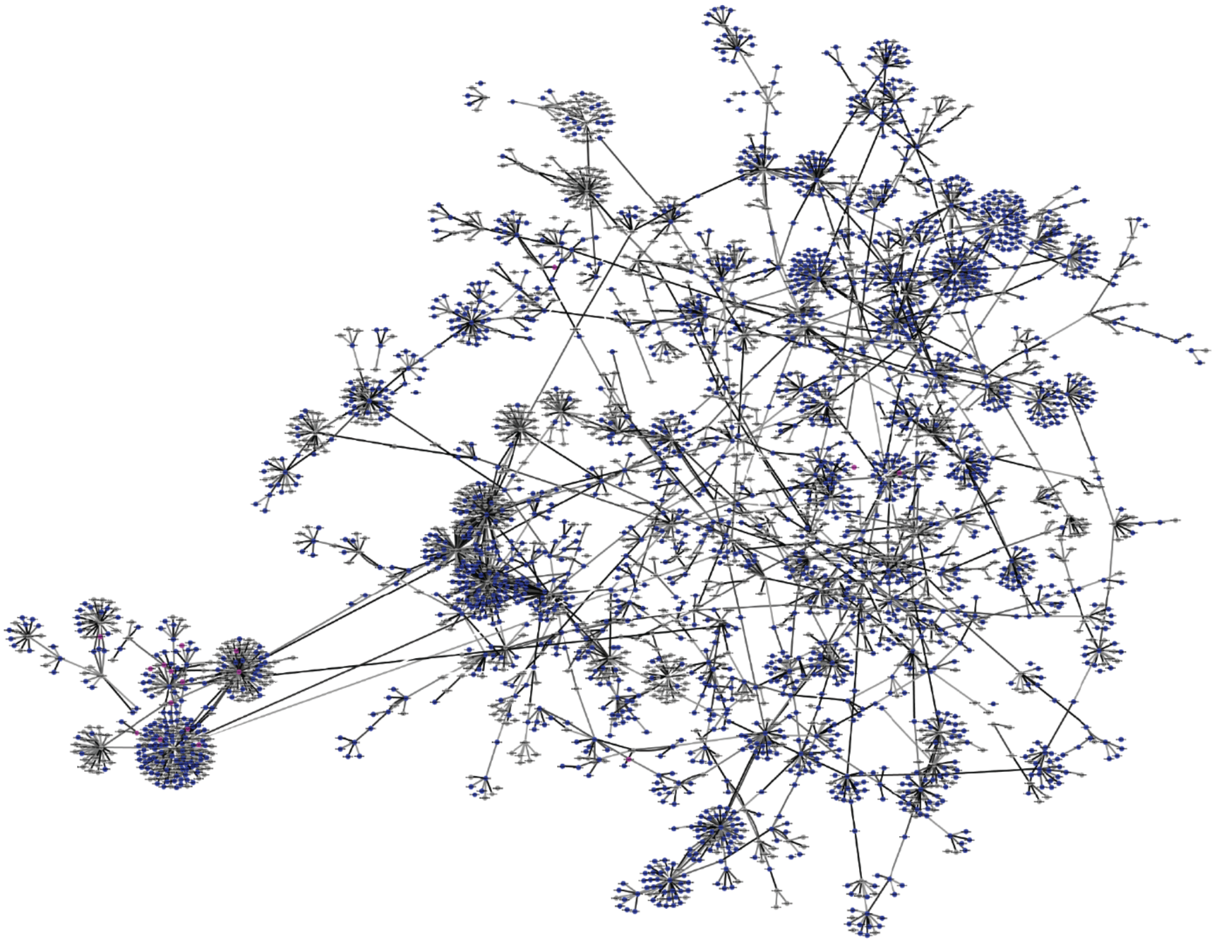


From Data to Wisdom

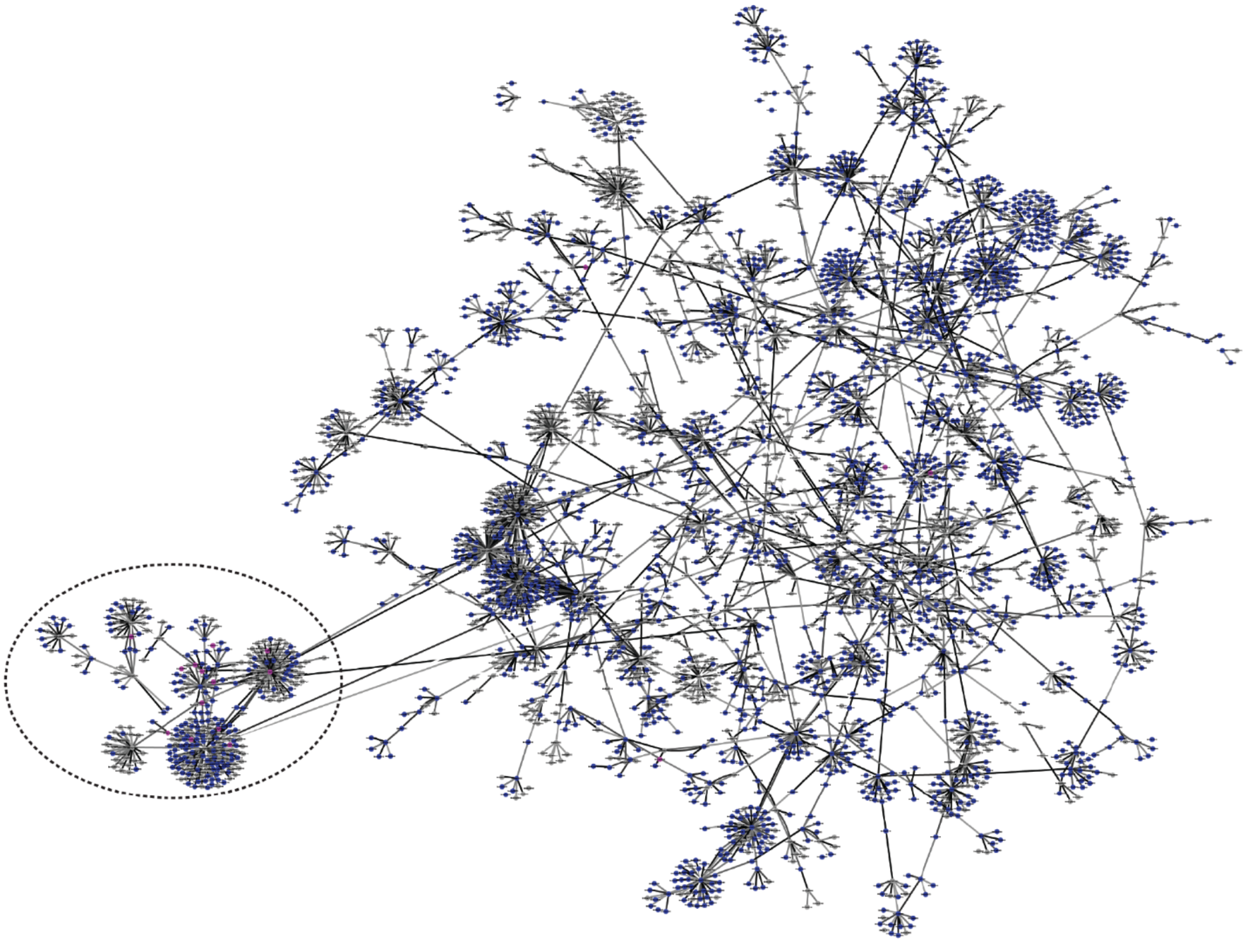


Ackoff, R. L., "From Data to Wisdom", Journal of Applied Systems