Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Plant Biology

Translating *Medicago truncatula* genomics to crop legumes

Nevin Dale Young[1,2] and Michael Udvardi[3]

Genomic resources developed in the model legume, *Medicago truncatula*, have the potential to accelerate practical advances in crop legumes. *M. truncatula* is closely related to many economically important legumes, frequently displaying genome-scale synteny. Translating genome data from *M. truncatula* should be highly effective in marker development, gene discovery, and positional cloning in crop legumes. The *M. truncatula* genome sequence also provides valuable insights about gene families of practical importance, especially those that are legume-specific. The *M. truncatula* genome sequence should also simplify the assembly of next-generation sequence data in closely related taxa, especially alfalfa. Genomic resources, such as whole-genome arrays, make it possible to pursue detailed questions about gene expression in both *M. truncatula* and related crop species, while tagged mutant populations simplify the process of determining gene function.

Addresses
[1] Department of Plant Pathology, 495 Borlaug Hall, University of Minnesota, St. Paul, MN 55108, USA
[2] Department of Plant Biology, 220 BioScience Building, University of Minnesota, St. Paul, MN 55108, USA
[3] Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

Corresponding author: Young, Nevin Dale (neviny@umn.edu)

Current Opinion in Plant Biology 2009, **12**:193–201

This review comes from a themed issue on
Genome studies and molecular genetics
Edited by Masahiro Yano and Roberto Tuberosa

Available online 21st January 2009

1369-5266/$ – see front matter

DOI 10.1016/j.pbi.2008.11.005

Introduction

Model species are valuable not only because they lead to discoveries about basic biology, but also because they provide resources that facilitate improvements in crops of economic importance. Scientists are drawn to models because of their ease of manipulation, simple genome organization and patterns of inheritance, rapid life cycle, or other desirable experimental properties. In legumes, *Medicago truncatula* is one of the two models with abundant genomic tools and a vibrant international research community (the other model is *Lotus japonicus* [1•]). With the recent development of a comprehensive gene expression atlas [2••], large-scale mutant knockout populations [3••], and the imminent release of its genome sequence (www.medicago.org/genome), *M. truncatula* is poised to provide a wealth of genomic discoveries to the applied legume research community.

M. truncatula as a model legume

M. truncatula is a galegoid legume, closely related to alfalfa (*M. sativa*). It originated in the Mediterranean basin and is now cultivated as an annual forage, especially in Australia. As a cool season legume, it is related to important crops like red and white clover (*Trifolium pratense* and *T. repens*, respectively), pea (*Pisum sativum*), chickpea (*C. arietinum*), lentil (*Lens culinaris*), and broad bean (*V. faba*). These genera are thought to have separated from *M. truncatula* within the last 25 Mya [4]. *M. truncatula* is also phylogenetically close to *L. japonicus* (estimated separation time, 50 Mya [4]), as well as soybean (*Glycine max*) and common bean (*Phaseolus vulgaris*) (estimated separation time, 60 Mya) [4]. Interest in *M. truncatula* as a legume model began with research into its symbiotic relationship with *Sinorhizobium meliloti* in the early 1990s. The scope of *M. truncatula* research later expanded to the study of mycorrhizal symbioses, root development, secondary metabolism, disease resistance, and ultimately to genomics, transcriptomics, proteomics, and metabolomics [2••,5•,6••]. Among the many reasons *M. truncatula* was chosen as a model legume are its modest genome size of 500–550 million base pair (Mbp), simple diploid genetics, short seed-to-seed generation time, workable levels of transformation, excellent mutant populations, and large collections of diverse ecotypes. A highlight of model legume research (including both *M. truncatula* and *L. japonicus*) has been the elucidation of key factors in the rhizobium recognition and nodulation signal transduction pathway (reviewed in [7]).

A growing array of genomic tools are being developed for *M. truncatula*. There are nearly 250 000 expressed sequence tags (ESTs) deposited in Genbank (October 2008) and an Affymetrix® GeneChip with approximately 51 000 probe sets for *M. truncatula* genes. This genechip forms the basis for a recently developed gene expression atlas [2••]. For reverse genetics, *M. truncatula* has large and characterized *Tnt1* [3••], fast neutron bombardment (FNB) [8], and targeting-induced local lesions in genomes (TILLING) (Cook, unpublished) populations available, as well as RNA interference (RNAi)-based gene silencing [9]. Finally, *M. truncatula* is the target of an international BAC-by-BAC genome sequencing initiative, scheduled for completion in 2009. These resources are complemented by an extensive array of websites and bioinformatic tools (Table 1).

Table 1

Medicago and related legume genome websites.

Medicago genome sequence sites	
Medicago Genome Sequence Resources	http://www.medicago.org
TIGR *Medicago* Database	http://www.tigr.org/tdb/e2k1/mta1/
University of Oklahoma Sequencing Project	http://www.genome.ou.edu/medicago.html
European *Medicago* Genome Database	http://mips.gsf.de/proj/plant/jsf/medi/index.jsp
Medicago functional genomics sites	
Noble Foundation *Medicago* Gene Expression Atlas	http://bioinfo.noble.org/gene-atlas/
Noble Foundation *Medicago* Mutant Database	http://bioinfo4.noble.org/mutant/
Noble Foundation *Medicago* Metabolic Pathways	http://mediccyc.noble.org/
Medicago Handbook	http://www.noble.org/MedicagoHandbook/
Comparative legume genomics sites	
Legume Information System (LIS)	http://www.comparative-legumes.org/
INRA/CNRS Integrative Legume Resources	http://www.legoo.org/
Legume Genome Website Links at LIS	http://www.comparative-legumes.org/lis/lis_links.html
Other legume genome sites	
Lotus japonicus Genome Browser	http://www.kazusa.or.jp/lotus/
Lotus japonicus Legume Base	http://www.shigen.nig.ac.jp/bean/lotusjaponicus/top/top.jsp
Phytozome *Glycine max* Genome	http://www.phytozome.net/soybean
Soybase and Soybean Breeders Toolbox	http://soybase.org/
Tree of Life: Fabaceae	http://www.tolweb.org/Fabaceae
International Legume Database	http://www.ildis.org/

The *Medicago* genome sequencing initiative

The decision to sequence *M. truncatula* stemmed in large part from the observation that its chromosomes were organized into gene-rich euchromatin separate from gene-poor heterochromatin [10]. This observation is borne out by the most recent release of the genome sequence (January 2008), which consisted of a sequence assembly 210 Mbp in size (60–70% of the estimated euchromatin) capturing approximately 60% of all genes (medicago.org/genome). At completion, the *Medicago* Genome Sequencing Consortium anticipates an assembly of approximately 300 Mbp, covering nearly all the euchromatin and capturing ~90% of all genes.

