An RNA-Seq based gene expression atlas of the common bean

O'Rourke *et al.*

BMC
Genomics

RESEARCH ARTICLE   **Open Access**

An RNA-Seq based gene expression atlas of the common bean

Jamie A O'Rourke[1,7*], Luis P Iniguez[2], Fengli Fu[1], Bruna Bucciarelli[1,3], Susan S Miller[1,3], Scott A Jackson[4], Philip E McClean[5], Jun Li[6], Xinbin Dai[6], Patrick X Zhao[6], Georgina Hernandez[2] and Carroll P Vance[1]

Abstract

Background: Common bean (*Phaseolus vulgaris*) is grown throughout the world and comprises roughly 50% of the grain legumes consumed worldwide. Despite this, genetic resources for common beans have been lacking. Next generation sequencing, has facilitated our investigation of the gene expression profiles associated with biologically important traits in common bean. An increased understanding of gene expression in common bean will improve our understanding of gene expression patterns in other legume species.

Results: Combining recently developed genomic resources for *Phaseolus vulgaris,* including predicted gene calls, with RNA-Seq technology, we measured the gene expression patterns from 24 samples collected from seven tissues at developmentally important stages and from three nitrogen treatments. Gene expression patterns throughout the plant were analyzed to better understand changes due to nodulation, seed development, and nitrogen utilization. We have identified 11,010 genes differentially expressed with a fold change ≥ 2 and a P-value < 0.05 between different tissues at the same time point, 15,752 genes differentially expressed within a tissue due to changes in development, and 2,315 genes expressed only in a single tissue. These analyses identified 2,970 genes with expression patterns that appear to be directly dependent on the source of available nitrogen. Finally, we have assembled this data in a publicly available database, The *Phaseolus vulgaris* Gene Expression Atlas (*Pv* GEA), http://plantgrn.noble.org/PvGEA/ . Using the website, researchers can query gene expression profiles of their gene of interest, search for genes expressed in different tissues, or download the dataset in a tabular form.

Conclusions: These data provide the basis for a gene expression atlas, which will facilitate functional genomic studies in common bean. Analysis of this dataset has identified genes important in regulating seed composition and has increased our understanding of nodulation and impact of the nitrogen source on assimilation and distribution throughout the plant.

Keywords: *Phaseolus vulgaris* cv Negro jamapa, Common bean, RNA-Seq, Symbiotic nitrogen fixation, Expression atlas, SRP046307

Background

Common bean (*Phaseolus vulgaris* L.), *Pv,* is an important source of proteins, micronutrients and calories for over three hundred million people worldwide, mostly throughout Latin America and Africa where beans are an important component of traditional diets. The high levels of dietary protein (between 20 and 25%) and micronutrients

in beans complement the high carbohydrates found in maize and cassava [1]. In addition to their important contribution to human health, legumes are also important contributors to biological nitrogen (N). N is a primary nutrient limiting plant production [2], with the acquisition and assimilation of N second only to photosynthesis for plant growth and development [1]. Despite the international importance of *Pv*, both in terms of economics and consumption, it has lagged behind *Medicago truncatula, Lotus japonicus,* soybean, and other legumes in terms of genetic resources. cDNA libraries have been used to investigate phosphate stress, resistance to bean rust, and leaf development [3-7]. Sequence information for *Pv* was greatly

* Correspondence: jamie.orourke@ars.usda.gov
[1]Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA
[7]Current Address: USDA-ARS, Corn Insect Crop Genetics Research Unit, Iowa State University, Ames, IA 50011, USA
Full list of author information is available at the end of the article

enhanced by using Roche 454 technology coupled with mRNA sequences to assemble 59,295 unigene sequences, [8], though these data are not yet publicly available. Most recently, the genome sequence and predicted gene calls for *Pv* G 19833 has been made publicly available (www. phytozome.net). This resource provides a platform for *Pv* genomic and comparative genomic analyses [9]. Sequence conservation and genetic colinearity between *Pv* and soybean (*Glycine max* L. merr) [10,11] which diverged from a common ancestor approximately 19 million years ago [12,13], allows genomic information to be leveraged from one species to the other.

In this study we utilized RNA-seq to characterize expression profiles for the transcriptome of common bean (*Pv* cv. Negro Jamapa). Gene expression profiles were analyzed from 24 unique samples from seven distinct tissues; roots, nodules, stems, flowers, leaves, pods, and seeds throughout development. Our data was used as the foundation for The *Pv* Gene Expression Atlas (*Pv* GEA) database, available at http://plantgrn.noble.org/ PvGEA/. We utilized the expression profiles of all predicted genes in *Pv* to examine the biological processes related to seed and pod development, nodulation and symbiosis, and changes in gene expression due to nitrogen availability.

Results and discussion
Phaseolus vulgaris gene expression atlas (Pv GEA)
To facilitate additional use of the RNA-Seq data generated for these analyses, we have developed a web-accessible database, The *Pv* Gene Expression Atlas (*Pv* GEA), available at http://plantgrn.noble.org/PvGEA/. This database was built using a similar database structure, web application, architecture and tools as the LegumeIP platform [14] to retrieve and visualize the gene expression patterns using RNA-seq data. To facilitate the mining of the data included in *Pv* GEA, we have provided the capability to: (i) visualize expression profiles of genes of interest, (ii) identify genes exhibiting certain expression patterns in specific tissues, (iii) identify genes and gene expression patterns based on www.phytozome.net annotation terms; and (iv) download the entire data set, either raw or normalized, in tabular form to facilitate the analysis of more complicated biological questions. Using the predicted gene calls of the G 19833 *Pv* genome to build the *Pv* GEA database means it can be easily expanded to integrate RNA-Seq data from future experiments. Currently, *Pv* GEA includes gene expression profiles from 24 samples isolated from roots, root nodules, stems, leaves, flowers, pods, and seeds at various developmental stages under ideal growth conditions. Included in this dataset are transcripts from eight samples including nodule, root, and leaf tissues for plants having either fix + or fix- root nodules; providing preliminary data

on the impact of nodulation and N fixation on gene expression, an important biological process for legumes.

The 26,964 transcriptionally active genes identified in our data (RPKM ≥ 3 in at least one tissue) represent 78% of the 31,638 predicted genes in *Pv*; confirming that the tissues, time points, and treatments used in this study (as described in Table 1) affected a majority of the genes in the genome and provide an excellent foundation of gene expression in future experimental comparisons. Pair-wise analyses identified differentially expressed genes between both tissues and samples (Table 2 and Additional file 1, respectively). These comparisons identified 11,010 genes differentially expressed between tissues. Additionally, we identified genes differentially expressed within tissues from samples collected at developmentally important time-points (between seeds from three developmental stages (DS): 2,694, four pod DS: 13,125, five leaf DS: 5,401, six root DS: 1,458, and three nodule DS: 1,551). Finally, we identified genes exhibiting either tissue specific expression, (Additional file 2 and Additional file 3a), or sample specific expression, (Additional file 3b and Additional file 4).

To ensure our transcriptome analysis was reliable we visualized the expression profile of the purine and ureide biosynthesis pathways (Additional file 5) and conducted qPCR on 85 additional genes (for details see experimental procedures and Additional file 6). In warm season legumes, such as soybean and *Pv*, purine biosynthesis is known to be highly up regulated in nodules [15]. Our data is consistent with this as the genes in this pathway are highly up regulated in nitrogen fixing nodules compared to all other tissues (Additional file 5). The enzyme uricase (Additional file 3c) degrades ureate to allantoin, which is the main supply of N for plant nutrition. In our *Pv* data allantoinase, the enzyme responsible for allantoin degradation, is highly expressed early in seed and pod development, likely providing N to developing seeds (Additional file 3d). Expression of uricase and allantoinase in aerial tissues suggests ureides are degraded after being transported from the nodules. Leaves, seeds, and pods can then utilize the released NH_3 and CO_2 in a variety of cellular processes. These results are consistent with reports of high ureide levels observed in developing *Pv* seeds [16] and high allantoinase enzyme activity throughout pod development and seed filling as measured by Thomas *et al.* [17]. Additionally, Pellissier et al. [18] reported allantoin transporters highly expressed in developing pods and seed coats of *Pv*. The results of these three studies combined with the gene expression patterns observed in our data highlight the importance of ureide metabolism in aerial tissues to provide N for developing tissues.

Gene expression analysis
Genes exhibiting tissue specific expression (Additional file 2 and Additional file 3a) are involved in a variety of gene

Table 1 Tissue samples isolated from *Phaseolus vulgaris* cv. Negro jamapa for RNA-Seq analysis

Organ	Sample ID[a]	Sample description	Reads sequenced[b]	Reads mapped[c]	Genes expressed[d]
Leaves	YL	Fully expanded 2nd trifoliate leaf tissue from plants provided with fertilizer	24,118,479	21,395,546	19,466
	L5	Leaf tissue collected 5 DAI with effective rhizobium	25,297,092	19,762,749	18,966
	LF	Leaf tissue from fertilized plants collected at the same time of LE and LI	22,692,275	17,581,928	16,103
	LE	Leaf tissue collected 21 DAI with effective *rhizobium*	23,366,279	17,714,097	16,081
	LI	Leaf tissue collected 21 DAI with ineffective *rhizobium*	23,257,968	18,820,812	18,523
Stem	YS	All stem internodes above the cotyledon collected at the 2nd trifoliate stage	27,696,970	22,718,724	20,299
	ST	Shoot tip, including the apical meristem, collected at the 2nd trifoliate stage	25,826,838	22,358,410	21,142
Flower	FY	Young flowers, collected prior to floral emergence	23,334,037	16,503,930	20,055
Pods	PY	Young pods, collected 1 to 4 days after floral senescence. Samples contain developing embryos at globular stage	26,234,498	17,381,695	12,115
	PH	Pods approximately 9 cm long, associated with seeds at heart stage (pod only)	24,986,174	19,130,051	19,065
	P1	Pods between 10 and 11 cm long, associated with stage 1 seeds (pod only)	24,349,622	16,400,591	15,831
	P2	Pods between 12 and 13 cm long associated with stage 2 seeds (pod only)	21,647,774	18,018,224	18,311
Seeds	SH	Heart stage seeds, between 3 and 4 mm across and ~7 mg	28,222,798	22,972,385	18,668
	S1	Stage 1 seeds, between 6 and 7 mm across and ~50 mg	21,395,296	17,004,466	16,949
	S2	Stage 2 seeds, between 8 and 10 mm across and 140–150 mg	24,696,630	19,949,204	15,363
Roots	RT	Root tips, 0.5 cm of tissue, collected from fertilized plants at 2nd trifoliate stage of development	24,536,948	21,680,391	18,514
	YR	Whole roots, including root tips, collected at the 2nd trifoliate stage of development	25,140,904	20,962,342	19,170
	R5	Whole roots separated from 5 day old pre-fixing nodules	27,423,246	23,034,248	19,865
	RF	Whole roots from fertilized plants collected at the same time as RE and RI	25,121,968	21,858951	20,305
	RE	Whole roots separated from fix + nodules collected 21 DAI	24,449,104	21,325,105	20,450
	RI	Whole roots separated from fix- nodules collected 21 DAI	27,834,770	24,188,904	20,697
Nodules	N5	Pre-fixing (fix+) nodules collected 5 DAI	23,909,973	20,774,095	19,102
	NE	Fix + nodules collected 21 DAI	26,350,845	23,035,381	17,011
	NI	Fix- nodules collected 21 DAI	24,875,317	21,877,973	19,278

[a]A two-letter ID assigned to each sample. [b]Number of single end 36 bp reads generated for each sample. [c]Number of RNA-Seq reads mapping to the genome using Bowtie. [d]Number of predicted genes expressed with an RPKM ≥ 3 in each sample.

ontology biological processes. Genes uniquely expressed in leaves are involved in amino acid phosphorylation, DNA and protein binding. Genes uniquely expressed in seed tissues include processes such as carbohydrate metabolism, as exemplified by a starch branching enzyme

Table 2 Differentially expressed genes between tissue types

	Seed	Pod	Stem	Leaf	Root	Nodule
Seed	0	909	2,521	446	1,690	840
Pod	665	0	227	844	504	455
Stem	1,299	946	0	1,847	602	773
Leaf	1,354	1,111	690	0	1,112	1,084
Root	1,986	1,867	731	3,003	0	679
Nodule	1,554	936	526	1,789	375	0

The number of genes in each cell represents genes up regulated in the column tissue compared to the row tissue. For the comparison Seeds vs Pods 1,574 genes are differentially expressed; 909 up regulated in pods and 665 up-regulated in seeds.