Analysis, Volume 16, 1989 p 3-9

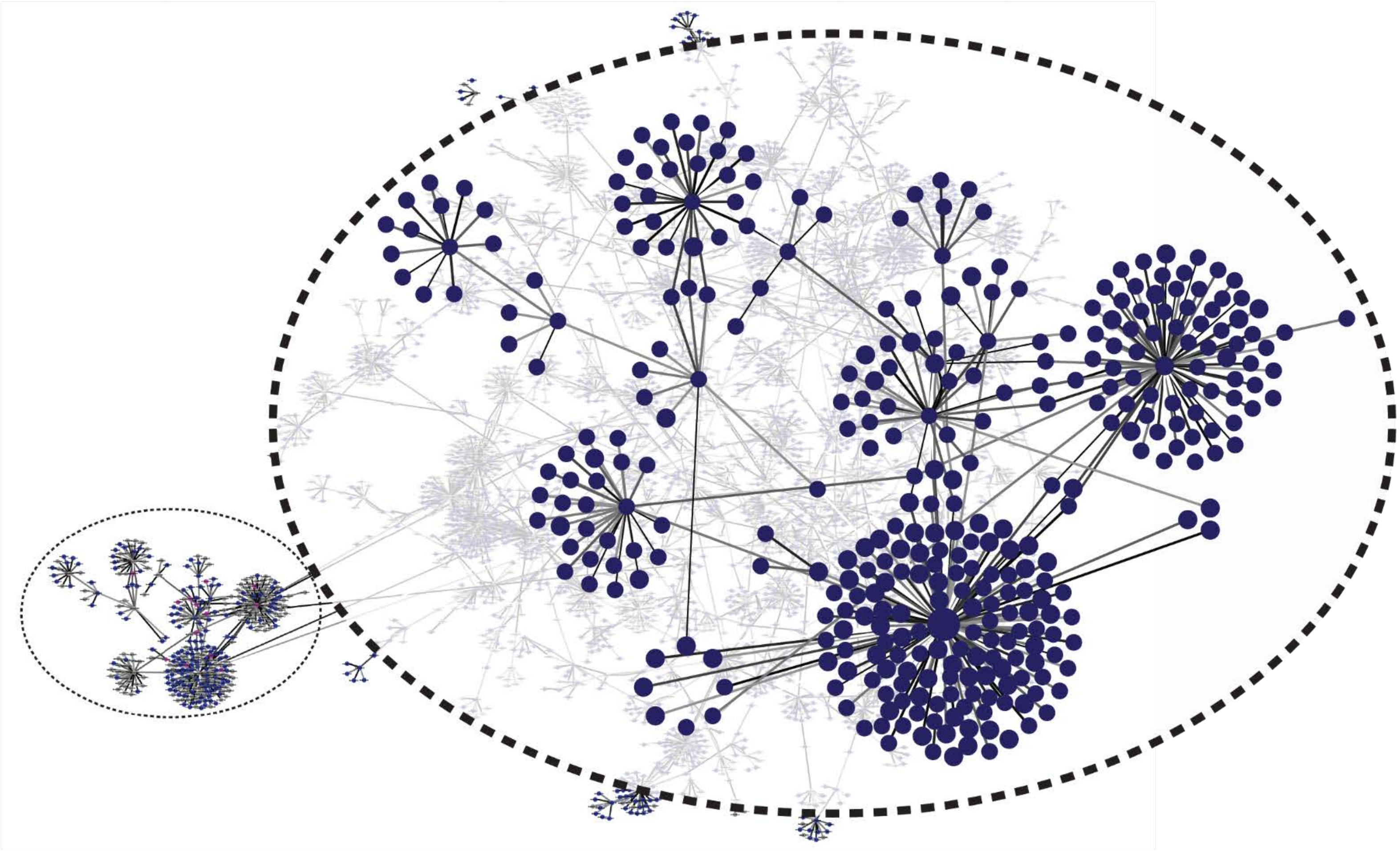
Inferring protein and phospho-protein regulation in the nodules using transcript co-expression