To be useful as a model for legume genomics, it is essential that *M. truncatula* exhibits genome conservation with other crop legumes. A growing number of studies demonstrate large-scale synteny among legumes. In a detailed comparison between *M. truncatula* and alfalfa, marker relationships were uniformly syntenic [10]. *M. truncatula* also exhibits nearly genome-wide colinearity with clover [11], pea [12], and lentil [13]. These are all members of the inverted-repeat lacking clade (IRLC), which is estimated to have descended from a common ancestor between 18 and 30 Mya [4]. *M. truncatula* also displays extensive synteny with *L. japonicus* and *G. max* at both macrosyntenic and microsyntenic scales [10,14••]. In comparison with *L. japonicus*, for example, 10 large-scale synteny blocks span ~75% of both genomes [14••]. Together with the recent publication of an extensive sequence for *L. japonicus* [1•] and the anticipated whole-genome shotgun sequence for *G. max* (http://www.phytozome.net/soybean), genome-scale comparisons will soon be possible, revealing the key genome features of the common ancestor for contemporary legume crops.

Practical applications of the *M. truncatula* genome

The vast amount of sequence data available in *M. truncatula* make it an excellent resource for translational genomics. Of course, the *M. truncatula* genome sequence provides a ready source for easily accessible DNA markers across legume species, enabling pan-legume comparisons, facilitating linkage mapping, comparative genomics, and marker-assisted selection [10]. The *M. truncatula* genome also reveals the scale and scope of biologically important gene families, including disease resistance genes [15] and legume-specific and legume-over-represented families [16,17•]. Analysis of upstream regions identifies regulatory sequences, which can be mined and utilized to fine-tune gene expression as part of genetic engineering. The *M. truncatula* sequence creates a platform for gene discovery and positional cloning. Finally, the essentially complete *M. truncatula* genome sequence accelerates physical mapping in related species and provides a scaffold for next-generation sequencing in close relatives like alfalfa.

DNA marker development

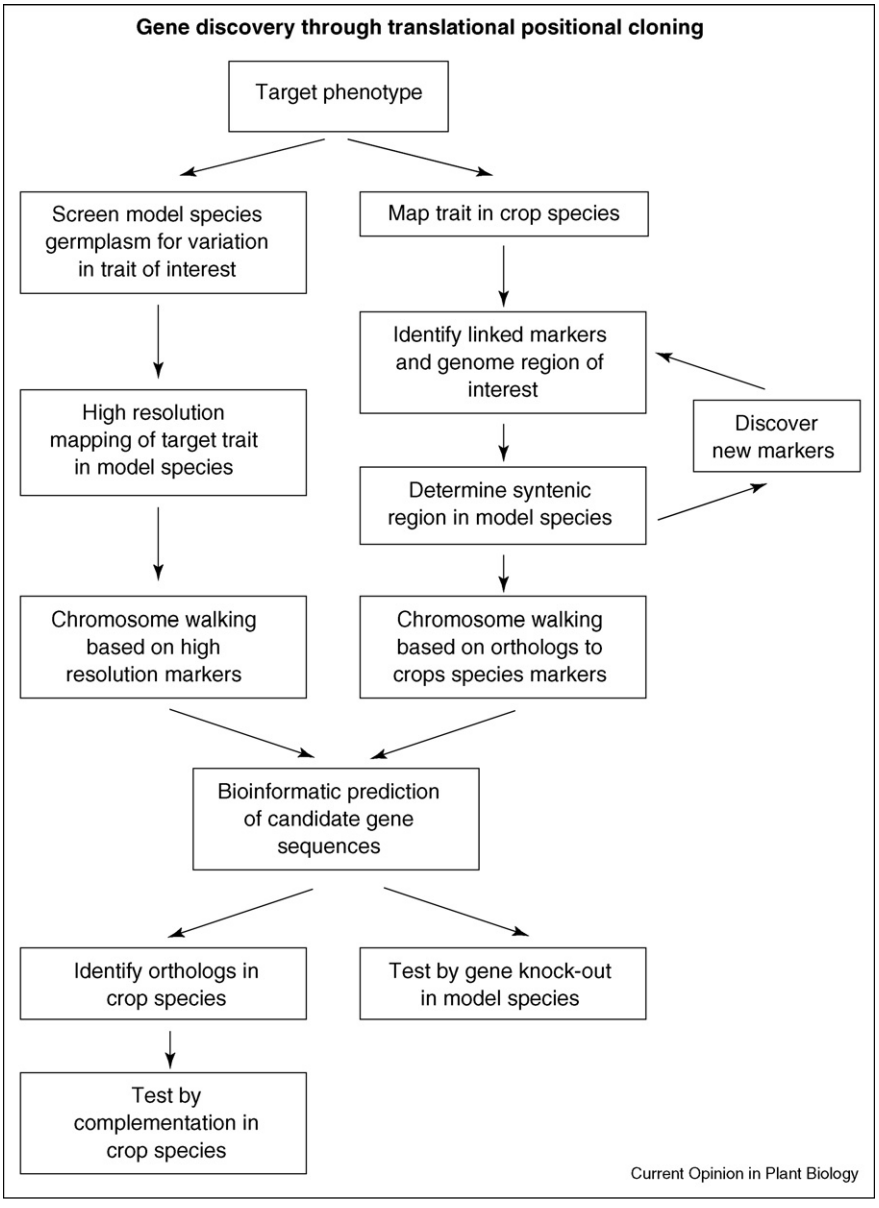
M. truncatula genome and EST sequence resources are already widely exploited for DNA marker discovery. The works of Choi *et al.* [10,18] illustrate the use of *M. truncatula* EST and bacterial artificial chromosome (BAC) sequence data to discover cross-species markers as a basis for creating linkage maps in legume species. These studies emphasized the use of intron-targeted markers, gene-based PCR products designed to amplify

