CSIRO PUBLISHING
www.publish.csiro.au/journals/fpb *Functional Plant Biology*, 2006, **33**, 783–788

The Affymetrix *Medicago* GeneChip® array is applicable for transcript analysis of alfalfa (*Medicago sativa*)

Mesfin Tesfaye[A], *Kevin A. T. Silverstein*[B], *Bruna Bucciarelli*[D], *Deborah A. Samac*[A,D] *and Carroll P. Vance*[C,D,E]

[A]Department of Plant Pathology, University of Minnesota, Saint Paul, MN 55108, USA.
[B]Department of Plant Biology, University of Minnesota, Saint Paul, MN 55108, USA.
[C]Department of Agronomy and Plant Genetics, University of Minnesota, Saint Paul, MN 55108, USA.
[D]USDA-ARS Plant Science Research Unit, Saint Paul, MN 55108, USA.
[E]Corresponding author. Email: vance004@umn.edu

This paper originates from a presentation at the Third International Conference on Legume Genomics and Genetics, Brisbane, Queensland, Australia, April 2006.

Abstract. The recently released Affymetrix GeneChip® *Medicago* Genome Array contains approximately 52 700 probe sets representing genes in both the model legume *Medicago truncatula* Gaertn. and the closely related crop species *Medicago sativa* L. (alfalfa). We evaluated the utility of the *Medicago* GeneChip® for monitoring genome-wide expression of *M. truncatula* and alfalfa seedlings grown to the first trifoliate leaf stage. We found that approximately 40–54% of the *Medicago* probes were detected in leaf or root samples of alfalfa or *M. truncatula*. Approximately 45–59% of the detected *Medicago* probes were called 'present' in all replicate GeneChips of *Medicago* species, indicating a considerable overlap in the number and type of *Medicago* probes detected between root and leaf organs. Nevertheless, gene expression differences between roots and leaf organs accounted for approximately 17% of the total variation, regardless of the *Medicago* species from which the samples were harvested. The result shows that the *Medicago* GeneChip® is applicable for transcript analysis for both alfalfa and *M. truncatula*.

Keywords: GeneChip, gene expression, medic, *Medicago*, transcript analysis.

Introduction

The plant genus *Medicago* contains approximately 60 species including perennial crops, such as alfalfa (*Medicago sativa*), and the annual species *M. truncatula*. Alfalfa is an allogamous (out-crossing) and an autotetraploid species ($2n = 4\times = 32$) whose cultivars are heterogeneous synthetic populations (Zhu *et al*. 2005). These traits make alfalfa a genetically complex species, recalcitrant to direct genomics studies. By comparison, *M. truncatula* is an autogamous (self-fertilising), diploid species ($2n = 2\times = 16$) with a small genome approximately 470 Mbp. Recent comparative genetic and physical mapping studies have shown a high degree of conservation of gene content and synteny between alfalfa and *M. truncatula* (Choi *et al*. 2004; Zhu *et al*. 2005).

Because of the ease of genetic manipulation and small genome size, *M. truncatula* has become a model species for genomic studies of the Fabaceae. In recent years, international collaborations have led to the development of genomic tools for *M. truncatula* including a large collection (over 226 000) of expressed sequence tags (ESTs) released into the public domain (http://www.tigr.org/tdb/mtgi; verified 23 May 2006) and an ongoing *Medicago* genome sequencing effort (VandenBosch and Stacey 2003 and references therein; Young *et al*. 2005; http://www.medicago.org; verified 23 May 2006). cDNA-based arrays and long oligonucleotide spotted microarrays are also available for the parallel assessment of thousands of genes in a single *Medicago* microarray experiment (Manthey *et al*. 2004; Hohnjec *et al*. 2005; Lohar *et al*. 2006). In addition to spotted arrays, a *Medicago* GeneChip® containing ~10 000 *Medicago* probes was used successfully for gene expression analysis during bacterial symbiosis (Barnett *et al*. 2004;

Abbreviations used: ESTs, expressed sequence tags; MM, mismatch; PM, perfect match.

