

Genome Wide Identification of Nodule-Specific Transcripts
in *Medicago truncatula*

M. Fedorova¹, J. van de Mortel¹, P. Matsumoto¹, J. Cho², C. D. Town²,
K. A. VandenBosch³, J. S. Gantt³, C. P. Vance^{1,4}

¹Department of Agronomy and Plant Genetics, Univ. Minnesota, St. Paul, MN

²The Institute for Genomic Research, Rockville, MD

³Department of Plant Biology, Univ. Minnesota, St. Paul, MN

⁴USDA-ARS-Plant Science Research, St. Paul, MN

The *Medicago truncatula* EST database (Gene Index) contains more than 140,000 sequences from 30 cDNA libraries. This resource offers the possibility of identifying previously uncharacterized genes and assessing the frequency and tissue specificity of their expression *in silico*. Because *Medicago truncatula* forms symbiotic root nodules, unlike *Arabidopsis thaliana*, this is a particularly important approach in investigating genes specific to nodule development and function in legumes. Our analyses have revealed 340 putative gene products, or tentative consensus sequences (TCs), expressed solely in root nodules. These TCs were represented by 2 to 379 ESTs. Of these TCs, 3% appear to encode novel proteins, 57% encode proteins with a weak similarity to the GenBank accessions and 40% encode proteins with strong similarity to the known proteins. Nodule-specific TCs were grouped into nine categories based on the predicted function of their protein product. Besides previously characterized nodulins, other examples of highly abundant nodule-specific transcripts include plantacyanin, agglutinin, embryo-specific protein, and purine permease. Six nodule-specific TCs encode calmodulin-like proteins that possess a unique cleavable transit sequence potentially targeting the protein into the peribacteroid space. *In situ* hybridization localized calmodulin-like mRNA to infected cells of the root nodule. Surprisingly, 114 nodule-specific TCs show similarity to the pea ENOD3 gene and encode small cysteine cluster proteins with a cleavable transit peptide. To determine the validity of the *in silico* analysis, expression of 91 putative nodule-specific TCs was analyzed by macroarray and RNA blot hybridizations. Nodule-enhanced expression was confirmed experimentally for the TCs comprised of five or more ESTs, while the results for those TCs containing fewer ESTs were variable. This analysis of nodule-specific TC has defined several previously unrecognized nodule-specific genes. These unique genes will provide additional insight into nodule development and function.