from one exon to the next while spanning intron borders. Intron-targeted markers tend to reveal higher levels of polymorphism and are thereby more useful in cross-species mapping [18]. *M. truncatula*-based intron-spanning markers have been applied to mapping in *L. culinaris* [13]. *M. truncatula* sequence data have also been used to develop cross-species simple sequence repeat (SSR) markers. For example, *M. truncatula*-based SSRs provided genetic markers for linkage mapping in both alfalfa [19] and more distantly related crop legumes [11,20]. Within the *Medicago* genus, nearly 70% of SSRs predicted based on *M. truncatula* ESTs were positive in other *Medicago* species [19]. With more distantly related legume relatives like *Vicia faba* (faba bean), chickpea, and pea, transferability of EST-based SSRs ranged from 39 to 43% [20], although [11] reported success of only 20% when transferring *M. truncatula* SSRs to *Trifolium*.

Gene discovery and positional cloning

Another opportunity arising from the genome sequence of *M. truncatula* is targeting genes underlying specific phenotypes. There are two strategies for accomplishing this goal (Figure 1). One approach, exemplified by disease resistance, is where phenotypes of interest in crop

Figure 1



Gene discovery through translational positional cloning. Key steps in the use of a model species to clone a gene of interest based on genome location (see text for details).

relatives are mapped and characterized within *M. truncatula* using the genomic resources available there. This strategy begins with a search for phenotypic variation in *M. truncatula*. Most legume phenotypes should be amenable to this approach, but it is important to remember that some phenotypes, such as perenniality, simply do not exist in *M. truncatula*. This is a trait of major importance related to winter survival in alfalfa — but it is also apparently lacking in *M. truncatula*. The second approach is where a trait of interest is mapped within the crop plant and then the corresponding region of the *M. truncatula* genome is identified through microsynteny, enabling candidate genes to be cloned in the model.

Examples of phenotype mapping in *M. truncatula* range from disease resistance genes to quantitative trait loci (QTL) for morphology, development, and metabolism. Resistance against *Colletotrichum trifolii* (anthracnose) [21,22••], *Phoma medicaginis* (black stem and leaf spot) [23], and *Erysiphe pisi* (powdery mildew) [21] have all been mapped in *M. truncatula*. Some of these resistance loci have been positioned at high resolution and candidate nucleotide binding site-leucine rich repeat (NBS-LRR) gene sequences described [21,22••]. Indeed, the cloned *RCT1* gene of *M. truncatula* has been shown to function in alfalfa [22••].

By contrast, it should also be possible to clone genes of interest originally mapped in a crop legume species and then targeted through synteny with the *M. truncatula* genome sequence. An excellent example is *SYM2* of pea, whose gene product contributes to recognition and specificity of Nod factors [24]. Here, the original gene was mapped and characterized in pea, microsyntenic regions in *M. truncatula* were identified [25], and finally cloned by chromosome walking on *M. truncatula* BACs. A similar story comes from the cloning of nodulation receptor kinase (*NORK*) [26], another gene product involved in Nod factor recognition. Today, there are numerous traits being mapped in legume relatives of *M. truncatula* with the eventual goal of utilizing the *M. truncatula* sequence to clone the underlying gene(s). Examples include aluminum tolerance [27], seed morphology [28], as well as multiple examples of disease resistance.

Physical mapping and whole-genome sequencing

Just as the *M. truncatula* sequence can be used to uncover DNA markers in crop legumes, it can also provide a framework for physical mapping. This is best illustrated by the 'OMAP' project, where the rice genome is being used as a scaffold for the construction of genome-wide physical maps in 12 *Oryza* relatives [29]. Here, the idea is to carry out deep BAC-end sequencing combined with DNA fingerprinting in related rice species and then align the maps to the completed rice genome. The aligned physical maps can be populated with different types of genomic data as a framework for translational genomics. Now this approach is feasible for close relatives of *M. truncatula* and initial steps have been taken for *M. truncatula* and red clover, *T. pratense* (M. Abberton and G. Oldroyd, unpublished).

The capacity to act as a genome scaffold is not limited to physical mapping. With the advent of next-generation sequencing like '454', Illumina, and ABI SOLiD, massive genome sequencing of alfalfa (*M. sativa*) followed by alignment to the *M. truncatula* genome sequence is becoming realistic. As such, alfalfa sequencing would become an example of cross-species resequencing. Resequencing is frequently carried out to discover SNP markers by alignment to an already sequenced reference. In the case of alfalfa, the goal would be quite different, relying on sufficient sequence conservation to order and orient millions of reads resulting from next-generation sequencing. The fact that *M. truncatula* and *M. sativa* are so closely related makes this application of translational genomics especially compelling. Nevertheless, important questions remain. Will it be possible to unambiguously align relatively short sequence reads to a heterologous genome sequence? Will the fact that alfalfa is a heterozygous allotetraploid with its own allelic diversity lead to sequence data that is even more difficult to align? In any case, the prospect of quickly constructing a whole-genome shotgun sequence for alfalfa is incredibly exciting.

M. truncatula functional genomics

Although genome sequence reveals the RNA-coding and protein-coding *potential* of an organism, additional data are required to determine how much of this potential is realized and what biological function(s) gene products fulfil. To answer these questions, high-throughput methods that identify and measure most of or all RNA transcripts (transcriptome), proteins (proteome), or metabolites (metabolome) are utilized. Information about where and when genomic DNA is transcribed into RNA, translated into protein and what metabolites are produced is crucial in formulating hypotheses about the possible function of each gene/gene product. Such hypotheses can be tested by reverse genetics, which begins with a deleterious mutation in a gene of interest and ends with characterization of the resulting phenotype. The functional genomics toolkit for *M. truncatula* is second to none among legumes.