 10.1071/FP06065 1445-4408/06/080783

Mitra *et al*. 2004). In September 2005, a new *Medicago* GeneChip® was released by Affymetrix for use in whole-genome transcript profiling that includes over 52 000 *Medicago* probe sets designed from 32 167 *M. truncatula* ESTs, 18 733 gene predictions from *M. truncatula* genome sequences and 1896 cDNAs from alfalfa (Affymetrix). The Affymetrix *Medicago* GeneChip® also contains 8305 gene predictions from the genome of the microsymbiont *Sinorhizobium meliloti*.

Expression profiling when combined with sequence similarity analysis can assist in the identification of candidate genes for particular traits worthy of further investigation and hypothesis building. Microarrays can also be used for the parallel analysis of gene expression in plant species that have limited genomic information available, thereby elucidating commonalities and differences highlighted by the cross-species comparisons (Horvath et al. 2003; Close *et al*. 2004; Hammond *et al*. 2006). Using *Arabidopsis thaliana* L. GeneChips, cross-species microarray experiments have been used successfully to identify genes from related plant species involved in several complex traits (Becher *et al*. 2004; Weber *et al*. 2004). Recently, a barley chip microarray analysis applied to wheat–barley chromosome addition lines facilitated a large-scale physical mapping of barley genes (Cho *et al*. 2006).

Alfalfa is grown as a forage crop on 32 million hectares throughout the world. Non-traditional uses of alfalfa include phytoremediation of contaminated soils (Wang *et al*. 2005), production of recombinant enzymes and feedstocks for industrial applications (Saruul *et al*. 2002; Tesfaye *et al*. 2005) and for human consumption as a source of flavonoids in health food products, in sprouts and juice. Alfalfa is also being developed as a biomass feedstock for energy production. While alfalfa is an extremely important crop worldwide, its genetic complexity has hindered progress in understanding the genes controlling important agronomic traits. The high degree of DNA sequence identity and conserved gene order between *M. truncatula* and alfalfa suggest that the Affymetrix *Medicago* GeneChip® would be a useful tool for understanding gene function and expression in alfalfa. By undertaking a comparative transcript analysis of *M. truncatula* and alfalfa leaves and roots, the aim of this study was to assess whether or not the Affymetrix *Medicago* GeneChip® array would be useful for transcript studies of alfalfa. Here we provide data that show the recently released *Medicago* GeneChip® array is suitable for transcript profiling of both model and crop species, and will aid in the discovery of agronomically important genes in alfalfa in particular, thereby helping to overcome the genetic complexities and limitations of this crop.

Materials and methods

Medicago truncatula Gaertn. and alfalfa (*Medicago sativa* L.) seedlings were grown to the first trifoliate leaf stage and total RNA was extracted from the first trifoliate leaf and whole root using the RNeasy Plant RNA mini kit (Qiagen, Valencia, CA). Tissues were harvested from three experimental replicates with up to 30 plants in each replicate. Extracted RNA was treated with DNase I enzyme to remove any contaminating DNA. Total RNA (10 µg) was used to produce biotin-labelled cRNA with Affymetrix kits and following suggested procedures for eukaryotic reactions (Affymetrix, Santa Clara, CA). Biotin-labelled cRNA (15 µg) fragmented as suggested by Affymetrix was hybridised to the GeneChip® *Medicago* Genome Array. The integrity and quality of total RNA, labelled and fragmented biotin-labelled cRNA was verified with the Agilent 2100 Bioanalyser in RNA 6000 Nano LabChip (Agilent Technologies, Palo Alto, CA). GeneChips were hybridised, washed, stained and scanned by the University of Minnesota Affymetrix core facility.

Probe sets in the *Medicago* GeneChip® array are composed of 11 pairs of 25-mer oligonucleotides, with each pair consisting of a perfect match (PM) oligonucleotide and a mismatch (MM) control containing a single nucleotide substitution at the thirteenth base position (Affymetrix). The purpose of MM controls is to help distinguish background non-specific hybridisation from true low-level expression of genes (Affymetrix). For each probe set, we calculated both its detection call ('present' *v.* 'absent') and its 'expression value', i.e. a number that reflects the concentration of the probe set's target RNA species. For detection calls, we used the default MAS 5.0 parameters except changing $\alpha_1 = 0.05$ as suggested for 11 probe pairs per probe set by Liu *et al*. (2002). Microarray data analysis was also conducted using Genedata Expressionist® Pro version 1.0 (Genedata, San Francisco, CA). Signal intensity values were normalised by global median scaling. In all of the data analyses, signals corresponding to *Sinorhizobium* probes were excluded.