Inferring protein and phospho-protein regulation in the nodules using transcript co-expression

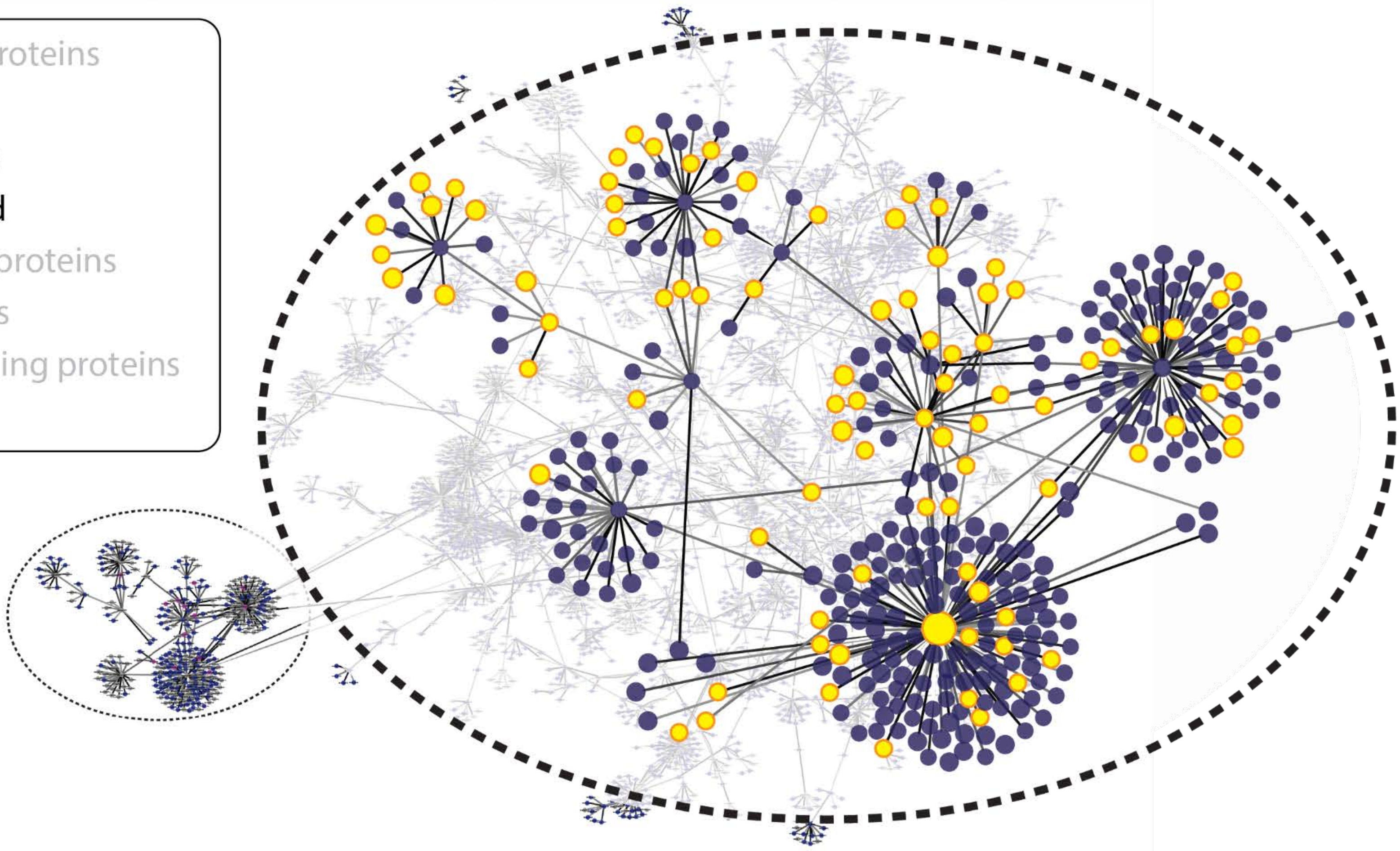


Inferring protein and phospho-protein regulation in the nodules using transcript co-expression