Transcriptomic analysis

Transcriptomic tools for *Medicago* have evolved rapidly from *in silico* analysis of ESTs, through nylon-filter and glass-slide arrays of PCR-amplified cDNA probes [30,31], arrays of gene-specific oligonucleotides on glass [32–34], to an Affymterix GeneChip with over 50 000 probe-sets (each consisting of multiple gene-specific oligonucleotides) designed from cDNA and genomic DNA [2••,35,36]. The latter contains probes for most, although not all, *Medicago* genes. A platform for quantitative reverse-

transcription polymerase chain reaction (qRT-PCR) analysis of all *Medicago* transcription factor genes, which is more sensitive than hybridization-based array analyses, is also under development [37].

M. truncatula DNA microarrays/chips have been applied to a wide range of studies in legume biology, most notably beneficial nitrogen-fixing and arbuscular mycorrhizal (AM) symbioses. Thousands of genes have been discovered that are induced or repressed during the development of nodules and symbiotic nitrogen fixation [2••,31,32,38] and hundreds more have been implicated in AM symbioses [30,33,39]. Other work has investigated the transcriptional basis of seed development, differentiation, and desiccation [34,40•,41•], plant responses to aluminum toxicity [27,42], or changes in nitrogen nutrition [36]. Although transcriptome analysis often generates more questions than answers, these studies provide an unprecedented view into the genes and networks involved in plant processes affecting performance and yield.

The cross-species utility of *M. truncatula* microarrays and DNA chips is well established through transcriptomic experiments where alfalfa mRNA has been assayed using *M. truncatula* platforms. Nearly identical percentages of probe sets on the Affymetrix *Medicago* GeneChip detected transcripts from both *M. truncatula* and alfalfa [35]. Moreover, expression patterns between the two species were nearly identical. Long-oligo arrays of *M. truncatula* have also been used effectively with alfalfa. For example, *M. truncatula* arrays were successful in identifying transcripts upregulated in alfalfa trichomes as a basis for exploring insect defense secretions [43].

Next-generation sequencing offers a tantalising alternative to DNA arrays for whole-genome expression analysis, and one that is not limited by prior knowledge of transcribed sequences or predicted genes. Illumina and SOLiD sequencing platforms generate millions of ESTs per run, thereby capturing the majority of expressed genes with sufficient sampling depth for statistical analysis of expression. Recently, we generated approximately 75 million ESTs from RNA from *M. truncatula* nodules, seed pods, leaves, vegetative buds, flowers, or roots (between 10 and 15 million per organ), using an Illumina/Solexa instrument (G. May, C. Town, M. Udvardi, unpublished). The resulting quantitative expression data complement and extend Affymetrix GeneChip data [2••]. More significantly, the new EST data help to define the complete transcriptome of *M. truncatula* and will be valuable in ongoing efforts to identify and annotate genes, including fine-structure like transcription start and stop sites, introns, and alternative-splicing. Next-gen sequencing may also become an attractive option for transcriptomics of nonmodel species where DNA arrays are unavailable, especially if sequence lengths can be increased to facilitate alignment and contig assembly. It remains to be seen, however, whether deep EST sequencing using next-gen technologies will displace DNA microarrays as the method of choice for quantitative transcriptomics, given the sequencing and computing costs inherent in the former.

Proteomic analysis

Proteomics complements transcriptomics by measuring the amount of specific proteins, providing insight into whether or not changes in transcript levels lead to changes in the level of the corresponding proteins. Mass spectrometry (MS) is the main analytical tool of proteomics, generally used in conjunction with liquid chromatography (LC) or gel electrophoresis. MS identifies and quantifies proteins based on their partial sequence or the pattern and mass of peptide fragments generated by defined proteolysis. In both cases, complete genome sequence is an important prerequisite to successful protein identification. A number of proteomic studies have been performed on *M. truncatula*, taking advantage of *M. truncatula* EST and genomic sequence data together with protein sequence information from other plant species [44]. Completion of the *M. truncatula* genome sequence will aid protein identification, not only in this species, but also in related legumes like alfalfa. Proteomics has been applied to *Medicago* in studies of roots [45,46], seed development and germination [41•,47], plant–pathogen interactions [48], embryogenesis [49], cell walls [50], nodule response to drought [51], plasma membrane lipid rafts [52], symbiosis [53], and responses to hormones [53].

Metabolomic analysis

Metabolomics often utilizes gas chromatography coupled to mass spectrometry (GC–MS) or LC–MS to separate, measure, and identify cellular metabolites. Although several hundred metabolites are typically measured in metabolomics experiments, this is just a fraction of the total number of metabolites produced in any one species. Fortunately, many of the intermediates of primary as well as major secondary metabolism are among the compounds typically identified and measured, so metabolomics has become a useful tool for molecular phenotyping as well as monitoring the status of primary and secondary metabolism [6••]. Unlike transcriptomics and proteomics, metabolomics is not dependent upon genome sequence information, although knowing the protein-coding potential of the genome helps to identify enzymes that may be responsible for metabolism of specific compounds.

Metabolomics has been utilized in *Medicago* studies of saponins [54], flavonoids and isoflavonoids [6••], and on the effects of biotic and abiotic elicitors [6••]. Metabolomic tools have also been used in alfalfa to study the effect of transgenic manipulation of lignin biosynthesis on phenolic compounds [55] and the effect of overexpression of the TF gene, *WXP1*, regarding metabolism related to

wax accumulation [56]. Thus, metabolomics has emerged as a valuable source of phenotypic information about molecular changes wrought by genetic manipulation of crop legumes. It will be interesting to see to whether these tools are adopted by legume breeders to select superior genotypes from among natural variants in non-transgenic breeding programs.

Reverse genetics

Resources for reverse genetics are crucial to unlocking genomic secrets by linking targeted genes to their mutant phenotypes. Reverse genetics relies on large mutant populations that saturate the genome with deleterious mutations. Useful mutagens include ethyl methane sulfonate (EMS), which primarily generates point mutations, FNB, which generates DNA deletions and point mutations, and transposon or T-DNA insertions, which disrupt the normal sequence of genes. All of these mutagens potentially cause complete loss of gene/protein function, and EMS, FNB, and transposons have been utilized to generate separate mutant populations in *M. truncatula* [3••,8,57].