(SBE), which is important in amylopectin synthesis, a carbohydrate precursor [19]. In our data, *Pv SBE* expression (Phvul.005G040300.1) is highest in developing seeds (Additional file 3e). These results are consistent with the high carbohydrate composition of *Pv* seeds reported by Broughton et al. [1]. Nodule specific transcripts annotate as involved in oxidoreductase activity, amino acid phosphorylation and membrane transport/signaling are highly and uniquely expressed; reflecting the importance of nutrient transport to the nodule and the high energy cost of N fixation. In root tissues, genes involved in pectinecterase, carbohydrate metabolism, iron ion binding, oligopeptide membrane transport, and lipid metabolisms are uniquely expressed. Expression of these genes illustrates the role of the root in nutrient acquisition and the importance of root growth for plant health.

The 6,667 known transcription factors (TF) in soybean (downloaded from SoyDB) [20] were compared to *Pv* genes using TBLASTN (e-value 1e-30). This analysis identified

3,726 putative TFs in the *Pv* genome, representing 52 of the 64 transcription factor families in soybean (Additional file 7). The 3,726 TFs identified in *Pv* is almost exactly half the number identified in soybean, as expected since soybean has undergone a whole genome duplication event not experienced by *Pv* [12,13]. The average expression of TFs in all seed tissues was much lower than that of other tissues, including developing pods. Fisher tests confirmed 26 TF families exhibited statistically enhanced or repressed expression by tissue type (Figure 1). Twenty-one TF families exhibited altered expression patterns in a single tissue.

Five other TF families (AS2, AUX, NAC, SBP, and ZD-HD) showed expression patterns that were statistically significant in multiple tissues.

Seed development and metabolism

Pv seeds are an integral component of diets around the world. Unlike soybean seeds, which are valued for high oil and protein content, *Pv* beans provide high levels of protein and carbohydrates making it a highly nutritious food for human consumption. The seed and pod samples represent an extended time-course collection of the
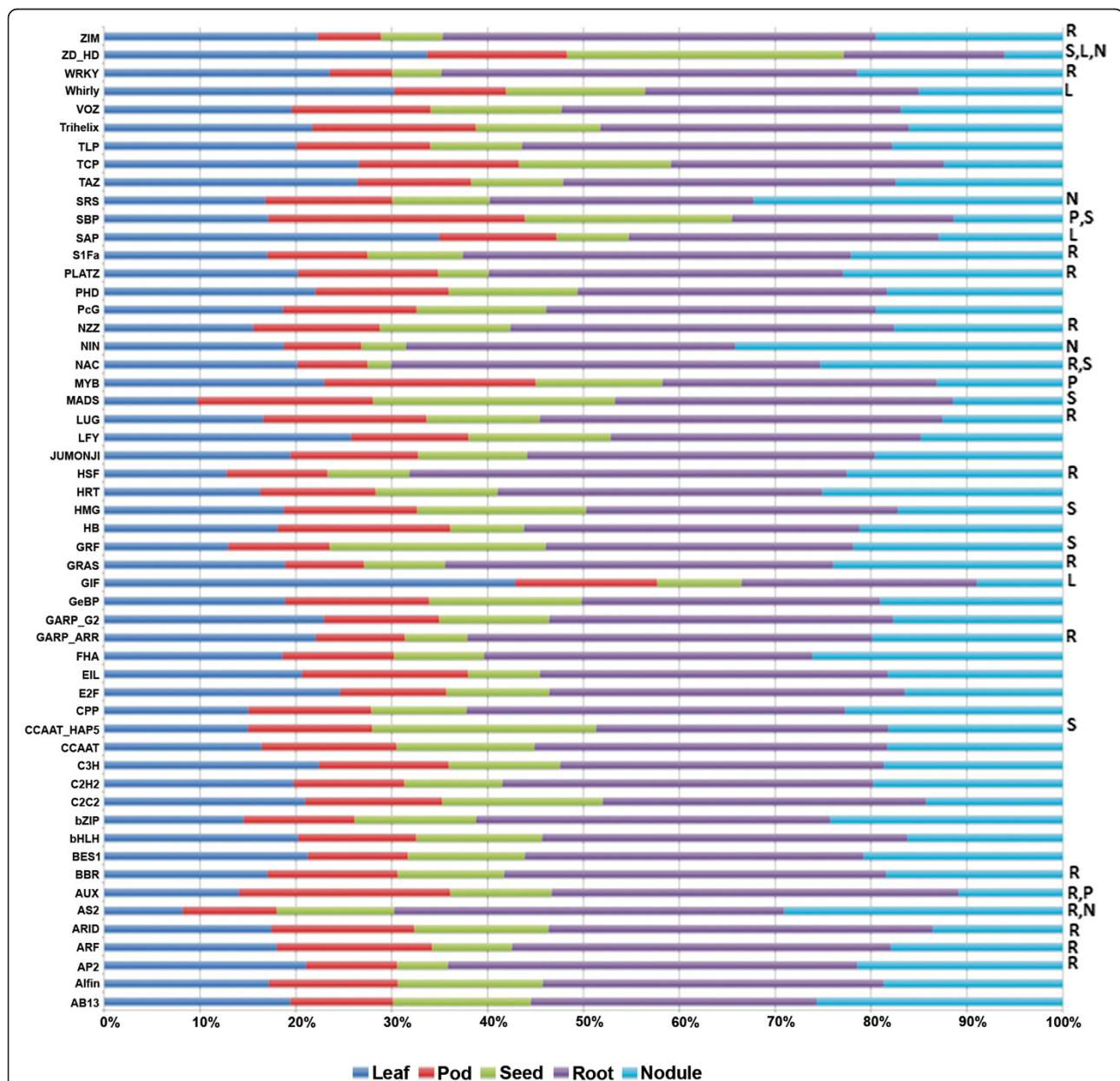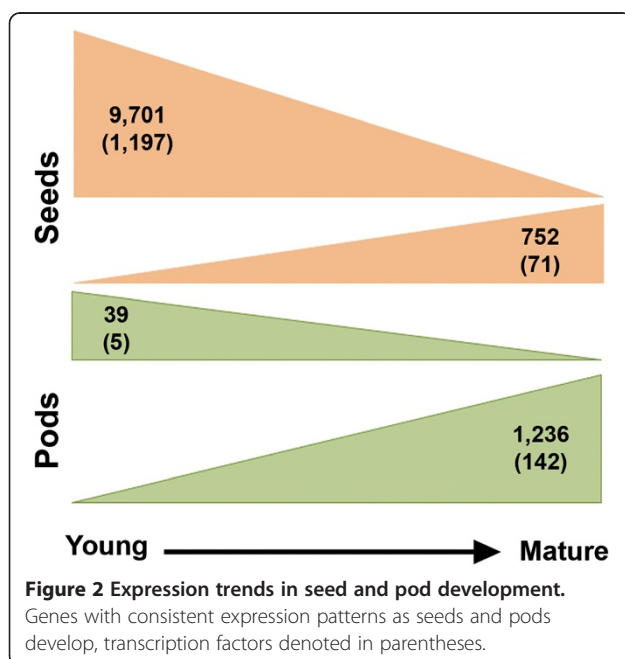

Figure 1 Transcription factor family expression profile by tissue. Fisher's test identified 26 transcription factor families with higher or lower than expected expression in a specific tissue (leaf, blue; pod, red; seed, green; root, purple; nodule, teal). Tissues with statistically significant gene expression patterns are denoted to the right of the graph; Leaf, L; Pod, P; Seed, S; Root, R; Nodule, N.

same tissue spanning great developmental changes. Approximately half (16,292) of the 31,638 predicted genes are expressed in *Pv* seeds with 12,182 genes expressed in all three stages of seed development examined. In *Pv* pods 17,248 genes are expressed in at least one of the four developmental stages examined. We identified genes differentially expressed between seeds and pods at the three stages of development (8,189 genes; Additional file 8, Additional file 9, and Additional file 10), and genes differentially expressed within seeds (2,694) and pods (13,125) throughout development (Additional file 11 and Additional file 12). Additionally, we identified 9,702 genes with consistently decreasing expression levels as the seed develops, including 1,196 TFs. By comparison, 753 genes were identified with increasing expression levels as the seed develops (Figure 2) including 70 TFs from 25 families, including HB (Phvul.007G064100.1), MYB (Phvul.001G025200.1), and ARF (Phvul.011G080100.1) (Additional file 3f). Members of three of these TF families (HB, MYB, and ARF) are seed specific in Arabidopsis, *M. truncatula*, and our *Pv* data [21,22]. Additionally, members of these TF families are among the TFs identified by Le *et al.* [23] as differentially expressed between structures of developing *Phaseolus coccineus* (scarlet runner bean) seeds that are also differentially expressed in *Pv* seeds at different developmental stages. All these results are consistent with a study by Hajduch *et al.* [24], which determined the expression of proteins involved in primary and secondary metabolism, cell growth and division, signal transduction, and protein synthesis all decrease as the seeds develop. Conversely, in the pods we identified 39 genes with expression levels

steadily decreasing and 1,236 genes with consistently increasing expression patterns (Figure 2).