Results

Our laboratory has cloned and characterised several nitrogen assimilation genes from alfalfa nodules. Using homologous gene sequences in the public domain for *M. truncatula* and other leguminous plants, we performed pair-wise comparisons of nucleotide sequence identities of five nitrogen assimilation genes between *M. truncatula* and other leguminous plants including alfalfa (Table 1). We observed 93% or more DNA sequence identity between protein coding regions of selected homologous genes in alfalfa and *M. truncatula* (Table 1). Percent nucleotide identity between *M. sativa* and *M. truncatula* of five selected genes involved in nitrogen assimilation was 97.5%

Table 1. Percent nucleotide identity of genes involved in nitrogen assimilation between *Medicago truncatula* and related legumes

AAT, aspartate amino transferase; PEPC, phosphoenol pyruvate carboxylase; MDH, malate dehydrogenase; AS, asparagine synthetase; GS, glutamine synthetase

	Medicago truncatula				
	AAT	PEPC	MDH	AS	GS
Medicago sativa	97.5	97.9	93.4	99.8	98.6
Pisum sativum	90.1	91.9	92.1	93.0	89.0
Lotus japonicus	87.6	86.6	–	89.4	79.3
Glycine max	86.1	84.7	86.3	86.5	87.7
Phaseolus vulgaris	85.5	85.8	92.1	84.7	88.0
Lupinus albus	86.3	84.9	72.2	90.7	86.5

for aspartate amino transferase, 97.9% for phosphoenol pyruvate carboxylase, 93.4% for malate dehydrogenase, 99.8% for asparagine synthetase and 98.6% for glutamine synthetase (Table 1).

In an effort to evaluate the utility of the GeneChip® for *Medicago* studies, we performed three GeneChip® hybridisation replicate experiments for leaf and root tissue using RNA extracted from *Medicago* seedlings for each replicate, resulting in a total of 12 GeneChip hybridisation experiments. Signal intensity values of the experimental replicates derived from the same plant tissues were assessed for variability of gene expression data. Strong correlation coefficients ($R^2 > 0.92$) were obtained among all the experimental replicates for all plant tissues, indicating that three experimental replicates provide adequate statistical power to allow an easy identification of differentially expressed probe sets (Fig. 1).

With regard to *Medicago* GeneChip® probe sets that successfully hybridised to each *Medicago* species and tissue category, approximately 54% (28 668 probe sets) of the over 52 000 plant gene probes of the *Medicago* GeneChip® produced 'present' calls when hybridised with biotin-labelled cRNA from *M. truncatula* root. An average of 46% (24 371 probe sets) of the plant genes probe sets also were detected in first trifoliate samples of *M. truncatula*. In comparison, 44% (23 202 probe sets) and 41% (21 526 probe sets) of the total plant gene probes were called 'present' in alfalfa root and first trifoliate, respectively. The detected probe sets in alfalfa tissues correspond to 88 and 81% of the baseline level of detected probe sets using the *M. truncatula* first trifoliate and root samples, respectively.

One distinct feature of the *Medicago* GeneChip® over other *Medicago* microarray platforms is the presence of probes for over 1800 cDNAs from alfalfa. Approximately 75 and 77% of the alfalfa probe sets included in the GeneChip® were detected in *M. truncatula* first trifoliate leaf and root samples, respectively (Fig. 2), whereas approximately 86 and 83% of the alfalfa probe sets included in the GeneChip® were called 'present' in alfalfa first trifoliate and root samples, respectively (Fig. 2). In all, 1091 (61%) alfalfa probe sets were called 'present' for leaf and root samples of both *Medicago* spp.