M. truncatula EMS mutant populations have been used for both forward and reverse genetics. Forward genetic screens have focused mostly on genes required for symbiotic nitrogen fixation and AM symbiosis (reviewed in [7]). Reverse genetic screening of *M. truncatula* EMS mutants is made possible by a procedure called TILLING, which has been used recently to assign function to the *MtPT4* gene [58•]. *MtPT4* encodes a phosphate transporter found to be essential, not only for phosphate acquisition from AM fungi, but also for stability/longevity of the symbiosis [58•].

FNB populations of *M. truncatula* have been generated and used to identify mutants defective in symbiosis and plant development [59]. PCR based and other approaches are currently being developed to enable reverse-genetic screening of these populations (G. Oldroyd and R. Chen, unpublished).

Probably the most facile of resources for reverse genetics are T-DNA or transposon-insertion mutant populations, which not only lead to complete inactivation of disrupted genes, but also provide a convenient 'tag' — the inserted foreign DNA, with which to locate the affected gene. PCR-based approaches are effective in screening such populations for insertions in genes of interest. However, the utility of such populations is increased greatly by systematic sequencing of flanking insertion sites. The resulting flanking sequence tags (FSTs) in the form of a searchable database facilitate identification of insertions/mutations in genes of interest (e.g. *Arabidopsis* T-DNA Express; http://signal.salk.edu/cgi-bin/tdnaexpress). The tobacco retrotransposon, *Tnt1* has been introduced into *M. truncatula* and mobilized during tissue culture to generate >9000 lines to date, each containing approximately 25 insertions [3••]. Nearly 9000 FSTs have been generated from 810 *Tnt1* mutant lines and deposited in a searchable database (http://bioinfo4.noble.org/mutant/). Work is underway to expand this to 100 000 FSTs. The *M. truncatula Tnt1* mutant population has been used in forward genetic screens to identify developmental and symbiotic mutants [3••,60•]. FST sequencing and PCR screening of the *Tnt1* population has identified dozens of insertions into known symbiosis genes (Pislariu *et al.*, unpublished).

Post-transcriptional gene silencing (PTGS) is an alternative to DNA mutation for reverse-genetics. RNAi revealed a role for *MtHAP2-1* in transcriptional regulation of nodule development [61•] and for the TF gene *MtSERF1* in somatic embryogenesis [62]. Antisense repression of the sucrose synthase gene *MtSucS1* compromised nitrogen fixation in nodules [38]. Although PTGS can side-step lethal phenotypes associated with complete loss-of-function alleles of some genes, it is difficult to scale up for whole-genome analysis.

Plant biotechnology

There are also a growing number of examples using *M. truncatula* genes in transgenic alfalfa to improve specific traits for agriculture. One example is the *RCT1* resistance to anthracnose mentioned earlier. Antisense-constructs or RNAi-constructs to downregulate lignin biosynthesis genes have been introduced into alfalfa, reducing lignin content and increasing biomass digestibility and weight-gain in cows [63]. Lignin modification/reduction also improves sugar yields from cellulose deconstruction, making alfalfa more attractive to the biofuel industry [64••]. Constitutive expression of isoflavone synthase, *MtIFS1*, from *M. truncatula* resulted in increased levels of genestein and other isoflavones in transgenic alfalfa, although concentrations varied substantially in different organs, highlighting the challenges with metabolic engineering of natural products [65]. In a final example, *M. truncatula* genes were used to probe transcriptional changes of cold-responsive genes to low temperature in a comparative study of *M. truncatula* and *M. falcata*, indicating that differences in the copy number and promoters of cold-acclimation-specific genes contribute to cold tolerance in *M. falcata* [66•].

Conclusions

With the growing array of genomic tools in *M. truncatula*, new avenues of research are opening up in legume crops. Successfully translating these tools to practical agriculture will lead to new discoveries about agricultural processes and accelerate crop improvement. To achieve these goals, we need to know more about the fine-structure of synteny between *M. truncatula* and its crop relatives as well as the usefulness of *M. truncatula* microarrays with more distant legume relatives. Microarrays also need to be tested as

platforms for other applications, like single feature polymorphism discovery and methylome analysis. Finally, bioinformatic resources for *M. truncatula* need to be expanded to make them better suited for translational genomics.

Acknowledgements

We thank Rick Dixon, Lloyd Sumner, and Steven Cannon for helpful discussions regarding the manuscript. We regret that it was not possible to cite the works of many additional scientists who contributed to this field as a result of space limitations. This work was supported, in part, by NSF grant DBI-0606966 to NY and NSF grant DBI-0703285 and USDA CSREES-NRI grant 2006-35300-17143 to MU.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. • Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K *et al.*: **Genome structure of the legume, *Lotus japonicus***. *DNA Res* 2008, **15**:227-239.
The first large-scale sequence of a legume genome sequence, focusing on the *Lotus* gene space of approximately 300 Mbp. The sequence provides evidence for expansion of several genes families as well as traces of an ancient whole-genome duplication.

2. •• Benedito VA, Torres-Jerez I, Murray JD, Andriankaja A, Allen S, Kakar K, Wandrey M, Verdier J, Zuber H, Ott T *et al.*: **A gene expression atlas of the model legume *Medicago truncatula***. *Plant J* 2008, **55**:504-513.
The first global gene expression atlas for a legume based on extensive DNA chip hybridization. The results demonstrate massive differences in gene expression between organs and numerous cases of legume-specific genes that are preferentially expressed in nodules.

3. •• Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M *et al.*: **Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume *Medicago truncatula***. *Plant J* 2008, **54**:335-347.
A report about the successful integration of the *Tnt1* transposon system into *M. truncatula* as a basis for reverse genetic analysis. Altogether, the authors generated 7600 insertions lines with 190 000 total insertion events. Nearly 2500 flanking segments were sequenced, demonstrating a preference for gene-containing regions.

4. Lavin M, Herendeen PS, Wojciechowski MF: **Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary**. *Syst Biol* 2005, **54**:575-594.

5. • Repetto O, Rogniaux H, Firnhaber C, Zuber H, Küster H, Larré C, Thompson R, Gallardo K: **Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling**. *Plant J* 2008, **56**:398-410.
A large-scale proteomic analysis of *M. truncatula* seeds at 12 days after pollination, the time of switch to seed filling. Proteins associated with ribosomal and trafficking functions were especially notable at this stage, along with transcriptional regulation and RNA transporting and processing.