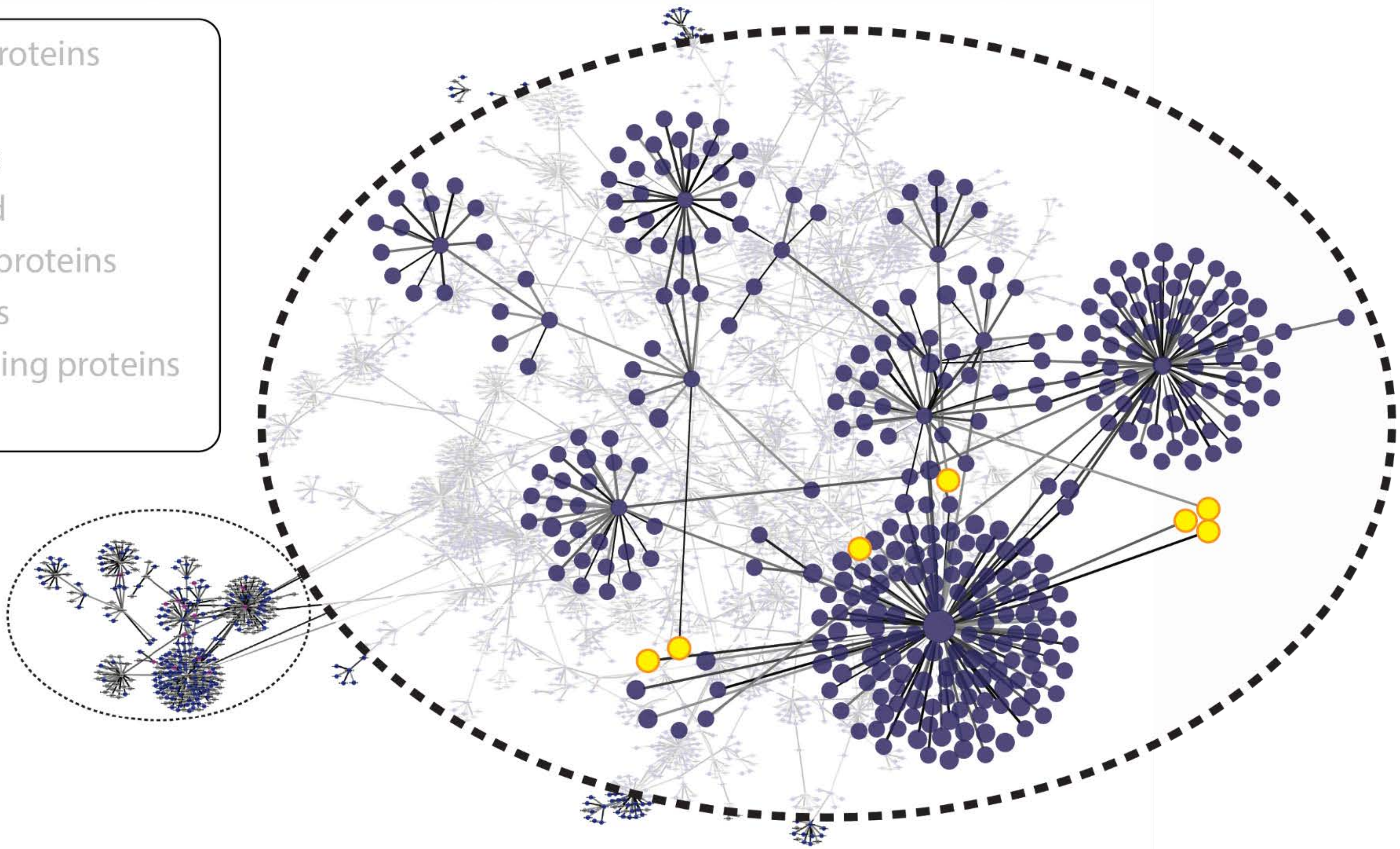


Inferring protein and phospho-protein regulation in the nodules using transcript co-expression

Phospho motif proteins
NCR Peptides
Leghemoglobins
Protein identified
Phosphorylated proteins
Acetyled proteins
Calmodulin-binding proteins
Kinases