It is anticipated that transcriptional changes will accompany morphological changes during the course of new organ and tissue development in plants. The MAS 5.0 expression data containing presence and absence calls was used to determine the conservation of detected probe sets or to identify differentially regulated probe sets among first trifoliate and root tissues of *Medicago* spp. We found that approximately15 000 probe sets were called 'present' in leaf tissues of both alfalfa and *M. truncatula* (Fig. 3*A*). In comparison, approximately 17 780 probe sets were called 'present' in root tissue of both alfalfa and *M. truncatula* (Fig. 3*B*). Most (12 770 probe sets) of the probe sets were called 'present' consistently in all replicates of both leaf and root samples (Fig. 3*C*). Despite the large overlap in gene expression between roots and leaves, unique gene expression patterns were also found in leaves and roots of both species. Of the 612 probes detected only in leaf tissue, 196 probes were unique to *M. truncatula* while 50 were unique to alfalfa leaf samples (Fig. 3). It was also found that 941 and 73 probe sets

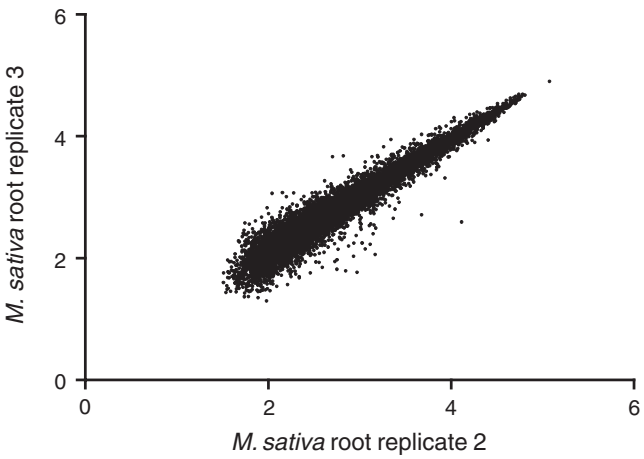

Fig. 1. Scatter plot of log-transformed signal intensities of GeneChip® *Medicago* Genome Array data. The figure illustrates the variation between experimental replicates. Signal intensity values from probe sets called 'present' by the Affymetrix Microarray Suite 5.0 (MAS 5.0) algorithm (Liu *et al*. 2002) were used to construct the scatter plot.

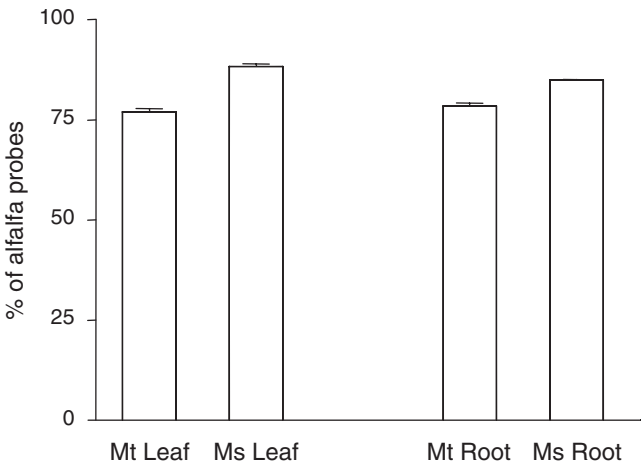

Fig. 2. Percentage of detected alfalfa probes on the GeneChip® *Medicago* Genome Array. A probe set was called 'present' when at least eight of the 11 PM and MM probe pairs within a probe set exceed the probe pair threshold parameters set at $\tau = 0.015$, $\alpha 1 = 0.05$ and $\alpha 2 = 0.065$, as determined by the Affymetrix MAS 5.0 (Liu *et al.* 2002). The first trifoliate leaf and whole-root tissue from *Medicago truncatula* (Mt) and alfalfa (Ms) seedlings were used for gene expression analysis.

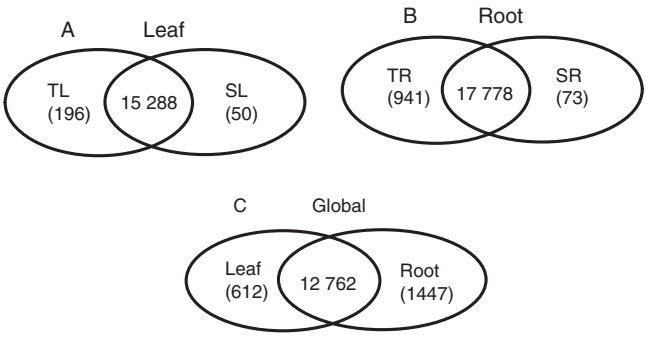

Fig. 3. Comparisons of detected *Medicago* GeneChip® probes between alfalfa and *M. truncatula* (*A*) leaf, (*B*) root as well as (*C*) global comparison between leaf and root. Numbers in parentheses indicate probes unique to an organ. Abbreviations are: TL, *M. truncatula* leaf; TR, *M. truncatula* root; SL, *M. sativa* leaf; SR, *M. sativa* root.