6. •• Farag MA, Huhman DV, Dixon RA, Sumner LW: **Metabolomics reveals novel pathways and differential mechanistic and elicitor-specific responses in phenylpropanoid and isoflavonoid biosynthesis in *Medicago truncatula* cell cultures**. *Plant Physiol* 2008, **146**:387-402.
The metabolome of *M. truncatula* cell suspension cultures was analyzed through LC/MS to reveal the array of secondary products produced in response to the elicitors, yeast elicitor and methyl jasmonate. The results indicate significant metabolic flexibility in the isoflavonoid pathway as well as multiple potential pathways leading to medicarpin.

7. Oldroyd GE, Downie JA: **Coordinating nodule morphogenesis with rhizobial infection in legumes**. *Annu Rev Plant Biol* 2008, **59**:519-546.

8. Starker CG, Parra-Colmenares AL, Smith L, Mitra RM, Long SR: **Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression**. *Plant Phys* 2006, **140**:671-680.

9. Limpens E, Ramos J, Franken C, Raz V, Compaan B, Franssen H, Bisseling T, Geurts R: **RNA interference in *Agrobacterium rhizogenes*-transformed roots of *Arabidopsis* and *Medicago truncatula***. *J Exp Bot* 2004, **55**:983-992.

10. Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P, Penmetsa RV, Seres A, Kulikova O *et al.*: **A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa***. *Genetics* 2004, **166**:1463-1502.

11. Zhang Y, Sledge MK, Bouton JH: **Genome mapping of white clover (*Trifolium repens L.*) and comparative analysis within the Trifolieae using cross-species SSR markers**. *Theor Appl Genet* 2007, **114**:1367-1378.

12. Kaló P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis TH, Kiss GB: **Comparative mapping between *Medicago sativa* and *Pisum sativum***. *Mol Genet Genomics* 2004, **272**:235-246.

13. Phan HT, Ellwood SR, Hane JK, Ford R, Materne M, Oliver RP: **Extensive macrosynteny between *Medicago truncatula* and *Lens culinaris* ssp. *culinaris***. *Theor Appl Genet* 2007, **114**:549-558.

14. •• Cannon SB, Sterck L, Rombauts S, Sato S, Cheung F, Gouzy JP, Wang X, Mudge J, Vasdewani J, Schiex T *et al.*: **Legume evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes**. *Proc Natl Acad Sci U S A* 2006, **103**:14959-14964.
A whole-genome synteny analysis of *M. truncatula* and *L. japonicus* based on extensive sequence analysis. The results indicate that the two genomes share a minimum of 10 large-scale synteny blocks as well as an absence of any recent large-scale duplication in either genome.

15. Ameline-Torregrosa C, Wang BB, Denny RL, O'Bleness M, Despande S, Zhu H, Roe B, Young ND, Cannon SB: **Identification and characterization of NBS-LRR encoded genes in the model plant *Medicago truncatula***. *Plant Phys* 2008, **146**:5-21.

16. Graham MA, Silverstein KA, Cannon SB, VandenBosch KA: **Computational identification and characterization of novel genes from legumes**. *Plant Phys* 2004, **135**:1179-1197.

17. • Alunni B, Kevei Z, Redondo-Nieto M, Kondorosi A, Mergaert P, Kondorosi E: **Genomic organization and evolutionary insights on *GRP* and *NCR* genes, two large nodule-specific gene families in *Medicago truncatula***. *Mol Plant Microbe Interact* 2007, **20**:1138-1148.
NCRs and GRPs are two large gene families that code for short secreted peptides with possible antimicrobial activity. The genes are specific to *M. truncatula* and closely related legume taxa. RT-PCR demonstrates their transcripts are found only in nodules.

18. Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis THN, Doyle J, Kiss GB *et al.*: **Estimating genome conservation between crop and model legume species**. *Proc Natl Acad Sci U S A* 2004, **101**:15289-15294.

19. Sledge MK, Ray IM, Jiang G: **An expressed sequence tag SSR map of tetraploid alfalfa (*Medicago sativa* L.)**. *Theor Appl Genet* 2005, **111**:980-992.

20. Gutierrez MV, Vaz Patto MC, Huguet T, Cubero JI, Moreno MT, Torres AM: **Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops**. *Theor Appl Genet* 2005, **110**:1210-1217.

21. Ameline-Torregrosa C, Cazaux M, Danesh D, Chardon F, Cannon SB, Esquerré-Tugayé MT, Dumas B, Young ND, Samac DA, Huguet T *et al.*: **Genetic dissection of resistance to anthracnose and powdery mildew in *Medicago truncatula***. *Mol Plant-Microbe Interact* 2008, **21**:61-69.

22. •• Yang S, Gao M, Xu C, Gao J, Deshpande S, Lin S, Roe BA, Zhu H: **Alfalfa benefits from *Medicago truncatula*: the *RCT1* gene from *M. truncatula* confers broad-spectrum resistance to anthracnose in alfalfa**. *Proc Natl Acad Sci U S A* 2008, **105**:12164-12169.
A TIR-NBS-LRR resistance gene cloned in *M. truncatula* and transferred to alfalfa provides effective resistance to an agriculturally important disease pathogen.

23. Kamphuis LG, Lichtenzveig J, Oliver RP, Ellwood SR: **Two alternative recessive quantitative trait loci influence resistance to spring black stem and leaf spot in *Medicago truncatula***. *BMC Plant Biol* 2008, **26(8)**:30.

24. Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R: ***LysM* domain receptor kinases regulating rhizobial Nod factor-induced infection**. *Science* 2003, **302**:630-633.

25. Gualtieri G, Kulikova O, Limpens E, Kim DJ, Cook DR, Bisselin T, Geurts R: **Microsynteny between pea and *Medicago truncatula* in the *SYM2* region**. *Plant Mol Biol* 2002, **50**:225-235.

26. Endre G, Kereszt A, Kevei Z, Mihacea S, Kaló P, Kiss GB: **A receptor kinase gene regulating symbiotic nodule development**. *Nature* 2002, **417**:962-966.