Soybean and *Pv*, although closely related species, have distinct seed compositions. While soybean accumulates oil and protein, *Pv* accumulates carbohydrates and protein [1,25]. Comparing the 500 most highly expressed genes in both *Pv* and soybean seeds [26] allowed us to identify genes important for general seed development (Figure 3a). Genes involved in carbohydrate biosynthesis are highly up regulated in *Pv* seeds while genes involved in fatty acid biosynthesis are highly up regulated in developing soybean seeds (Figure 3b and c). The starch synthase (STS) genes, particularly Phvul.001G082500.1, are highly expressed in *Pv* seeds (RPKMs: S1 = 240, S2 = 286) but not in soybean seeds (RPKM < 15) (Figure 3b and c). STS is required for amylopectin biosynthesis, a component of carbohydrates. Similarly, sucrose acts as a key regulator of seed carbon flux; high sucrose synthase (SS) activity in developing seeds may channel available carbon towards carbohydrate biosynthesis and away from fatty acid biosynthesis [27]. In our *Pv* RNA-Seq data, genes encoding SS are highly expressed in developing seeds (Figure 3b). In Arabidopsis, SS loss of function mutants favor fatty acid and protein biosynthesis over starch biosynthesis in seed development, resulting in a 55% increase of fatty acids and a near 70% reduction in starch content of mature seeds [28]. The role of SS in regulating carbohydrate synthesis is further supported by the low expression in developing soybean seeds (Figure 3c) (Severin et al. [26]), which are valued for oil and protein. The synthesis of polyunsaturated fatty acids is regulated by FATTY ACID DESATURASE 2 (FAD2) [29]. In soybeans, *FAD2* is highly expressed in developing seeds while in *Pv FAD2* is expressed early in seed and pod development, but at a much lower level (Figure 3b and c).

Seed development in multiple species is regulated by four master TFs: LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), ABSCISIC ACID INSENSITIVE3 (ABI3), and WRINKLED 1 (WRI1*)* [21,22,30]. Using BLASTP, we queried the Arabidopsis protein sequences to identify homologous sequences in the *Pv* predicted genes (Figure 4). The homolog for *LEC2* was only weakly expressed (RPKM = 4) mid-seed development in *Pv*. Seeds of Arabidopsis loss of function *lec2* mutants accumulated 15% less protein and 30% less oil while the seed starch content increased five fold as compared to wild type plants [31]. The altered seed composition of *lec2* mutant plants closely resembles that of *Pv*, suggesting down regulation of *LEC2* may affect seed composition. *LEC2* controls the gene expression of WRI1, which also exhibits low expression patterns in *Pv* developing seeds (RPKM: SH = 5, S1 = 9, S2 = 5). *WRI1* expression modulates the expression of a set of genes controlling late glycolysis and
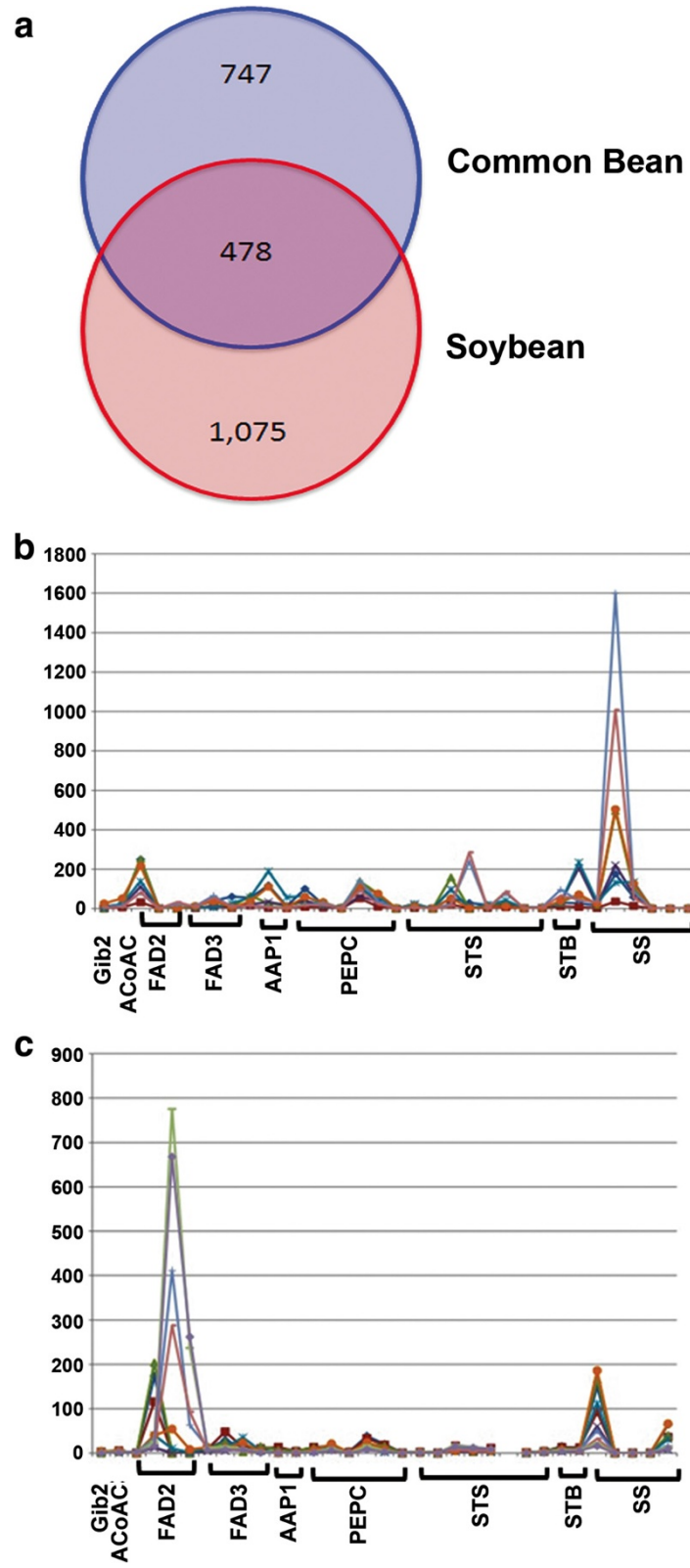

Figure 2 Expression trends in seed and pod development.
Genes with consistent expression patterns as seeds and pods develop, transcription factors denoted in parentheses.

Figure 3 (See legend on next page.)

(See figure on previous page.)

Figure 3 Comparison of soybean and common bean seeds. (a) Comparing the top 1,500 expressed genes (regardless of seed stage) in soybean (as reported by Severin *et al.* [26]) and common bean seeds. **(b and c)** Expression profiles of genes involved in fatty acid and starch biosynthesis pathways in developing seed tissues. Glb2 (GLABARA 2), ACoAC (Acetyl CoA Carboxylase), FAD2 and FAD3 (fatty acid desaturase), AAP1 and 2 (amino acid transporter), PEPC (phosphoenolpyruvate carboxylase), STS (starch synthase), STB (starch branching enzyme), SS (sucrose synthase). **(b)** Gene expression profiles in common bean, **(c)** Gene expression profiles in soybean (as reported by Severin *et al.* [26]).

fatty acid biosynthesis. The low expression of both *LEC2* and *WRI1* may relate to the lower oil composition of *Pv*.

Abscisic acid (ABA) is a key hormone in seed development, important in developing desiccation tolerance and entrance into dormancy [30]. ABA accumulation in seeds is both temporally and spatially regulated [27]. We found high expression of ABA biosynthesis genes (Phvul.002G018700.1 and Phvul.005G031500.1) in developing seeds, with expression decreasing as the seeds matured (Additional file 3g). ABA biosynthesis is regulated by 9-cis-EPOXYCAROTENOID DIOXYGENASE (NCED9) [32]. The expression pattern of *NCED9* (Additional file 3 h) and ABA biosynthesis genes

(Additional file 3 g) in *Pv* developing seeds is consistent with those from developing seeds in Arabidopsis [33-35].

Trehalose biosynthesis is important in regulating both seed composition and nodule metabolism [36,37]. In seeds, TREHALOSE 6 PHOSPHATE SYNTHASE 1 (TPS1), the enzyme responsible for converting glucose-6-phosphate to trehalose-6-phosphate is thought to regulate sugar utilization [32]. In *Arabidopsis thaliana tps1* null mutants, both sucrose and starch content of seeds dramatically increased [38]. In *Pv*, *TPS1* expression drops dramatically as the seed develops (Additional file 3i), corresponding with increased SS expression. We hypothesize that, as in Arabidopsis, the reduced *TPS1*
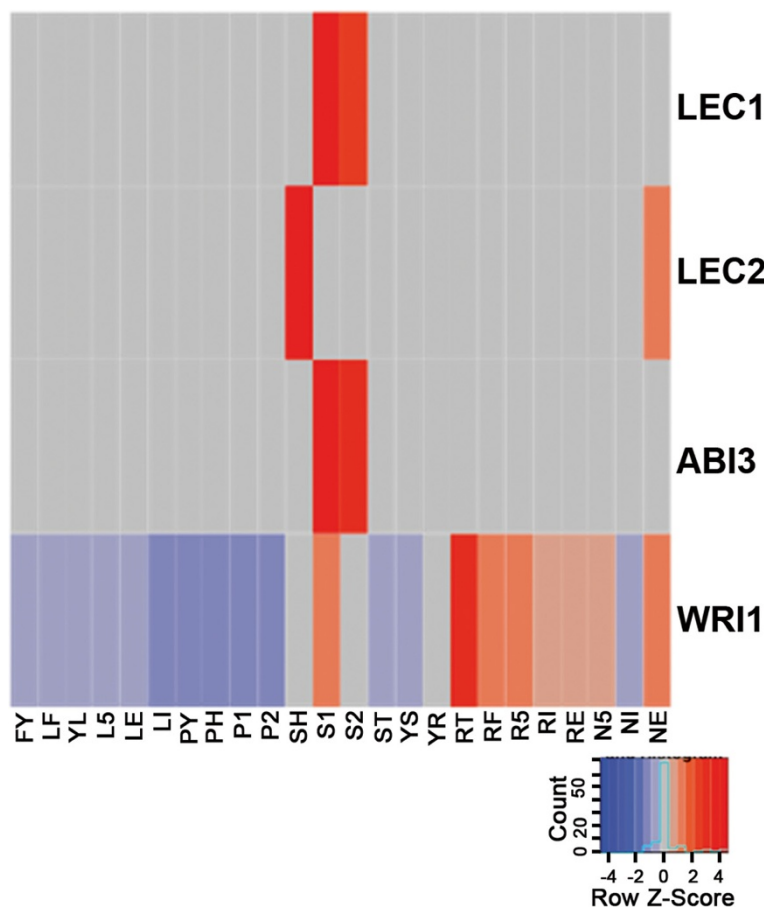
Figure 4 Seed master transcription factor expression. The expression profiles (as Z-scores: red = high, blue = low) of four transcription factors that regulate seed development in multiple species. Note the low expression of LEC2 (RPKM = 4) and WRI1 (RPKM = 5–9) in developing seeds. See Table 1 for tissue descriptions.

expression promotes increased carbohydrate biosynthesis in *Pv* seeds.