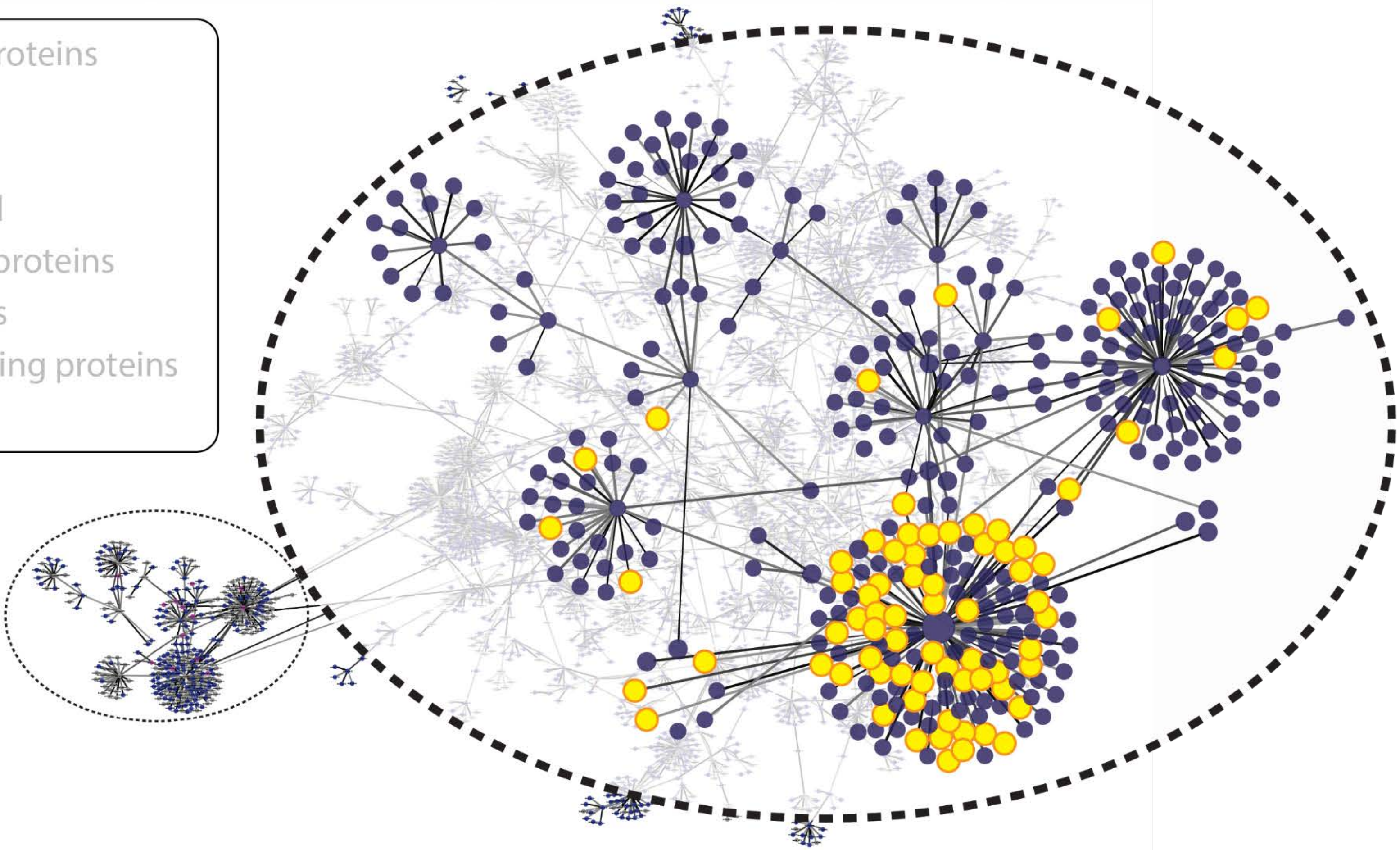
Inferring protein and phospho-protein regulation in the nodules using transcript co-expression



Phospho motif proteins
NCR Peptides
Leghemoglobins
Protein identified
Phosphorylated proteins
Acetyled proteins
Calmodulin-binding proteins
Kinases