27. Narasimhamoorthy B, Bouton JH, Olsen KM, Sledge MK: **Quantitative trait loci and candidate gene mapping of aluminum tolerance in diploid alfalfa**. *Theor Appl Genet* 2007, **114**:901-913.

28. Burstin J, Marget P, Huart M, Moessner A, Mangin B, Duchene C, Desprez B, Munier-Jolain N, Duc G: **Developmental genes have pleiotropic effects on plant morphology and source capacity, eventually impacting on seed protein content and productivity in pea**. *Plant Physiol* 2007, **144**:768-781.

29. Kim H, Hurwitz B, Yu Y, Collura K, Gill N, SanMiguel P, Mullikin JC, Maher C, Nelson W, Wissotski M *et al.*: **Construction, alignment and analysis of twelve framework physical maps that represent the ten genome types of the genus *Oryza***. *Genome Biol* 2008, **9**:R45.

30. Liu JY, Blaylock LA, Endre G, Cho J, Town CD, Vanden Bosch KA, Harrison MJ: **Transcript profiling coupled with spatial expression analyses reveals genes involved in distinct developmental stages of an arbuscular mycorrhizal symbiosis**. *Plant Cell* 2003, **15**:2106-2123.

31. Kuester H, Hohnjec N, Krajinski F, El Yahyaoui F, Manthey K, Gouzy J, Dondrup M, Meyer F, Kalinowski J, Brechenmacher L *et al.*: **Construction and validation of cDNA-based Mt6k-RIT macroarray and microarray to explore root endosymbioses in the model legume *Medicago truncatula***. *J Biotechnol* 2004, **108**:95-113.

32. Jones KM, Sharopova N, Lohar DP, Zhang JQ, VandenBosch KA, Walker GC: **Differential response of the plant *Medicago truncatula* to its symbiont *Sinorhizobium meliloti* or an exopolysaccharide-deficient mutant**. *Proc Natl Acad Sci U S A* 2008, **105**:704-709.

33. Hohnjec N, Vieweg ME, Puhler A, Becker A, Kuster H: **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 Phys* 2005, **137**:1283-1301.

34. Buitink J, Leger JJ, Guisle I, Vu BL, Wuillème S, Lamirault G, Le Bars A, Le Meur N, Becker A, Küster H *et al.*: **Transcriptome profiling uncovers metabolic and regulatory processes occurring during the transition from desiccation-sensitive to desiccation-tolerant stages in *Medicago truncatula* seeds**. *Plant J* 2006, **47**:735-750.

35. Tesfaye M, Silverstein KAT, Bucciarelli B, Samac DA, Vance CP: **The Affymetrix *Medicago* GeneChip® array is applicable for transcript analysis of alfalfa (*Medicago sativa*)**. *Funct Plant Biol* 2006, **33**:783-788.

36. Ruffel S, Freixes S, Balzergue S, Tillard P, Jeudy C, Martin-Magniette ML, van der Merwe MJ, Kakar K, Gouzy J, Fernie AR *et al.*: **Systemic signaling of the plant nitrogen status triggers specific transcriptome responses depending on the nitrogen source in *Medicago truncatula***. *Plant Phys* 2008, **146**:2020-2035.

37. Kakar K, Wandrey M, Czechowski T, Gaertner T, Scheible W-R, Stitt M, Torres-Jerez I, Xiao Y, Redman J, Wu H *et al.*: **A community resource for high-throughput quantitative RT-PCR analysis of transcription factor gene expression in *Medicago truncatula***. *Plant Methods* 2008, **4**:18.

38. Baier MC, Barsch A, Kuster H, Hohnjec N: **Antisense repression of the *Medicago truncatula* nodule-enhanced sucrose synthase leads to a handicapped nitrogen fixation mirrored by specific alterations in the symbiotic transcriptome and metabolome**. *Plant Phys* 2007, **145**:1600-1618.

39. Manthey K, Krajinski F, Hohnjec N, Firnhaber C, Puhler A, Perlick AM, Küster H: **Transcriptome profiling in root nodules and arbuscular mycorrhiza identifies a collection of novel genes induced during *Medicago truncatula* root endosymbioses**. *Mol Plant-Microbe Interact* 2004, **17**:1063-1077.

40. • Verdier J, Kakar K, Gallardo K, Le Signor C, Aubert G, Schlereth A, Town CD, Udvardi MK, Thompson RD: **Gene expression profiling of *Medicago truncatula* transcription factors identifies putative regulators of grain legume seed filling**. *Plant Mol Biol* 2008, **67**:567-580.

The first expression profiling of *M. truncatula* transcription factors expressed during seed development that employed a dedicated qRT-PCR platform for TF transcript quantification. The study goes on to identify 17 TFs that are preferentially expressed in specific seed tissues and their possible targets.

41. • Gallardo K, Firnhaber C, Zuber H, Hericher D, Belghazi M, Henry C, Kuster H, Thompson R: **A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds**. *Mol Cell Proteom* 2007, **6**:2165-2179.

A nice example of integrative functional genomics in which transcriptome and proteome data were analyzed together to reveal transcriptional and post-transcriptional regulation of gene expression during seed development.

42. Chandran D, Sharopova N, Ivashuta S, Gantt JS, Vandenbosch KA, Samac DA: **Transcriptome profiling identified novel genes associated with aluminum toxicity, resistance and tolerance in *Medicago truncatula***. *Planta* 2008, **228**:151-166.

43. Aziz N, Paiva NL, May GD, Dixon RA: **Transcriptome analysis of alfalfa glandular trichomes**. *Planta* 2005, **221**:28-38.

44. Lei Z, Elmer AM, Watson BS, Dixon RA, Mendes PJ, Sumner LW: **A two-dimensional electrophoresis proteomic reference map and systematic identification of 1367 proteins from a cell suspension culture of the model legume *Medicago truncatula***. *Mol Cell Proteomics* 2005, **4**:1812-1825.

45. Mathesius U, Keijzers G, Natera SHA, Weinman JJ, Djordjevic MA, Rolfe BG: **Establishment of a root proteome reference map for the model legume *Medicago truncatula* using the expressed sequence tag database for peptide mass fingerprinting**. *Proteomics* 2001, **1**:1424-1440.