Nodule analysis

Legumes have established a unique symbiotic relationship with *Rhizobium*, which allows legumes to fix atmospheric N_2 into biologically useful NH_3. For this experiment plants were provided with nutrients containing NO_3^- nitrogen for optimal growth conditions or inoculated with either effective fix + *Rhizobium tropici* CIAT899 or ineffective fix- *Rhizobium giardini* 6917 to induce nodulation. Plants inoculated with normal (fix+) *R. tropici* appeared green and healthy, though smaller than plants provided with nitrate fertilizer (Additional file 13a). This phenotype is consistent with previous studies reporting the overall growth of N_2 fixing *Pv* plants is restricted compared to fertilized plants, likely due to altered carbon partitioning [39,40]. Fix + plants inoculated with *R. giardini* were nitrogen (N) deficient, exhibiting severe chlorosis and a stunted phenotype (Additional file 13a). Small, pre-fixing white nodules (N5) were isolated from root tissues of plants inoculated with effective fix + *R. tropici* five days after inoculation (DAI). At 21 DAI nodules were collected from plants inoculated with either fix + *R. tropici* (NE) or fix- *R. giardini* (NI) (Additional file 13b and c respectively). Microscopic imaging of fix-nodules 21 DAI (NI) revealed early senescing cells with few, if any, infected cells compared to fix + nodules formed 21 DAI (Additional file 13d and e). *In situ* hybridization analysis was used to visualize the localization pattern of leghemoglobin transcripts in these two nodule types. Fix- nodules collected 21 DAI (NI) exhibited little to no expression of leghemoglobin transcripts while fix + nodules collected 21 DAI (NE) exhibited high expression levels, likely mirroring the bacteroid colonization patterning (Additional file 13f and g) and directly reflecting the gene expression patterns observed in the RNA-Seq data (Figure 5). Nodule acetylene reduction assays failed to detect nitrogenase activity at 5 DAI in pre-fixing nodules and at 21 DAI with fix-nodules. Fix + nodules (and associated roots) from plants 21 DAI reduced 320 nm/hr/gfw (roots) of acetylene, indicating high N_2 fixation activity. Leaf tissue was also collected at 5 DAI for fix + plants and 21 DAI for both fix + and fix- inoculated plants (see Table 1).

Comparing 5 DAI pre-fixing nodules (N5) and 21 DAI fix + nodules (NE) revealed 2,932 differentially expressed genes (Additional file 14). Comparing 21 DAI fix + and fix-nodules (NE vs NI) identified 2,953 differentially expressed genes (Additional file 15). Additionally, we found 245 nodule specific genes; genes expressed in any and/or all the nodule tissues sampled, but not expressed in any other tissues (Additional file 4). Comparing these nodule specific genes to those identified in two soybean gene atlases

[26,41] identified 21 nodule specific homologs common to both species (Additional file 16) including seven TFs and four transporters. The conserved expression of these genes highlights the importance of regulating gene expression, but also the exchange of nutrients between nodules and the plant roots. Five of these sequences have no known annotation, though nodule specific expression in both species suggests these are important candidates for characterization in future nodulation and nitrogen fixation research.

Cognate genes involved in nodule development and the establishment of N fixation have been identified and cloned from multiple species [42,43]. Using BLASTN, the homologous sequences in common bean were identified and gene expression patterns were visualized as a heat-map (Figure 5). Upon further analysis of these nodule cognate genes, we detected three expression profiles: those highly up regulated early in nodule development (N5), those highly up regulated in 21DAI fix + nodules (NE), and those highly up regulated in 21 DAI fix- nodules (NI).

The autoregulation of nodulation (AON) pathway mediates nodule formation [44]. ASTRAY and UFD1a proteins, both expressed in leaves, function in the AON pathway [44]. *ASTRAY*, encodes a bZIP TF that interacts with a nodulation autoregulation receptor kinase (NARK) [44,45]. *UFD1a* expression indicates the presence of Q, a root derived signal induced upon compatible *rhizobial* infection [44]. In soybean, three candidates for Q have been identified, all of which are CLE peptides [46]. *Pv* encodes a single CLE homolog (Phvul.005G097000.1), which is highly up regulated in N5 and NE, but noticeably absent in NI (Figure 5). Surprisingly, we observed aerial AON genes (*ASTRAY and UFD1a*) expressed higher in leaves of plants inoculated with fix- *rhizobia* than in leaves of plants inoculated with fix + rhizobia (Figure 5). We hypothesize the fix- inoculated plant may up regulate the AON pathway to minimize resources allocated to nodules as part of a survival strategy.

Early in nodule development (N5) nod factor receptors (NFR) and nodulation signaling pathway (NSP) TFs are highly expressed (Figure 5). The early calcium spiking response induces both a calmodulin dependent protein kinase (CCaMK/DMI3), which is required and sufficient for nodule organogenesis [47-49], and nuclear porin proteins (NUPs) [44] (Figure 5). Additionally, genes involved in infection thread formation and elongation including *ERN1, FLOT, VPY, PUB1* and *RPG* [50-58] are highly up regulated in N5 (Figure 5). *VPY and PUB1*, which are up regulated in N5 and NE, are involved in *rhizobial* recognition, attachment, entry, and initiation of the infection thread [53,58]. Nodule organogenesis involves the altered differentiation and division of root cortical cells prior to the formation of the nodule primordia. In *Medicago truncatula*, these processes are dependent on *ENOD40* [59],
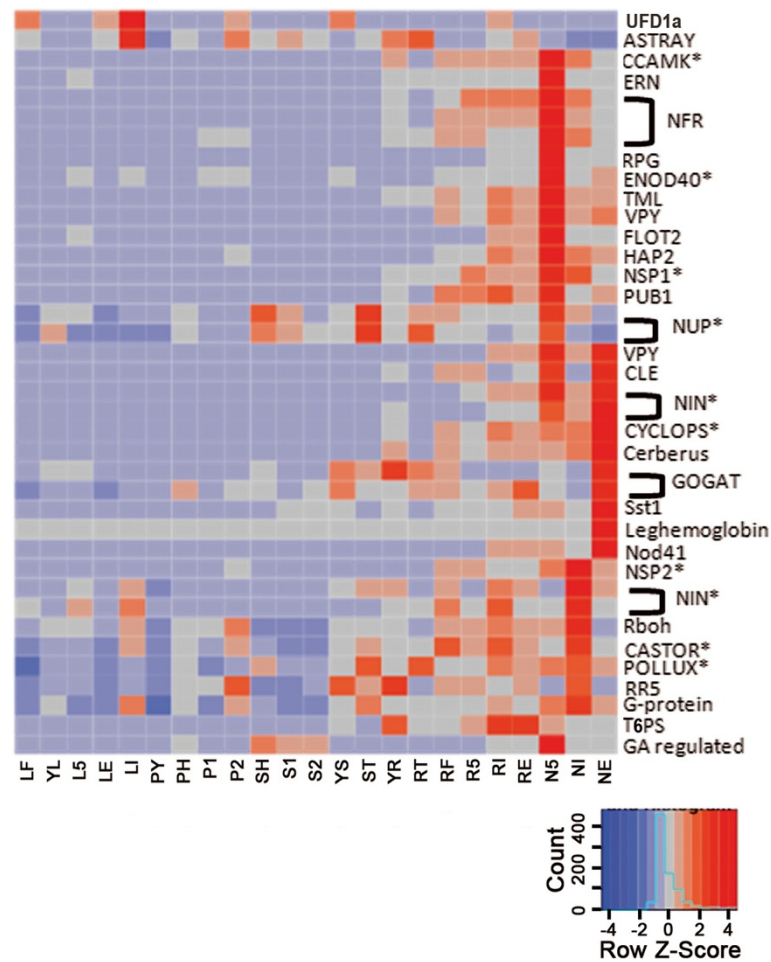
Figure 5 Nodulation gene expression patterns. Expression patterns (as Z-scores) of *Pv* homologs of genes involved in nodulation and symbiosis identified in *Lotus japonicus*, *Medicago truncatula*, and *Glycine max*. Red indicates a positive Z-score while blue indicates a negative Z-score. Genes common to both symbiotic nitrogen fixation and mycorrhizal symbiosis are denoted with an asterisk (*). See Table 1 for tissue descriptions.

which is highly expressed in N5 (Figure 5). *HAP2*, which promotes nodule development and the release of bacteria from the infection thread [60] is expressed highest in N5, but remains elevated in NE (Figure 5).

Genes highly expressed in NE are involved in processes such as amino acid biosynthesis, nitrogen metabolism, carbohydrate metabolism, membrane transport, and sulfur assimilation (cysteamine dioxygenase). We identified 402 genes highly expressed in NE as compared to all other tissues (Additional file 17 and Additional file 18). These genes are likely involved in the establishment of symbiosis and symbiotic nitrogen fixation (SNF). Among these 402 genes, 73 encoding a transmembrane domain, 49% of which relate to transport including Phvul.002G300900.1, which encodes a SWEET4 homolog. SWEET genes mediate sucrose transport to the phloem [61]. In Arabidopsis, SWEET4 is up regulated by pathogen infection, likely being co-opted to aid pathogen growth [62]. We hypothesize this function is conserved in *Pv* upon fix + *Rhizobum* infection.

Consistent with these results, the two most statistically significant GO categories among the 402 genes are GO:0005215 (P-value = 0.002), associated with transport activity and GO:0006857, associated with oligopeptide transport (P-value = 0.007). Increased expression of transporters in nodule tissues is consistent with previous reports [63] of nodule organogenesis gene expression. Also important to nodule function is carbohydrate metabolism, which is statistically over-represented in genes highly expressed in NE (Table 3, GO:0030246) [63]. Among the 402 genes up regulated in NE are 34 TFs belonging to 18 different families and 73 transporters. Four of the TFs belong to the Nodule Inception (NIN) family. NIN TFs mediate signals of rhizobial infection including: root hair curling, infection thread formation, and the initiation of the nodule primodia [64-66]. They are also involved in regulating gene expression in response to nitrate. Additional highly expressed genes in NE are members of the shi related sequence (SRS) TF family, with 32% of the familial expression from nodules

(Figure 1). SRS TFs mediate protein:protein interactions involved in ubiquitination for targeted proteolysis. Expression data from both *Pv* and soybean [26,41] indicates this family is highly expressed in both roots and nodules of legumes (Figure 1).

Among the nodule cognate genes, those most highly expressed in NE are: NIN transcription factors, *CYCLOPS*, and *CERBERUS* with expression profiles increasing from 5 DAI to 21 DAI in fix + nodules (2X, 2.4X, and 4X respectively, Figure 5). *CYCLOPS* expression is required for rhizobia infection. Nodule organogenesis is dependent on *CERBERUS* gene expression [49,67-69].