were unique to *M. truncatula* and alfalfa roots, respectively (Fig. 3). Of the 50 probes unique to the alfalfa trifoliate leaf, 18 were derived from alfalfa sequences. Similarly, of the 73 unique probes in alfalfa roots, 10 were derived from alfalfa sequences, suggesting that these may be species-specific transcripts.

We utilised principal component analysis to construct a graphical interpretation of gene expression data, shown in Fig. 4. The first two principal components explained approximately 50% of the total gene expression variation. Along the second principal component axis, which explained 17% of the total gene expression variation, gene expression of roots was distinctly separated from first trifoliate leaf regardless of the *Medicago* species from which the samples were harvested.

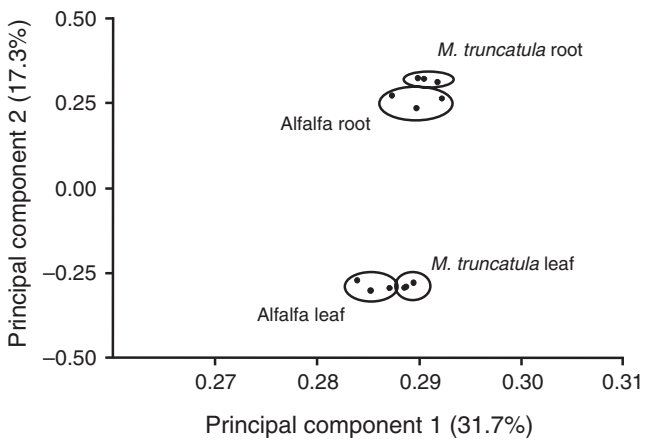

Fig. 4. Principal component analysis of gene expression data from first trifoliate and root samples of *Medicago* species. The first and second principal components explained 49% of the total gene expression variation. The percentages show the variation explained by each principal component.

Discussion

The Affymetrix GeneChip® *Medicago* genome array is a very recent, largely untested, addition to the genomics tools available for *Medicago* researchers. In this report, we have provided evidence that the GeneChip® is an effective platform for screening the transcriptomes of *M. truncatula* or alfalfa by demonstrating that (a) selected genes between alfalfa and *M. truncatula* showed between 93.4–99.8% nucleotide identity, (b) the total numbers and types of probe sets detected in the *Medicago* GeneChip® were comparable between alfalfa and *M. truncatula*, and (c) considerable gene expression differences were observed between root and leaf tissues, regardless of the *Medicago* species from which the samples were derived. Results also demonstrate that cross-species comparison of gene expression is possible with the *Medicago* GeneChip®.

The GeneChip® provides sensitive and reproducible ($R^2 > 0.92$) detection of the transcriptomes of both *Medicago truncatula* and alfalfa. Depending on the tissue surveyed, we detected 46–54% of the more than 52 000 plant probe sets for *M. truncatula*, and 41–44% in alfalfa. These percentages are comparable to, or higher than, those reported previously by other workers for other cross-species microarray hybridisations (Becher *et al*. 2004; Close *et al*. 2004; Weber *et al*. 2004). The relatively high numbers of 'present' calls (73–88% of the 1896 alfalfa probes) for leaf and root samples of both *Medicago* species are consistent with observations of very high nucleotide similarities between the coding regions of the *M. truncatula* and alfalfa genes. Close and associates (2004) examined gene expression of first leaf stage green seedlings of barley, wheat, oat, rice and sorghum with the 22K Barley1 GeneChip. On average, 45% (barley), 25% (wheat), 12% (oat), 9% (rice), 8% (sorghum) and 6% (maize) of the 22K Barley1 GeneChip probe array showed 'present' calls (Close *et al*. 2004). The percent of detected probe sets was increased to 58% when pathogen-challenged barley leaf tissue was used (Close *et al*. 2004). In another study, an *Arabidopsis* cDNA microarray was used in a broad cross-species comparison with wild oat (*Avena fatua* L.), poplar (*Populus deltoides* Bartram ex Marshall), and leafy spurge (*Euphorbia esula* L.) to identify genes involved in shoot growth (Horvath *et al*. 2003). Over 23, 34 and 47% of the 11 522 *Arabidopsis* probes produced signals that were greater than the threshold when hybridised with wild oat, poplar and leafy spurge cRNA, respectively, (Horvath *et al*. 2003).