46. Valot B, Gianinazzi S, Eliane DG: **Sub-cellular proteomic analysis of a *Medicago truncatula* root microsomal fraction**. *Phytochemistry* 2004, **65**:1721-1732.

47. Boudet J, Buitink J, Hoekstra FA, Rogniaux H, Larre C, Satour P, Leprince O: **Comparative analysis of the heat stable proteome of radicles of *Medicago truncatula* seeds during germination identifies late embryogenesis abundant proteins associated with desiccation tolerance**. *Plant Phys* 2006, **140**:1418-1436.

48. Colditz F, Nyamsuren O, Niehaus K, Eubel H, Braun HP, Krajinski F: **Proteomic approach: identification of *Medicago truncatula* proteins induced in roots after infection with the pathogenic oomycete *Aphanomyces euteiches***. *Plant Mol Biol* 2004, **55**:109-120.

49. Imin N, De Jong F, Mathesius U, van Noorden G, Saeed NA, Wang XD, Rose RJ, Rolfe BG: **Proteome reference maps of *Medicago truncatula* embryogenic cell cultures generated from single protoplasts**. *Proteomics* 2004, **4**:1883-1896.

50. Soares NC, Francisco R, Ricardo CP, Jackson PA: **Proteomics of ionically bound and soluble extracellular proteins in *Medicago truncatula* leaves**. *Proteomics* 2007, **7**:2070-2082.

51. Larrainzar E, Wienkoop S, Weckwerth W, Ladrera R, Arrese-Igor C, Gonzalez EM: ***Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress**. *Plant Phys* 2007, **144**:1495-1507.

52. Lefebvre B, Furt F, Hartmann MA, Michaelson LV, Carde JP, Sargueil-Boiron F, Rossignol M, Napier JA, Cullimore J, Bessoule JJ *et al.*: **Characterization of lipid rafts from *Medicago truncatula* root plasma membranes: a proteomic study**

reveals the presence of a raft-associated redox system. *Plant Phys* 2007, **144**:402-418.

53. van Noorden GE, Kerim T, Goffard N, Wiblin R, Pellerone FI, Rolfe BG, Mathesius U: **Overlap of proteome changes in *Medicago truncatula* in response to auxin and *Sinorhizobium meliloti***. *Plant Phys* 2007, **144**:1115-1131.

54. Huhman DV, Berhow M, Sumner LW: **Quantification of saponins in aerial and subterranean tissues of *Medicago truncatula***. *J Agric Food Chem* 2005, **53**:1914-1920.

55. Chen F, Reddy MSS, Temple S, Jackson L, Shadle G, Dixon RA: **Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (*Medicago sativa* L.)**. *Plant J* 2006, **48**:113-124.

56. Zhang J-Y, Broeckling CD, Blancaflor EB, Sledge MK, Sumner LW, Wang Z-Y: **Overexpression of *WXP1*, a putative *Medicago truncatula AP2* domain-containing transcription factor gene, increases cuticular wax accumulation and enhances drought tolerance in transgenic alfalfa (*Medicago sativa*)**. *Plant J* 2005, **42**:689-707.

57. Penmetsa RV, Cook DR: **Production and characterization of diverse developmental mutants of *Medicago truncatula***. *Plant Phys* 2000, **123**:1387-1397.

58. • Javot H, Penmetsa RV, Terzaghi N, Cook DR, Harrison MJ: **A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis**. *Proc Natl Acad Sci U S A* 2007, **104**:1720-1725.

An example of the use of TILLING to assign gene function in *M. truncatula*. In this case, the plant phosphate transporter *MtPT4* was proven to play an important role not only in P transfer from the fungus to the plant, but also in the integrity of the symbiosis as a whole.

59. Wang HL, Li GM, Chen RJ: **Fast neutron bombardment (FNB) mutagenesis for forward and reverse genetic studies in plants**. In *Floriculture, Ornamental and Plant Biotechnology: Advances and Special Issues*, vol 1. Edited by Teixeira de Silva J. Global Science Books; 2006:629-639.

60. • Wang HL, Chen JH, Wen JQ, Tadege M, Li GM, Liu Y, Mysore KS, Ratet P, Chen RJ: **Control of compound leaf development by *FLORICAULA/LEAFY* ortholog *SINGLE LEAFLET1* in *Medicago truncatula***. *Plant Phys* 2008, **146**:1759-1772.

A good example of the ease of mutant gene isolation following *Tnt1* insertion mutagenesis.

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

RNA interference indicated a key role of MtHAP2-1 in nodule development. Interestingly, this TF was found to be subject to negative regulation by microRNA169, one of the first demonstrations of this type of control in legumes.

62. Mantiri FR, Kurdyukov S, Lohar DP, Sharopova N, Saeed NA, Wang XD, VandenBosch KA, Rose RJ: **The transcription factor *MtSERF1* of the ERF subfamily identified by transcriptional profiling is required for somatic embryogenesis induced by auxin plus cytokinin in *Medicago truncatula***. *Plant Phys* 2008, **146**:1622-1636.

63. Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA: **Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.)**. *Proc Natl Acad Sci U S A* 2005, **102**:16573-16578.

64. •• Chen F, Dixon RA: **Lignin modification improves fermentable sugar yields for biofuel production**. *Nat Biotechnol* 2007, **25**:759-761.

Serendipity and the power of legume translational genomics: following gene discovery in *M. truncatula* and transgenic interference with lignin biosynthesis in alfalfa, it was found that reduction in lignin increased recovery of sugars from cell walls. This has important implications for the fledgling biofuel industry.

65. Deavours BE, Dixon RA: **Metabolic engineering of isoflavonoid biosynthesis in alfalfa**. *Plant Phys* 2005, **138**:2245-2259.

66. • Pennycooke JC, Cheng H, Stockinger EJ: **Comparative genomic sequence and expression analyses of *Medicago truncatula* and alfalfa subspecies *falcata* cold-acclimation-specific genes**. *Plant Phys* 2008, **146**:1242-1254.

A nice example of the power of comparative genomics with the prospect of translating the results to plant breeding.