The primary function of nodules is to fix N_2 to NH_3. Nitrogenase, the enzyme responsible for nitrogen fixation, requires sulfur and a near anaerobic environment to function. The gene encoding SYMBIOSOME SULFATE TRANSPORTER 1 (SST1), which transports sulfur into bacteroids [70], is expressed 15 fold higher in NE than in NI (Figure 5). The genes encoding leghemoglobin, which sequesters oxygen [71], are only expressed in NE (Figure 5). Once N_2 is reduced to usable ammonia it must be assimilated for use and distribution throughout the plant. NADH-dependent glutamate synthase (NADH-GOGAT) is a key enzyme in ammonia assimilation [72]. Two NADH-GOGAT genes (Phvul.009G053900.1 and Phvul.001G076400.1) are expressed 5 and 10-fold higher in NE than in NI (Figure 5), reflecting the difference in effectiveness. Transcripts encoding glutamine synthase, uricase, and allantoinase (Additional file 3 j, c, and d), each involved in primary ammonia assimilation, exhibit similar expression patterns. Nod41 was identified in *Pv* by Olivares *et al.* [73] as important in controlling defense responses during symbiotic interactions and maintaining the integrity of the uninfected root nodule cells. Our data is consistent with this hypothesis as *Nod41* is expressed 7-fold higher in 21 DAI fix + nodules than in 21 DAI fix- nodules (Figure 5).

Genes up regulated in NI include those involved in the GO processes of autophagy and early senescence including ubiquitination, proteolysis, peptidyolysis and apoptosis. Respiratory burst oxidase homolog (RBOH) genes, which generate reactive oxygen species (ROS) [74], are up regulated four-fold in NI. Increased ROS production is a common defense response to pathogen attack (ie: ineffective rhizobia) and in response to abiotic stress, including nitrogen deficiency. Additionally, we observed high expression of leucine rich repeat (LRR) genes in NI (Additional file 19), likely reflecting a defense response as the plant reacts to invading bacteria [75]. Also highly expressed in NI are genes involved in oxidation-reduction processes, membrane transport, protein binding, and amino acid phosphorylation. Among the nodule cognate genes are those encoding a second group of NIN and NSP TFs (Figure 5). This result suggests a group of alternative TFs may be induced in NI versus NE.

Type A-response regulators (RRs) negatively regulate cytokinin signaling [76]. In lotus, RRs are rapidly induced following *rhizobial* inoculation in root hairs and dividing cortical cells [77], repressing the cytokinin signaling pathway [48]. Inhibition of the cytokinin-signaling pathway may contribute to plant and bacterial cell differentiation. In *Pv*, the gene encoding RR5 is more highly expressed in NI than in NE (Figure 5). Genes encoding both CASTOR and POLLUX, both of which are required for the activation of voltage gated calcium (Ca^{2+}) channels [74], are highly expressed in NI (Figure 5). The high expression of *CASTOR* and *POLLUX* genes in NI may suggest that at 21 DAI the plant is still attempting to establish SNF or may reflect the induction of Ca^{2+} channels by ROS as described in *Pisum sativum* [78,79].

Roots and nitrogen

Gene expression profiles of *Pv* roots were examined from plants grown under three conditions 1) those from NO_3^-

Table 3 Gene Ontology (GO) categories statistically over-represented in NE enhanced genes

GO ID	P-value	Number of genes[a]	Description
GO:0005215	0.00271	14	Transporter Activity
GO:0006857	0.00702	8	Oligopeptide Transport
GO:0004106	0.00871	3	Chorismate Mutase Activity
GO:0006188	0.01178	2	IMP Biosynthetic Process
GO:0006950	0.01773	7	Response to Stress
GO:0047800	0.01929	3	Cysteamine Dioxygenase Activity
GO:0030246	0.01929	3	Carbohydrate Binding
GO:0009073	0.01929	3	Aromatic Amino Acid Biosynthesis
GO:0006164	0.02765	2	Purine Nucleotide Biosynthesis
GO:0009113	0.02765	2	Purine Base Biosynthesis
GO:0030259	0.03735	3	Lipid Glycosylation

[a]The number of genes on NE enhanced list with GO ID of interest.

fertilized plants (RF), 2) those from plants with fix + effective nodules (RE), and 3) those derived from plants having fix- ineffective nodules (RI). RF and RE roots had adequate N for growth, while RI roots were N deficient. We identified 1,714 genes differentially expressed between the three 21 DAI root samples (RI, RE, and RF) (Additional file 20). The majority of these genes (1,668 genes) are differentially expressed between fertilized roots (RF) and nodulated roots (either RE or RI).
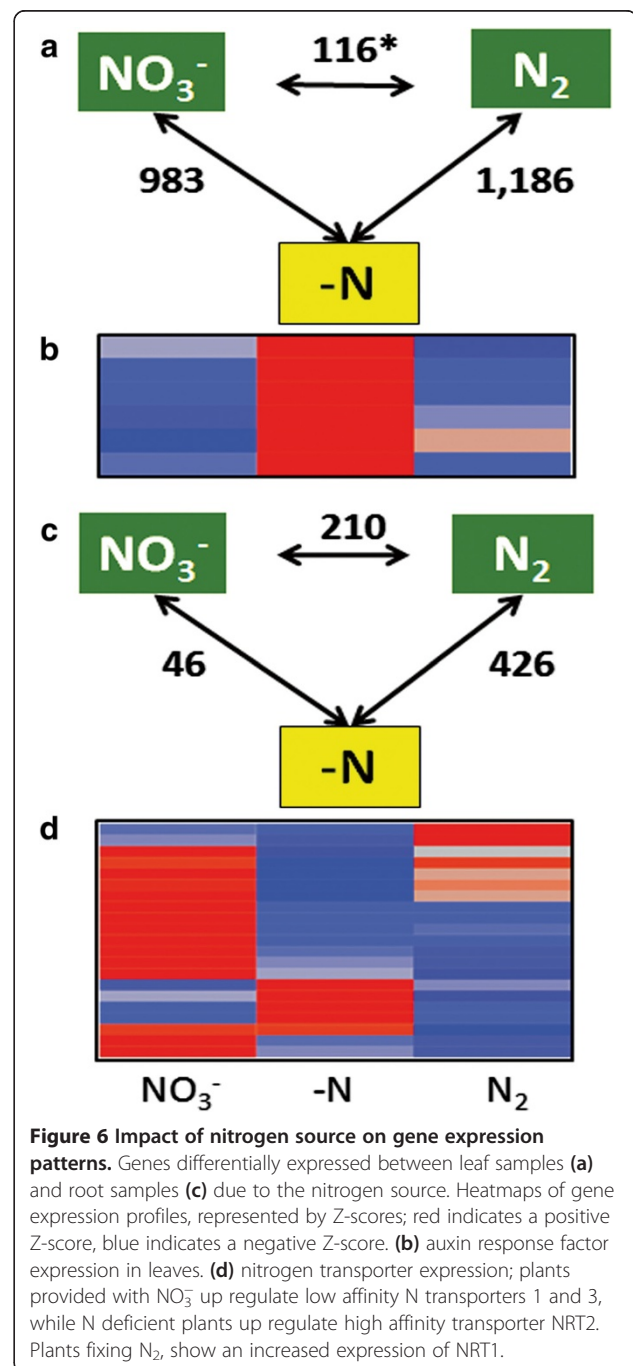
Comparing gene expression patterns between RI (roots from fix- plants) and both RF and RE identified 426 and 46 genes differentially expressed between root samples respectively (Figure 6c). Additionally, 210 genes were differentially expressed between RE and RF. These 210 sequences represent genes differentially expressed due to the N source (Figure 6c). A similar comparison of leaf tissues collected from each of the plants revealed 116 genes differentially expressed between + N leaves due to different N sources (Figure 6a). These 116 genes indicate that the source of N (either via N_2 fixation or NO_3^- fertilization) has a long-term impact on plant gene expression. Among all 2,641 genes differentially expressed between samples due to the N source are 340 TFs, the majority of which are up regulated in −N tissues (Additional file 21 a and b).

In the presence of abundant NO_3^-, plants will preferentially take up and utilize NO_3^- rather than develop SNF. NO_3^- transporters exhibit either low (NRT1) or high (NRT2 and NRT3) NO_3^- affinity [80-84]. Examination of the expression patterns of NO_3^- transporters in our root samples revealed plants provided with NO_3^- as a fertilizer induce *NRT1* and *NRT3* gene expression, reflecting the abundance of available NO_3^- (Figure 6d). In −N roots, only members of the high affinity NRT2 gene family are up regulated (Figure 6d). N-deficient plants may up regulate N transporters in an attempt to increase the N content of the plant to mitigate −N stress. Members of the low affinity NRT1 gene family are also up regulated in the roots of N_2 fixing plants (Figure 6d). The constitutive expression of *NRT1.1* is consistent with the recent evidence suggesting that it serves as both an N sensor and transporter [85-87]. Expression of *NRT1.1* in fix + plants may be involved in N sensing.

Once N is within the plant it must be assimilated. Glutamine synthetase (GS) functions as a primary enzyme for NH_4 assimilation produced from N_2 fixation or NO_3^- nutrition [88,89]; synthesizing glutamine from NH_3 and glutamate (Additional file 3j). In SNF plants, the majority of glutamine is committed to the *de-novo* purine biosynthesis pathway. Alternatively, glutamine may be reduced by GOGAT. Consistent with previous studies [72], *NADH-GOGAT* expression is highest in roots (particularly YR) and NE while *Fd-GOGAT* is expressed highest in leaf tissues (Additional file 3 k). Plants provided with NO_3^- fertilization utilize glutamine in the synthesis of asparagine via asparagine synthetase (AS) [88], which is most highly expressed in fertilized root (RF) tissues (Additional file 3 l).

Our data shows increased expression of auxin response factors unique to −N leaves (Figure 6b). This gene expression pattern indicates increased auxin levels in −N leaves, supporting auxin as the N signal. The availability of N for proper growth and development is likely monitored throughout the plant. Auxin has been proposed as an N status mediator, signaling from root to

Figure 6 Impact of nitrogen source on gene expression patterns. Genes differentially expressed between leaf samples **(a)** and root samples **(c)** due to the nitrogen source. Heatmaps of gene expression profiles, represented by Z-scores; red indicates a positive Z-score, blue indicates a negative Z-score. **(b)** auxin response factor expression in leaves. **(d)** nitrogen transporter expression; plants provided with NO_3^- up regulate low affinity N transporters 1 and 3, while N deficient plants up regulate high affinity transporter NRT2. Plants fixing N_2, show an increased expression of NRT1.

shoot [80]. Under low N and other nutrient stress conditions in the shoot, increased auxin is transported to the roots to enhance lateral root development, a hallmark response of –N plants [46,80,87].

Conclusion

This study provides a resource for global analysis of gene expression patterns in *Pv* of 24 samples from seven unique tissues across important developmental time points. The publicly available gene atlas, *Pv* GEA, will facilitate the use of this data for researchers querying gene expression patterns within various biological processes, as evidenced by Additional file 3. Additionally, by comparing gene expression patterns in developing seeds of *Pv* to those in *Glycine max*, we were able to identify differences potentially responsible for altered seed composition between the two closely related species. Finally, our analysis of N uptake and utilization revealed the N source is an important component of the N pathway and has a long-term effect on gene expression patterns.