Despite a much closer agreement in the number and types of detected probes between *M. truncatula* and alfalfa, the number of expressed genes in roots was higher than in leaves. Similar observations were made in *Arabidopsis* by Schmid *et al*. (2005). Gene expression profiles of roots and leaves in *Arabidopsis* were clearly distinct, with gene expression in root tissues showing generally higher transcript

abundance than in leaves (Schmid *et al*. 2005). Consistent with our observation of putative organ-specific or species-specific probe sets in alfalfa or *M. truncatula*, Schmid *et al*. (2005) also reported specific marker genes for the major organs of *Arabidopsis*. Additional data analysis or evaluation by a more sensitive method such as quantitative PCR (qPCR) analysis is needed to confirm putative marker probes sets in *Medicago* species.

Cross-species microarray transcript profiling with the *A. thaliana* GeneChip was used successfully to identify candidate genes involved in zinc tolerance in *Arabidopsis halleri* (L.) O'Kane & Al-Shebaz, a naturally selected zinc hyper-accumulator (Becher *et al*. 2004; Weber *et al*. 2004). *A. thaliana* shares approximately 94% identity with *A. halleri* at the nucleotide level within coding regions. This level of nucleotide similarity is comparable to what we observed for *M. truncatula* and alfalfa. Weber *et al*. (2004) reported that 55 and 26% of the probes on the *Arabidopsis* GeneChip were active in untreated plant roots of *A. thaliana* and *A. halleri*, respectively. Following zinc treatment, 63.9 and 36.8% of the total probes on the chip were shown to be active in leaves of *A. thaliana* and *A. halleri*, respectively (Becher *et al*. 2004). Candidate gene family members identified by transcript profiling were subsequently shown to confer enhanced heavy metal tolerance to transformed yeast (Becher *et al*. 2004; Weber *et al*. 2004). Our results also suggest that genome-wide transcript profiling with the GeneChip® *Medicago* genome array will aid in the identification of biologically meaningful gene expression patterns in *Medicago* spp. and will facilitate gene functional discovery in the cultivated crop, *M. sativa*.

It is estimated that only 10% of the gene annotation in *Arabidopsis* genome is based on experimental data on gene or protein function (Brown *et al*. 2005). The majority of genes (62%) are annotated based on conserved domains or assigned to a gene family based on homology that clusters the genes into broad categories. Currently, the *Medicago* GeneChip® lacks functional annotation. Future *Medicago* genomics efforts should include such an undertaking to enhance the usability of gene expression data.

References

Becher M, Talke IN, Krall L, Krämer U (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyper-accumulator *Arabidopsis halleri*. *The Plant Journal* **37**, 251–268.

Barnett MJ, Tolman CJ, Fisher RF, Long SR (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proceedings of the National Academy of Sciences USA* **101**, 16 636–16 641. doi: 10.1073/pnas.0407269101

Brown DM, Zeef LAH, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *The Plant Cell* **17**, 2281–2295. doi: 10.1105/tpc.105.031542

Cho S, Garvin DF, Muehlbauer GJ (2006) Transcriptome analysis and physical mapping of barley genes in wheat–barley chromosome addition lines. *Genetics* **172**, 1277–1283. doi: 10.1534/genetics.105.049908

Choi HK, Kim D, Uhm T, Limpens E, Lim H, *et al*. (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker co-linearity with *Medicago sativa*. *Genetics* **166**, 1463–1502. doi: 10.1534/genetics.166.3.1463

Close TJ, Wanamaker SI, Caldo RA, Turner SM, Ashlock DA, Dickerson JA, Wing RA, Muehlbauer GJ, Kleinhofs A, Wise RP (2004) A new resource for cereal genomics, 22K barley genechip comes of age. *Plant Physiology* **134**, 960–968. doi: 10.1104/pp.103.034462