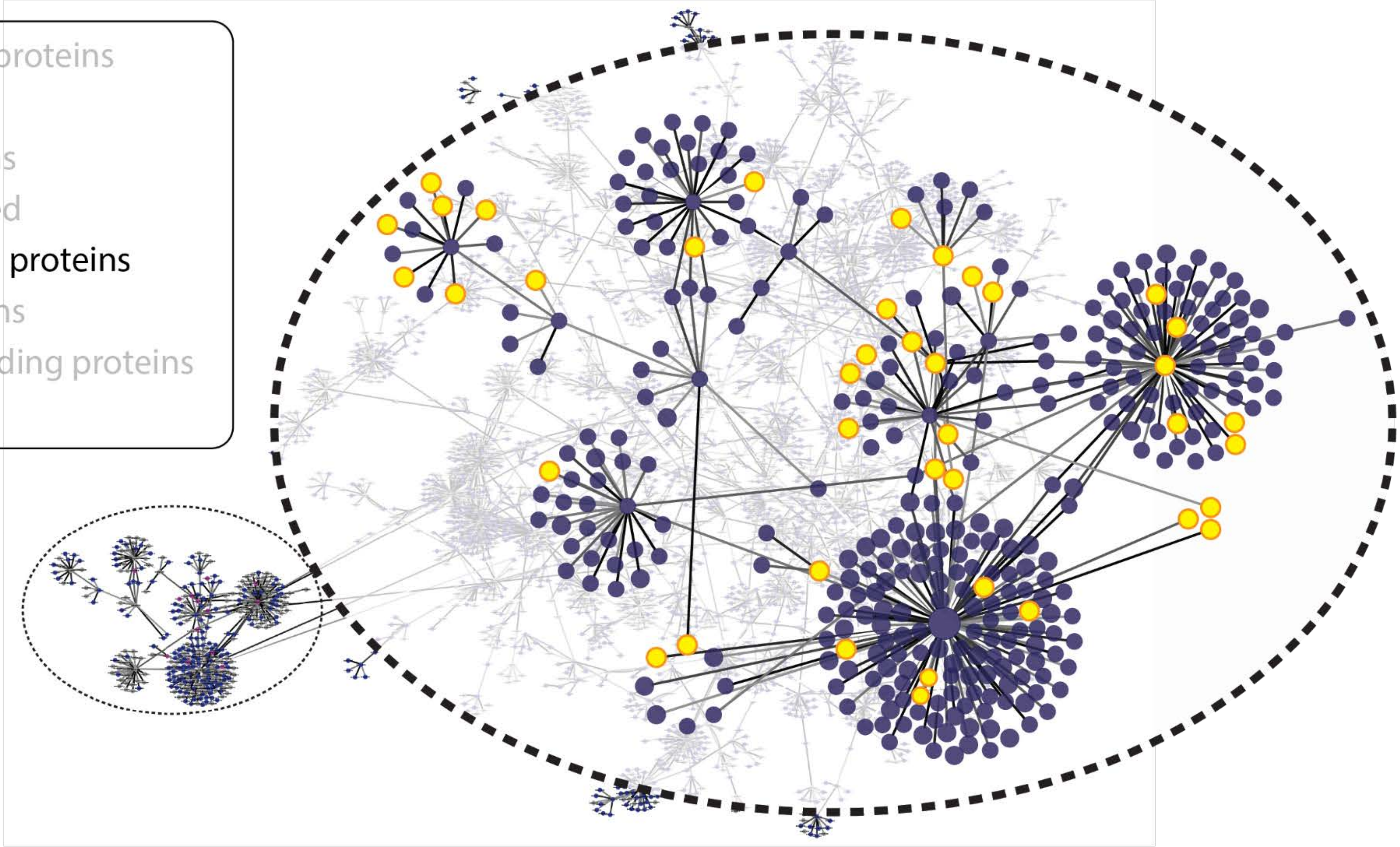
Inferring protein and phospho-protein regulation in the nodules using transcript co-expression

Phospho motif proteins
NCR Peptides
Leghemoglobins
Protein identified
Phosphorylated proteins
Acetyled proteins
Calmodulin-binding proteins
Kinases