Methods
Plant materials and growth conditions

Phaseolus vulgaris cv. Negro jamapa seeds were grown as described by O'Rourke et al. [90]. At the emergence of the unifoliate, pots were assigned to one of three nitrogen (N) treatments; inoculated with *Rhizobium tropici* CIAT899 (fix+), *Rhizobium giardini* 6917 (fix-), or fertilized with a full nutrient solution. Pots assigned to the fertilization treatment were watered daily with a nutrient solution of 9 mM KNO_3, 2.5 mM $Ca(NO_3)_2$, 1.0 mM $Ca(H_2PO_4)_2$, 1.0 mM $MgSO_4$, 12 µM Fe (as FeEDTA), 4.0 µM $MnCl_2$, 22.0 µM H_3BO_3, 0.4 µM $NaMoO_4$, and 1.6 µM $CuSO_4$. Twenty-four tissue samples were collected throughout development and across all N treatments (for details see Table 1). For each nitrogen treatment, a representative plant was chosen and leaf, root, and nodule tissue samples were collected. Two plants were maintained in full nutrient solution fertilized pots. From these plants, root, leaf, flower, stem, seed and pod tissues were collected (see Table 1 for details). All tissue collected for RNA -Seq analysis was immediately flash frozen in liquid nitrogen.

RNA extraction and expression analysis

Total RNA was purified from 24 tissue samples using RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). For RNA-Seq analysis, RNA samples were shipped on dry ice overnight to National Center for Genome Resources (NCGR, Santa Fe, NM) for sequencing as described by Severin et al. [26]. Illumina reads generated from all 24 samples are available at the NCBI SRA browser, accession number SRP046307. Illumina reads passing quality

control standards, approximately 25 million sequences per sample or 596 million 36 bp sequences total, were mapped to the *Phaseolus vulgaris* v1.0 genome available at www.phytozome.net using the program Bowtie [91]. Reads were also mapped to the predicted transcripts to account for splicing events. Reads that mapped to more than one location were counted at each mapping location. Of the 596 million reads generated, 406 million (89%) mapped to the genome with 14% of those mapping to non-coding regions. Raw gene expression counts were normalized using the RPKM (reads/Kb/Million) method [92,93] using custom R scripts. To ensure expression profiles were not statistical artifacts as described by earlier studies [94,95], we determined an RPKM of 3 represents a 2X coverage across the coding region, assuming equal distribution, and would be the minimum level at which a gene would be considered expressed; genes with an RPKM < 3 were considered silent. Transcripts differentially expressed between libraries were identified using NOIseq [96]. Differentially expressed transcripts were required to have > 2-fold change in expression between samples and a probability of differential expression > 0.9. Additionally, one of the two sequences was required to have an RPKM > 3. Heatmaps illustrating expression patterns of various subgroups of transcripts were generated in R as described by Severin *et al.* [26].

To identify genes exhibiting enhanced expression in NE, we determined the Euclidian distance between Z-scores for each gene. A threshold of two Euclidian distances was established as significant; genes within the threshold were identified as co-expressed.

Real time quantitative RT-PCR (qPCR) and housekeeping genes

RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) from three biological replicates of tissues grown in growth chambers under the same conditions described above. Transcript specific primers were designed using Primer3 (Frodo.wi.mit.edu). The qPCR analysis was run as described by O'Rourke et al. [90] for 85 genes identified as differentially expressed by NOIseq (Additional file 6). 92% of the qPCR experiments confirmed the differential expression measured by NOIseq analysis in at least two of the three biological replicates.

Genes exhibiting stable expression profiles between tissues and across growth conditions were identified as described by Severin et al. [26] (Additional file 22). The 10% of transcripts with the lowest CV were selected as potential housekeeping genes. This suite of stably expressed transcripts may be useful in future experiments for normalizing gene expression patterns across a variety of experimental conditions, or tissues [97]. One of these genes (Phvul.006G165300.1) was successfully utilized as a

housekeeping sequence in the qPCR analysis. Comparing the housekeeping genes proposed in this study to the eleven potential housekeeping genes identified for *Pv* under biotic and abiotic stress by Borges et al. [98] found seven genes common to both lists, illustrating the utility of this list for multiple experimental conditions.

Acetylene reductase assay
Acetylene reduction assays were performed as described by Vance et al. [99] with the following modifications. Plant roots from six biological replicates of each sample (roots inoculated with fix + *R. tropici* CIAT899 at 5 and 21 DAI and roots inoculated with fix − *R. giardini* 6917 21 DAI) were placed in 500 ml airtight glass containers equipped with serum stoppers. 50 ml of air was removed from each container and replaced with 50 ml of ethylene; samples were incubated at room temperature for one hour, at which time 10 ml of gas was withdrawn from the container for analyses as previously described by Vance et al. [99].

Nodule *In Situs*
A partial coding sequence for *Pv* leghemoglobin (645 bp) was PCR amplified and cloned into pBSSK+. Nodules inoculated with fix + *R. tropici* CIAT899 and fix − *R. giardini* 6917 were collected 21 DAI and analyzed as described by Sbabou et al. [100].

Availability of supporting data
The expression data used in this study is publicly available at the NCBI short read archive; accession SRP046307. Additionally, the raw and normalized datasets can be downloaded and explored at the *Phaseolus vulgaris* Gene Expression Atlas (*Pv* GEA) website, http://plantgrn.noble.org/PvGEA/.

Additional files

Additional file 1: Pairwise identification of differentially expressed genes.

Additional file 2: Tissue specific genes.

Additional file 3: Expression patterns of specific genes of interest. Graphs illustrating the expression patterns of genes of interest, individual tissues on the X-axis, RPKM values, on the Y-axis. (a) tissue specific, (b) sample specific genes, (c) uricase (d) allantoin degradation, (e) glutamate dehydrogenase, (f) transcription factors with increased expression as seeds develop, (g) ABA1 and ABA2, (h) NCED9, (i) trehalose 6 phosphate, (j) glutamine synthetase, (k) GOGAT, (l) asparagine synthetase.

Additional file 4: Sample specific genes.

Additional file 5: Purine Biosynthesis. Expression patterns of the *Pv* homologs in the purine biosynthesis pathway. Genes in this pathway are most highly expressed in effective nodules compared to the rest of the plant. For tissue descriptions see Table 1. Gene expression is represented by Z-scores; red indicates a positive Z-score while blue indicates a negative Z-score.

Additional file 6: qPCR results.

Additional file 7: Transcription factors.

Additional file 8: Genes differentially expressed between SH and PH.

Additional file 9: Genes differentially expressed between S1 and P1.

Additional file 10: Genes differentially expressed between S2 and P2.

Additional file 11: Genes differentially expressed as seeds develop.

Additional file 12: Genes differentially expressed as pods develop.

Additional file 13: Nitrogen Deficiency, Plants, Nodules, and *In situs*. (a) Plant phenotypes, plant on the left provided with nitrate (NO$_3^-$) fertilizer, middle plant inoculated with fix + *rhizobium*, plant on right inoculated with fix- *rhizobium* (nodules form, but no N$_2$ fixation). (b) Effective (fix+) nodules 21 DAI with fix + *R. tropici* CIAT 899. (c) Ineffective (fix-) nodules 21 DAI with fix- *R. giardini* 6917. (d and e) Cross section of fix + (d) and fix- (e) nodules stained with toluene blue. Bacteroid structures visible in fix + nodules. Disorganized cellular structure evident in (e). (f and g) Transcript abundance of leghemoglobin: high in fix + nodules (f) but barely detectable in fix- nodules (g), as assessed by *in situ* hybridizations.

Additional file 14: Genes differentially expressed between NE and N5.

Additional file 15: Genes differentially expressed between NE and NI.

Additional file 16: Nodule specific genes common to *Pv* and *Glycine max*.

Additional file 17: Genes highly expressed in NE compared to other tissues.

Additional file 18: Genes highly up regulated in NE.

Additional file 19: LRR gene expression. 151 genes containing an LRR domain expressed in our data. Expression, represented by Z-scores, of LRR containing sequences is low in developing seeds, but high in ineffective nodules. Red indicates a positive Z-score while blue indicates a negative Z-score. For tissue description see Table 1.

Additional file 20: Genes differentially expressed between root samples.

Additional file 21: Impact of N on TF expression. Transcription factors differentially expressed in leaf samples (a) and root samples (b) of plants provided with NO$_3^-$ fertilizer, −N plants (inoculated with fix- rhizobium), and N$_2$ fixing plants (inoculated with fix + rhizobium). Heatmaps of gene expression represented by Z-scores; red indicates a positive Z-score while blue indicates a negative Z-score.

Additional file 22: Housekeeping genes.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
JAO, LPI, and CPV: conceived the project, analyzed the data, and drafted the manuscript. FF: performed bioinformatic analyses and initiated database construction. BB: provided plant tissues and performed RNA extractions. SS: performed qPCR experiments. SAJ and PEM: conceived the project and critically reviewed the manuscript. JL, XD and PXZ: built the gene expression atlas database. GH: advised on the experimental plan and critically reviewed the manuscript. All authors have read and approved the final manuscript.

Author details
[1]Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA. [2]Centro de Ciencias Genomicas, Universidad Nacional Autonoma de Mexico, 66210 Cuernavaca, Mor, Mexico. [3]USDA-Agricultural Research Service, Plant Science Research Unit, St. Paul, MN 55108, USA. [4]Center for Applied Genetic Technologies, University of Georgia, Athens, GA 30602, USA. [5]Department of Plant Sciences, North Dakota State University, Fargo, ND 58105, USA. [6]Plant Biology Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA. [7]Current Address: USDA-ARS, Corn Insect Crop Genetics Research Unit, Iowa State University, Ames, IA 50011, USA.