Hammond JP, Bowen HC, White PJ, Mills V, Pyke KA, Baker AJM, Whiting SN, May ST, Broadley MR (2006) A comparison of the *Thlaspi caerulescens* and *Thlaspi arvense* shoot transcriptomes. *New Phytologist* **170**, 239–260. doi: 10.1111/j.1469-8137.2006.01662.x

Hohnjec N, Vieweg MF, Pühler A, Becker A, Küster H (2005) Overlaps in the transcriptional profiles of *Medicago truncatula* roots inoculated with two different *Glomus* fungi provide insights into the genetic program activated during arbuscular mycorrhiza. *Plant Physiology* **137**, 1283–1301. doi: 10.1104/pp.104.056572

Horvath DP, Schaffer R, West M, Wisman E (2003) *Arabidopsis* microarrays identify conserved and differentially expressed genes involved in shoot growth and development from distantly related plant species. *The Plant Journal* **34**, 125–134. doi: 10.1046/j.1365-313X.2003.01706.x

Liu WM, Mei R, Di X, Tyder TB, Hubbell E, *et al*. (2002) Analysis of high density expression microarrays with signed-rank call algorithms. *Bioinformatics* **18**, 1593–1599. doi: 10.1093/bioinformatics/18.12.1593

Lohar DP, Sharopova N, Endre G, Penuela S, Samac D, Town C, Silverstein KAT, VandenBosch KA (2006) Transcript analysis of early nodulation events in *Medicago truncatula. Plant Physiology* **140**, 221–234. doi: 10.1104/pp.105.070326

Manthey K, Krajinski F, Hohnjec N, Firnhaber C, Puhler A, Perlick AM, Kuster H (2004) Transcriptome profiling in root nodules and arbuscular mycorrhiza identifies a collection of novel genes induced during *Medicago truncatula* root endosymbioses. *Molecular Plant–Microbe Interactions* **17**, 1063–1077.

Mitra RM, Shaw SL, Long SR (2004) Six nonnodulating plant mutants defective for Nod factor-induced transcriptional changes associated with the legume–rhizobia symbiosis. *Proceedings of the National Academy of Sciences USA* **101**, 10 217–10 222. doi: 10.1073/pnas.0402186101

Saruul P, Sreinc F, Somers DA, Samac DA (2002) Production of a biodegradable plastic polymer poly-β-hydroxybutyrate in transgenic alfalfa (*Medicago sativa* L.). *Crop Science* **42**, 919–927.

Schmid M, Davidson TS, Henz SR, Pape TJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics* **37**, 501–506. doi: 10.1038/ng1543

Tesfaye M, Denton MD, Samac DA, Vance CP (2005) Transgenic alfalfa secretes a fungal endochitinase protein to the rhizosphere. *Plant and Soil* **269**, 233–243. doi: 10.1007/s11104-004-0520-0

VandenBosch KA, Stacey G (2003) Summaries of legume genomics projects from around the globe. Community resources for crops and models. *Plant Physiology* **131**, 840–865. doi: 10.1104/pp.103.020388

Wang L, Samac DA, Shapir N, Wackett LP, Vance CP, Olszewski NE, Sadowsky MJ (2005) Biodegradation of atrazine in transgenic plants expressing a modified bacterial atrazine chlorohydrolase (atzA) gene. *Plant Biotechnology Journal* **3**, 475–486. doi: 10.1111/j.1467-7652.2005.00138.x

Weber M, Harada E, Vess C, Roepenack-Lahaye EV, Clemens S (2004) Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identified nicotianamine synthase a ZIP transporter and other genes as potential metal hyperaccumulation factors. *The Plant Journal* **37**, 269–281. doi: 10.1111/j.1365-313X.2003.02013.x

Young ND, Cannon SB, Sato S, Kim DJ, Cook DR, Town CD, Roe BA, Tabata S (2005) Sequencing the genespaces of *Medicago truncatula* and *Lotus japonicus*. *Plant Physiology* **137**, 1174–1181. doi: 10.1104/pp.104.057034

Zhu H, Choi H-K, Cook DR, Shoemaker RC (2005) Bridging model and crop legumes through comparative genomics. *Plant Physiology* **137**, 1189–1196. doi: 10.1104/pp.104.058891

Manuscript received 29 March 2006, accepted 28 April 2006