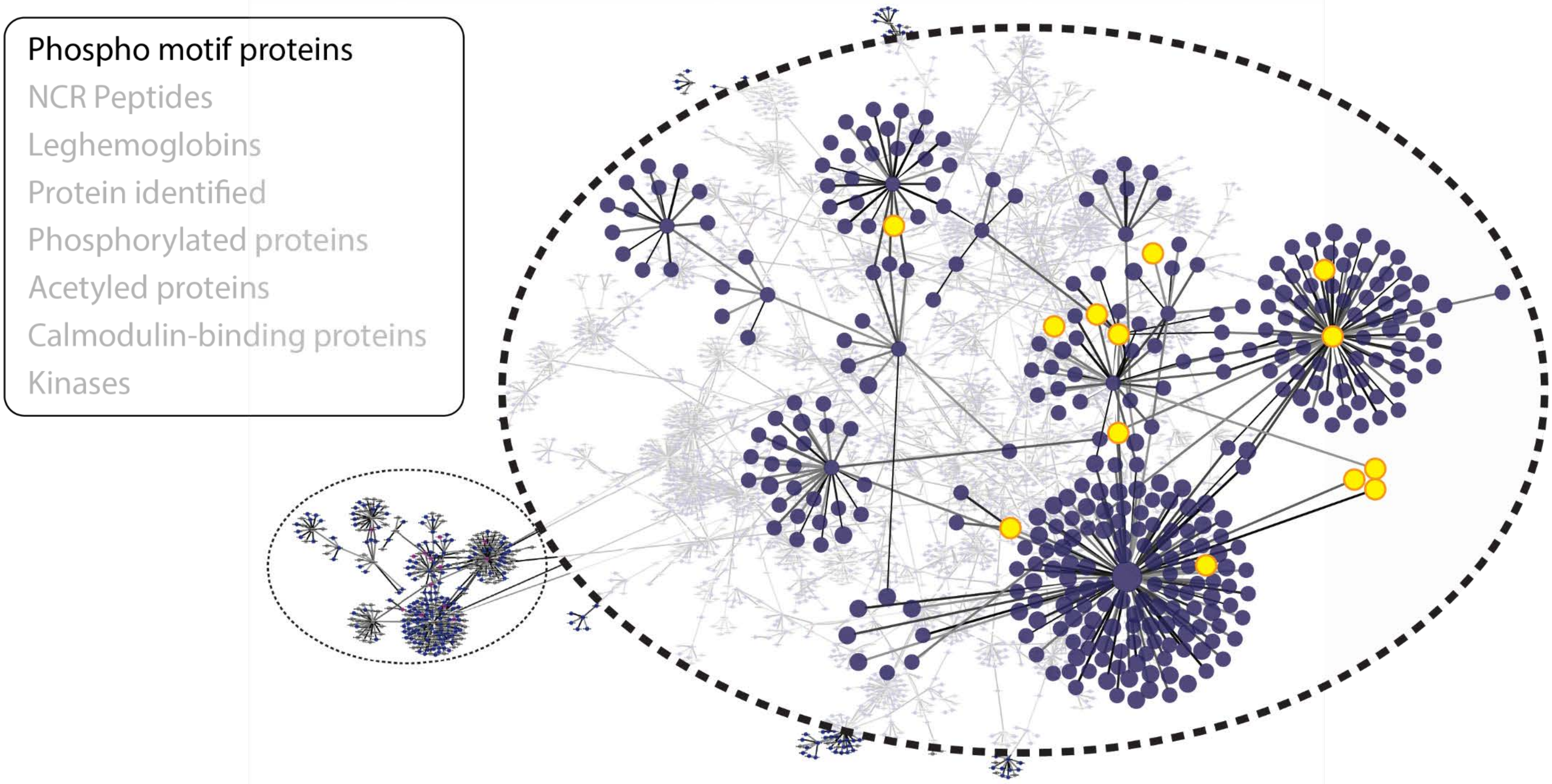


Inferring protein and phospho-protein regulation in the nodules using transcript co-expression

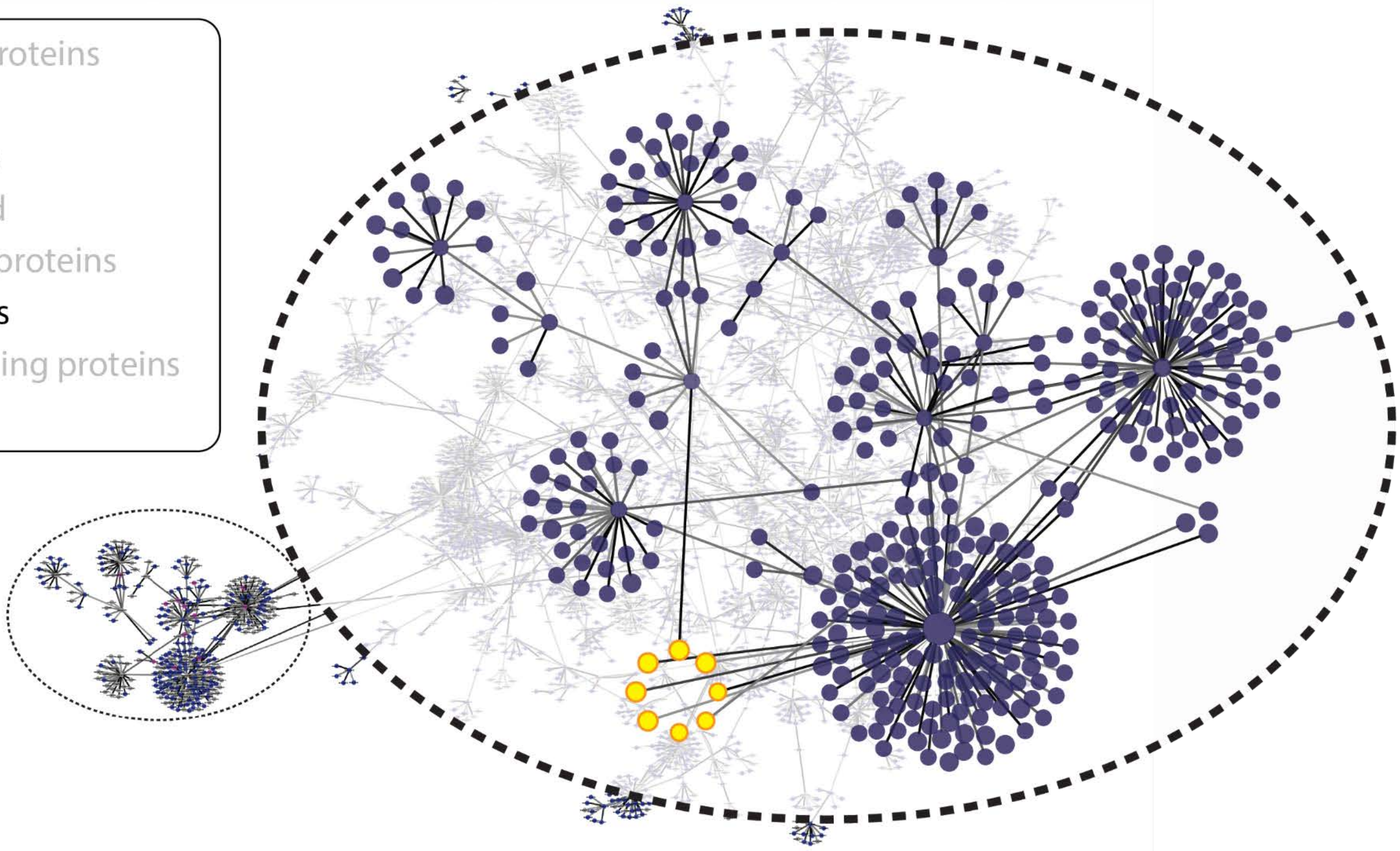
Phospho motif proteins
NCR Peptides
Leghemoglobins
Protein identified
Phosphorylated proteins
Acetyled proteins
Calmodulin-binding proteins
Kinases