References

1. Broughton WJ, Hernandez G, Blair M, Beebe S, Gepts P, Vanderleyden J: Beans (*Phaseolus* spp.) - model food legumes. *Plant Soil* 2003, **252**:55–128.
2. Graham PH, Vance CP: Legumes: importance and constraints to greater use. *Plant Physiol* 2003, **131**:872–877.
3. Ramirez M, Graham MA, Blanco-Lopez L, Silvente S, Medrano-Soto A, Blair MW, Hernandez G, Vance CP, Lara M: Sequencing and analysis of common bean ESTs, building a foundation for functional genomics. *Plant Physiol* 2005, **137**:1211–1227.
4. Tian J, Venkatachalam P, Liao H, Yan X, Raghothama K: Molecular cloning and characterization of phosphorus starvation responsive genes in common bean (*Phaseolus vulgaris* L.). *Planta* 2007, **227**:151–165.
5. Hernandez G, Ramirez M, Valdes-Lopez O, Tesfaye M, Graham MA, Czechowski T, Schlereth A, Wandrey M, Erban A, Cheung F, Wu HC, Lara M, Town CD, Kopka J, Udvardi MK, Vance CP: Phosphorus stress in common bean: root transcript and metabolic responses. *Plant Physiol* 2007, **144**:752–767.
6. Thibivilliers S, Joshi T, Campbell KB, Scheffler B, Xu D, Coopers B, Nguyen HT, Stacey G: Generation of *Phaseolus vulgaris* ESTs and investigation of their regulation upon *Uromyces appendiculatus* infection. *BMC Plant Biol* 2009, **9**:46.
7. Melotto M, Monteiro-Vitorello CB, Bruschi AG, Carmargo LE: Comparative bioinformatic analysis of genes expressed in common bean (*Phaseolus vulgaris* L.) seedlings. *Genome* 2005, **48**:562–570.
8. Kalavacharia V, Liu Z, Meyers BC, Thimmapuram J, Melmaiee K: Identification and analysis of common bean (*Phaseolus vulgaris* L.) transcriptomes by massively parallel pyrosequencing. *BMC Plant Biol* 2011, **11**:135. 107.
9. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C: A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet* 2014, **46**(7):707–713.
10. Lin J-Y, Stupar RM, Hans C, Hyten DL, Jackson SA: Structural and functional divergence of a 1-Mb duplicated region in the soybean (*Glycine max*) genome and comparison to an orthologous region from *Phaseolus vulgaris*. *Plant Cell* 2010, **22**:2545–2561.
11. McClean PE, Marnidi S, McConnell M, Chikara S, Lee R: Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. *BMC Genomics* 2010, **11**:184–194.
12. Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC: Mining EST databases to resolve evolutionary events in major crop species. *Genome* 2004, **47**:868–876.
13. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L. et al.: Genome sequence of the palaeopolyploid soybean. *Nature* 2010, **14**:178–183.
14. Li J, Dai X, Liu T, Zhao PX: LegumeIP: an integrative database for comparative genomics and transcriptomics of model legumes. *Nucleic Acids Res* 2011, **40**(D1):D1221–D1229.
15. Smith PMC, Atkins CA: Purine biosynthesis. Big in cell division, even bigger in nitrogen assimilation. *Plant Physiol* 2002, **128**:793–802.
16. Ishisuka J: Function of Symbiotically Fixed Nitrogen for Grain Production in Soybean. In *Proceedings of the International Seminar on Soil Intensive Agriculture*. Tokyo, Japan: Society of the Science of Soil and Manure; 1977:617–624.
17. Thomas RJ, Schrader LE: The assimilation of ureides in shoot tissues of soybeans: changes in allantoinase activity and ureide contents of leaves and fruits. *Plant Physiol* 1981, **67**:973–976.
18. Pelissier HC, Frerich A, Desimone M, Schumacher K, Tegeder M: PvUPS1, an allantoin transporter in nodulated roots of French bean. *Plant Physiol* 2004, **134**(2):664–675.
19. Chen J, Huang B, Li Y, Du H, Gu Y, Liu H, Zhang J, Huang Y: Synergistic influence of sucrose and abscisic acid on the genes involved in starch synthesis in maize endosperm. *Carbohydr Res* 2011, **346**:1684–1691.
20. Wang Z, Libault M, Joshi T, Valliyodan B, Nguyen HT, Xu D, Stacey G, Cheng J: SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biol* 2010, **10**:14.
21. Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Hovath S, Drews GN, Fischer RL, Okamuro JK, Harada JJ, Goldberg RB: Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proc Natl Acad Sci U S A* 2010, **107**:8063–8070.
22. Verdier J, Kaker K, Gallardo K, Le Signor C, Aubert G, Schlereth A, Town CD, Udvardi MK, Thompson RD: Gene expression profiling of *M. truncatula* transcription factors identifies putative regulators of grain legume seed filling. *Plant Mol Biol* 2008, **67**:567–580.
23. Le BH, Wagmaister JA, Bui AQ, Harada JJ, Goldberg RB: Using genomics to study legume seed development. *Plant Physiol* 2007, **144**:562–574.
24. Hajduch M, Ganapathy A, Stein JW, Thelen JJ: A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol* 2005, **137**:1397–1419.
25. Bolon Y-T, Vance CP: Characterization of the Linkage Group I Seed Protein QTL in Soybean. In *Designing Soybeans for the 21st Century*. Edited by Wilson RF. Urbana: AOCS Press; 2012:175–195.
26. Severin AJ, Woody JL, Bolon Y-T, Joseph B, Diers BW, Farmer AD, Muehlbauer GJ, Nelson RT, Grant D, Specht JE, Graham MA, Cannon SB, May GD, Vance CP, Shoemaker RC: RNA-Seq atlas of *Glycine max*: a guide to the soybean transcriptome. *BMC Plant Biol* 2010, **10**:160.
27. Weber H, Borisjuk L, Wobus U: Molecular physiology of legume seed development. *Annu Rev Plant Biol* 2005, **56**:253–279.
28. Angeles-Nunez JG, Tiessen A: Arabidopsis sucrose synthase 2 and 3 modulate metabolic homeostasis and direct carbon towards starch synthesis in developing seeds. *Planta* 2010, **232**(3):701–718.
29. Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J: Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* 1994, **6**(1):147–158.
30. Verdier J, Thompson RD: Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol* 2008, **49**(9):1263–1271.
31. Angeles-Nunez JG, Tiessen A: Mutation of the transcription factor LEAFY COTYLEDON 2 alters the chemical composition of Arabidopsis seeds, decreasing oil and protein content, while maintaining high levels of starch and sucrose in mature seeds. *J Plant Physiol* 2011, **168**:1891–1900.
32. Gutierrez L, Van Wuytswinkel O, Catelain M, Bellini C: Combined networks regulating seed maturation. *Trends Plant Sci* 2007, **12**:294–300.
33. Kanno Y, Jikumaru Y, Hanada A, Nambara E, Abrams SR, Kamiya Y, Seo M: Comprehensive hormone profiling in developing Arabidopsis seeds: examination of the site of ABA biosynthesis, ABA transport and hormone interactions. *Plant Cell Physiol* 2010, **51**(12):1988–2001.
34. Lefebvre V, North N, Frey A, Sotta B, Seo M, Okamoto M, Nambara E, Marion-Poll A: Functional analysis of Arabidopsis NCED6 and NCED9 genes indicates that ABA synthesize in the endosperm is involved in the induction of seed dormancy. *Plant J* 2006, **45**:309–319.
35. Frey A, Effroy D, Lefebvre V, Seo M, Perreau F, Berger A, Sechet J, To A, North HM, Marion-Poll A: Epoxycarotenoid cleavage by NCED5 fine-tunes ABA accumulation and affects seed dormancy and drought tolerance with other NCED family members. *The. Plant J* 2012, **70**(3):501–512.
36. Eastmond PJ, van Dijken AJ, Spielman M, Kerr A, Tissier AF, Dickinson HG, Jones JD, Smeekens SC, Graham IA: Trehalose-6-phosphate synthase 1, which catalyses the first step in trehalose synthesis is essential for Arabidopsis embryo maturation. *Plant J* 2002, **29**:225–235.
37. Barraza A, Estrada-Navarrete G, Rodriguez-Alegria ME, Lopez-Munguia A, Merino E, Quinto C, Sanchez F: Down-regulation of PvTRE1 enhances nodule biomass and bacteroid number in the common bean. *New Phytol* 2013, **197**:194–206.
38. Gomez LD, Gilday A, Feil R, Lunn JE, Graham IA: AtTPS1-mediated trehalose 6-phosphate synthesis is essential for embryogenic and vegetative growth and responsiveness to ABA in germinating seeds and stomatal guard cells. *Plant J* 2010, **64**:1–13.
39. Finke RL, Harper JE, Hageman RH: Efficiency of nitrogen assimilation by N2-fixing and nitrate grown soybean plants (*Glycine max* [L.] Merr.). *Plant Physiol* 1982, **70**:1178–1184.
40. Timpo EE, Neyra CA: Expression of nitrate and nitrite reductase activities under various forms of nitrogen nutrition in *Phaseolus vulgaris* L. *Plant Physiol* 1983, **72**:71–75.
41. Libault M, Farmer A, Joshi T, Takahashi K, Langley RJ, Franklin LD, He J, Xu D, May G, Stacey G: An integrated transcriptome atlas of the crop model

Glycine max, and its use in comparative analyses in plants. *Plant J* 2010, **63**:86–99.

42. Kouchi H, Imaizumi-Anraku H, Hayashi M, Hakoyama T, Nakagawa T, Umehara Y, Suganuma N, Kawaguchi M: How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant Cell Physiol* 2010, **51**(9):1381–1397.

43. Hayashi S, Reid DE, Lorenc MT, Stiller J, Edwards D, Gresshoff PM, Ferguson BJ: Transient Nod factor-dependent gene expression in the nodulation-competent zone of soybean (*Glycine max* [L.] merr.) roots. *Plant Biotechnol J* 2012, **10**:995–1010.

44. Reid DE, Hayashi S, Lorenc M, Stiller J, Edwards D, Gresshoff PM, Ferguson BJ: Identification of systemic responses in soybean nodulation by xylem sap feeding and complete transcriptome sequencing reveal a novel component of the autoregulation pathway. *Plant Biotechnol J* 2012, **10**(6):680–689.

45. Oka-Kira E, Kawaguchi M: Long-distance signaling to control root nodule number. *Curr Opin Plant Biol* 2006, **9**(5):496–502.

46. Reid DE, Ferguson BJ, Gresshoff PM: Inoculation and nitrate induced CLE peptides of soybean control NARK-dependent nodule formation. *Mol Plant Microbe Interact* 2011, **24**:606–618.

47. Rival P, de Billy F, Bono JJ, Gouch C, Rosenberg C, Bensmihen S: Epidermal and cortical roles of NFP and DMI3 in coordinating early steps of nodulation in *Medicago truncatula*. *Development* 2012, **139**:3383–3391.

48. Desbrosses GJ, Stougaard J: Root Nodulation: a paradigm for how plant-microbe symbiosis influences host developmental pathways. *Cell Host & Microbe* 2011, **10**:348–358.

49. Hayashi T, Shimoda Y, Sato S, Tabata S, Imaizumi-Anraku H, Hayashi M: Rhizobial infection does not require cortical expression of upstream common symbiosis genes responsible for the induction of Ca2+ spiking. *Plant J* 2014, **77**:146–159.

50. Haney CH, Riely BK, Tricoli DM, Cook DR, Ehrhardt DW, Long SR: Symbiotic rhizobia bacteria trigger a change in localization and dynamics of the *Medicago truncatula* receptor kinase LYK3. *Plant Cell* 2011, **23**:2774–2787.

51. Andriankaja A, Boisson-Dernier A, Frances L, Sauviac L, Jauneau A, Barker DG, de Carvlho-Niebel F: AP2-ERF transcription factors mediate nod factor dependent Mt ENOD11 activation in root hairs via a novel cis-regulatory motif. *Plant Cell* 2007, **19**:2866–2885.

52. Cerri MR, Frances L, Laloum T, Auriac MC, Niebel A, Oldroyd GE, Barker DG, Fournier J, de Carvlho-Niebel F: Medicago truncatula ERN transcription factors: regulatory interplay with NSP1/NSP2 expression dynamics throughout rhizobial infection. *Plant Physiol* 2012, **160**:2155–2172.

53. Murray JD, Muni RR, Torres-Jerez I, Tang Y, Allen S, Andriankaja M, Li G, Laxmi A, Cheng X, Wen J, Vaughan D, Schultze M, Sun J, Charpentier M, Oldroyd G, Tadege M, Ratet P, Mysore KS, Chen R, Udvardi MK: Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of Medicago truncatula. *Plant J* 2011, **65**:244–252.

54. Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi A, Kondorosi E: A novel family in Medicago truncatula consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* 2003, **132**:161–173.

55. Arrighi J-F, Godfroy O, de Billy F, Saurat O, Jauneau A, Gough C: The RPG gene of Medicago truncatula controls Rhizobium-directed polar growth during infection. *Proc Natl Acad Sci U S A* 2008, **105**(28):9817–9822.

56. Oldroyd GED, Downie JA: Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu Rev Plant Physiol Plant Mol Biol* 2008, **59**:519–546.

57. Haney CH, Long SR: Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proc Natl Acad Sci U S A* 2010, **107**:478–483.

58. Mbengue M, Camut S, de Carvlho-Niebel F, Deslandes L, Froidure S, Klaus-Heisen D, Moreau S, Rivas S, Timmers T, Herve C, Cullimore J, Lefebvre B: The Medicago truncatula E3 ubiquitin ligase PUB1 interacts with the LYK3 symbiotic receptor and negatively regulates infection and nodulation. *Plant Cell* 2010, **22**:3474–3488.

59. Fang Y, Hirsch AM: Studying early nodulin gene ENOD40 expression and induction by nodulation factor and cytokinin in transgenic alfalfa. *Plant Physiol* 1998, **116**:53–68.

60. Combier J-P, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernie T, Ott T, Gamas P, Crespi M, Niebel A: MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev* 2006, **20**:3084–3088.

61. Chen L-Q, Qu X-Q, Hou B-H, Sosso D, Osorio S, Fernie AR, Frommer WB: Sucrose efflux mediated by SWEET proteins ad a key step for phloem transport. *Science* 2012, **335**:207–211.

62. Chen L-Q, Hou B-H, Lalonde S, Takanaga H, Hartung ML, Qu X-Q, Guo W-J, Kim J-G, Underwood W, Chaudhuri B, Chermak D, Antony G, White FF, Somerville SC, Mudgett MB, Frommer WB: Sugar transporters for intercellular exhange and nutrition of pathogens. *Nature* 2010, **468**:527–532.

63. de Carvalho GA B, Stefania J, Batista S, Marcelino-Guimaraes FC, Costa do Nascimento L, Hungria M: Transcriptional analysis of genes involved in nodulation in soybean roots inoculated with *Bradyrhizobium japonicum* strain CPAC 15. *BMC Genomics* 2013, **14**:153.

64. Schauser L, Roussis A, Stiller J, Stougaard J: A plant regulator controlling development of symbiotic root nodules. *Nature* 1999, **402**:191–195.

65. Soyano T, Kouchi H, Hirota A, Hayashi M: NODULE INCEPTION directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLoS Genet* 2013, **9**:e1003352.

66. Konishi M, Yanagisawa S: Arabidopsis NIN-like transcription factors have a central role in nitrate signaling. *Nat Commun* 2013, **4**:1617.

67. Capoen W, Oldroyd G: How CYCLOPS keeps an eye on plant symbiosis. *Proc Natl Acad Sci U S A* 2008, **105**:20053–20054.

68. Madsen LH, Trichine L, Jurkiewicz A, Sullivan JT, Heckmann AB, Bek AS, Ronson CW, James EK, Stougaard J: The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat Commun* 2010, **1**. doi:10.1038/ncomms1009.

69. Yano K, Shibata S, Chen WL, Sato S, Kaneko T, Jurkiewicz A, Sandal N, Banba M, Imaizumi-Anraku H, Kojima T, Ohtomo R, Szczyglowski K, Stougaard J, Tabata S, Hayashi M, Kouchi H, Umehara Y: CERBERUS, a novel U-box protein containing WD-40 repeats, is required for formation of the infection thread and nodule development in the legume-rhizobium symbiosis. *Plant J* 2009, **60**:168–180.

70. Krusell L, Krause K, Ott T, Desbrosses G, Kramer U, Sato S, Nakamura Y, Tabata S, James EK, Sandal N, Stougaard J, Kawaguchi M, Miyamoto A, Suganuma N, Udvardi MK: The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *Plant Cell* 2005, **17**:1625–1636.

71. Ott T, van Dongen JT, Gunther C, Krusell L, Desbrosses G, Vigeolas H, Bock V, Czechowski T, Geigenberger P, Udvardi MK: Symbiotic leghemoglobins are crucial for nitrogen fixing in legume root nodules but not for general plant growth and development. *Curr Biol* 2005, **15**:531–535.

72. Blanco L, Reddy PM, Silvente S, Bucciarelli B, Khandual S, Alvarado-Affantranger X, Sanchez F, Miller S, Vance C, Lara-Flores M: Molecular cloning, characterization and regulation of two different NADH-glutamate synthase cDNAs in bean nodules. *Plant Cell Environ* 2008, **3**:454–472.

73. Olivares JE, Diaz-Camino C, Estrada-Navarrete G, Alvarado-Affantranger X, Rodriguez-Kessler M, Samudio FZ, Olamendi-Portugal T, Marquez Y, Servin LE, Sanchez F: Nodulin 41, a novel late nodulin of common bean with peptidase activity. *BMC Plant Biol* 2011, **11**:134.

74. Charpentier M, Bredemeier R, Wanner G, Takeda N, Schleiff E, Parniske M: *Lotus japonicus* CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. *Plant Cell* 2008, **20**:3467–3479.

75. DeYoung BJ, Innes RW: Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat Immunol* 2006, **7**(12):1243–1249.

76. Heyl A, Schmuling T: Cytokinin signal perception and transduction. *Curr Opin Plant Biol* 2003, **6**:480–488.

77. Lohar DP, Schaff JE, Laskey JG, Kieber JJ, Bilyeu KD, Bird DM: Cytokinins play opposite roles in lateral root formation, and nematode and Rhizobial symbioses. *Plant J* 2004, **38**:203–214.

78. Pottosin I, Velarde-Buendia AM, Bose J, Zepeda-Jazo I, Shabala S, Dobrovinskaya O: Cross-talk between reactive oxygen species and polyamines in regulation of ion transport across the plasma membrane: implications for plant adaptive responses. *J Exp Bot* 2014, doi:10.1093/jxb/ert1423.

79. Zepeda-Jazo I, Velarde-Buendia AM, Enriquez-Figueroa R, Bose J, Shabala S, Muniz-Murguia J, Pottosin II: Polyamines interact with hydroxy radicals in activating Ca2+ and K+ transport across the root epidermal plasma membranes. *Plant Physiol* 2011, **157**:2167–2180.

80. Kiba T, Kudo T, Kojima M, Sakakibara H: Hormonal control of nitrogen acquisition: roles of auxin, abscisic acid, and cytokinin. *J Exp Bot* 2011, **62**:1399–1409.

81. Okamoto M, Kumar A, Li W, Wang Y, Yaeesh-Siddiqi M, Crawford NM, Glass ADM: High-affinity nitrate transport in roots of Arabidopsis

depends on expression of the NAR2-like gene AtNRT3.1. *Plant Physiol* 2006, **140**:1036–1046.

82. Miller AJ, Fan X, Orsel M, Smith SJ, Wells DM: **Nitrate transport and signaling.** *J Exp Bot* 2007, **58**:2297–2306.

83. Tsay YF, Chiu CC, Tsai CB, Ho CH, Hsu PK: **Nitrate transporters and peptide transporters.** *FEBS Lett* 2007, **581**:2290–2300.

84. Gojon A, Nacry P, Davidian JC: **Root uptake regulation: a central process for NPS homeostasis in plants.** *Curr Opin Plant Biol* 2009, **12**:328–338.

85. Bouguyon E, Gojon A, Nacry P: **Nitrate sensing and signaling in plants.** *Semin Cell Dev Biol* 2012, **23**:648–654.

86. Gojon A, Krouk G, Perrine-Walker F, Laugier E: **Nitrate transceptor(s) in plants.** *J Exp Bot* 2011, **62**:2299–2308.

87. Krouk G, Lacombe B, Bielach A, Perrine-Walker F, Malinska K, Mounier E, Hoyerova K, Tillard P, Leon S, Ljung K, Zazimalova E, Benkova E, Nacry P, Gojon A: **Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants.** *Dev Cell* 2010, **18**:927–937.

88. Miflin BJ, Habash DZ: **The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops.** *J Exp Bot* 2002, **53**:979–987.

89. McGrath RB, Coruzzi GM: **A gene network controlling glutamine and asparagine biosynthesis in plants.** *Plant J* 1991, **3**:275–280.

90. O'Rourke JA, Yang SS, Miller SS, Bucciarelli B, Liu J, Rydeen A, Bozsoki Z, Uhde-Stone C, Tu ZJ, Allan D, Gronwald JW, Vance CP: **An RNA-Seq transcriptome analysis of orthophosphate-deficient white lupin reveals novel insights into phosphorus acclimation in plants.** *Plant Physiol* 2013, **161**:705–724.

91. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**:R25.

92. Mortazavi A, Williams BA, McCue K, Schaffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5**(7):621–628.

93. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M: **The transcriptional landscape of the yeast genome defined by RNA sequencing.** *Science* 2008, **320**(5881):1344–1349.

94. Zhao J, Oshumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, Kingston RE, Borowsky M, Lee JT: **Genome-wide identification of polycomb-associated RNAs by RIP-seq.** *Mol Cell* 2010, **40**:939–953.

95. Burke GR, Strand MR: **Deep sequencing identifies viral and wasp genes with potential roles in replication of *Microplitis demolitor* bracovirus.** *J Virol* 2012, **86**:3293–3306.

96. Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, Conesa A: **Differential expression in RNA-seq: a matter of depth.** *Genome Res* 2011, **21**:2213–2223.

97. Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible W-R: **Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis.** *Plant Physiol* 2005, **139**(1):5–17.

98. Borges A, Tsai SM, Caldes DGG: **Validation of reference genes for RT-qPCR normalization in common bean during biotic and abiotic stresses.** *Plant Cell Rep* 2012, **31**:827–838.

99. Vance CP, Heichel GH, Barnes DK, Bryan JW, Johnson LE: **Nitrogen fixation, nodule development, and vegetative regrowth of alfalfa (*Medicago sativa* L.) following harvest.** *Plant Physiol* 1979, **64**:1–8.

100. Sbabou L, Bucciarelli B, Miller SS, Liu J, Berhada F, Filali-Maltouf A, Allan D, Vance CP: **Molecular analysis of SCARECROW genes expressed in white lupin cluster roots.** *J Exp Bot* 2010, **61**(5):1351–1363.