Inferring protein and phospho-protein regulation in the nodules using transcript co-expression

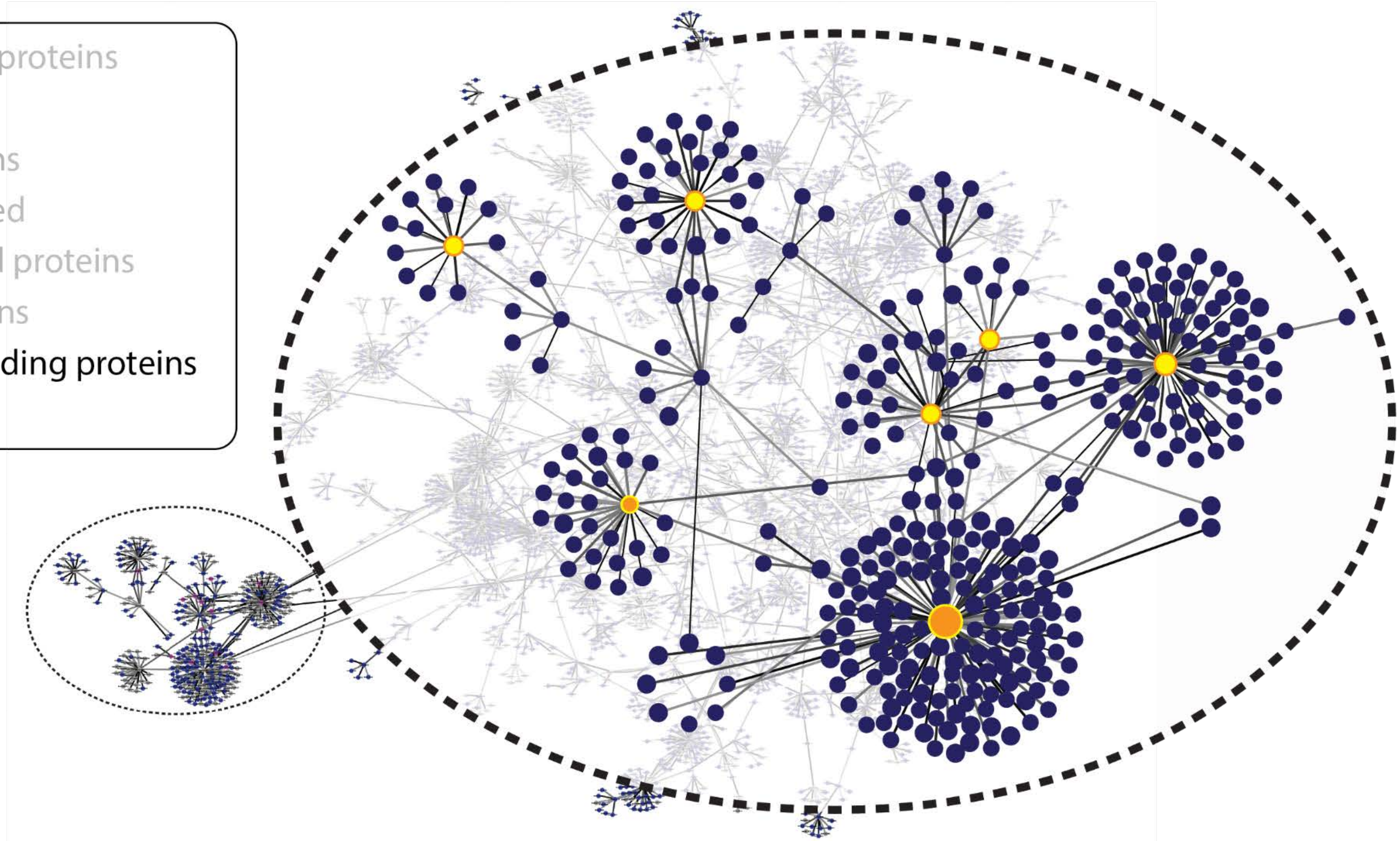


Inferring protein and phospho-protein regulation in the nodules using transcript co-expression



- Phospho motif proteins
- NCR Peptides
- Leghemoglobins
- Protein identified
- Phosphorylated proteins
- Acetyled proteins
- Calmodulin-binding proteins
- Kinases

Inferring protein and phospho-protein regulation in the nodules using transcript co-expression



Phospho motif proteins
NCR Peptides
Leghemoglobins
Protein identified
Phosphorylated proteins
Acetyled proteins
Calmodulin-binding proteins
Kinases

Future Directions

Make this resource available to the community
http://www.medicago.wisc.edu/compendium/

Experimental validation of hypotheses
(genetics, genome editing)

Conservation in leguminous crops
(evolutionary analyses)

Wisconsin Medicago Group

Jean-Michel Ané

Dhileep Jayaraman
Shanmugam Rajasekar
Junko Maeda
Kevin Garcia
Angel R. Del Valle-Echevarria

Joshua J. Coon

Harald Marx
Catherine E. Minogue
Alicia L. Richards
Nicholas W. Kwiecien
Michael S. Westphall

Michael R. Sussman

Jeremy Volkening

Sushmita Roy

Alireza F. Sihapirani

NSF IOS # 0701846