SHORT COMMUNICATION

Cloning, molecular characterization, and expression analysis of several red clover cDNAs

Michael L. Sullivan[1] and Sharon L. Thoma[2, 3]

[1]*US Dairy Forage Research Center, USDA-ARS, 1925 Linden Drive West, Madison, WI 53706, USA (e-mail, mlsulliv@wisc.edu); and* [2]*Natural Science Department, Edgewood College, 1000 Edgewood College Drive, Madison, WI 53711, USA.* *Received 28 February 2005, accepted 20 October 2005*

Sullivan, M. L. and Thoma, S. L. 2006. **Cloning, molecular characterization, and expression analysis of several red clover cDNAs**. Can. J. Plant Sci. **86**: 465–468. To begin gathering information regarding nucleotide sequence similarity between red clover genes and other plant species, especially the model legume *Medicago truncatula*, several random red clover cDNAs were sequenced. The analyzed cDNAs included genes encoding actin; several proteins involved in photosynthesis including PsaH, PsbR, PsbX, early light-induced protein (ELIP), ferredoxin, chlorophyll a/b binding protein; fructose-bisphosphate aldolase; chloroplastic superoxide dismutase; and GTP-binding protein typA. The gene set had a median sequence identity of 92% with their counterparts from *M. truncatula*, suggesting its available genomics tools can be applied to red clover. An expression analysis of the gene set in various red clover tissues indicates the genes show a wide range of expression patterns. Consequently, this set of cDNAs and associated data are proving useful as controls in molecular genetic experiments involving red clover.

Key words: Red clover, *Trifolium pratense*, *Medicago truncatula*, forage legume, genomics, inquiry-based learning

Sullivan, M. L. et Thoma, S. L. 2006. **Clonage, caractérisation moléculaire et analyse de l'expression de plusieurs ADNc du trèfle rouge**. Can. J. Plant Sci. **86**: 465–468. Les auteurs ont séquencé plusieurs ADNc du trèfle rouge sélectionnés au hasard afin d'en apprendre davantage sur les similitudes existant entre les gènes de cette espèce et ceux d'autres plantes, en particulier la légumineuse modèle *Medicago truncatula*, au niveau de la séquence des nucléotides. Les ADNc analysés incluaient les gènes codant l'actine, diverses protéines participant à la photosynthèse (dont PsaH, PsbR et PsbX), les protéines induites par une faible lumière (ELIP), la ferrédoxine, la protéine a/b de liaison de la chlorophylle, mais aussi la fructose-biphosphate aldolase, la superoxyde dismutase des chloroplastes et la protéine typA de liaison de la GTP. Le jeu de gènes est identique à 92 % (valeur médiane) à la séquence correspondante de *M. trunculata*, signe qu'on peut se servir des outils de génomique existants avec le trèfle rouge. L'analyse de l'expression des gènes dans les divers tissus du trèfle révèlent que les gènes s'expriment de façon très variable. Par conséquent, ce jeu d'ADNc et les données qui s'y rapportent présentent de l'utilité comme témoin dans les expériences de génétique moléculaire sur le trèfle rouge.

Mots clés: Trèfle rouge, *Trifolium pratense, Medicago trunculata*, légumineuse fourragère, génomique, méthode expérimentale

Red clover (*Trifolium pratense* L.), an important and versatile forage legume, is probably the second most important forage legume after alfalfa in Canada and the United States. In some regions, such as Scandinavia, red clover is the most widely used forage legume (Frame et al. 1998). There has also been recent interest in the use of red clover-derived supplements for treatment of various medical conditions, including post-menopausal symptoms, prostate cancer, and vascular disease, because of red clover's high content of phytoestrogen and isoflavanoid compounds (Saloniemi et al. 1995). Red clover has the potential to serve not only as a source of such useful secondary metabolites, but may also be a useful model system for studying secondary metabolic pathways and a source of the genes involved.

Molecular genetic and genomics approaches have great potential to complement traditional breeding approaches in red clover. For example, expression profiling using DNA microarrays could help identify specific genes involved in traits like persistence or disease resistance. These could in turn result in the development of useful genetic and biochemical markers for such desirable traits allowing even greater improvements to be made by traditional breeding. Unfortunately, sequencing the red clover genome will not likely happen in the near future since the large number of important legume species prevents their simultaneous in-depth characterization (Choi et al. 2004). If red clover genes are sufficiently similar to those of the model legume *M. truncatula*, however, genomics tools developed for *M. truncatula* (e.g., DNA microarrays) could be applied to red clover. Although Choi et al. (2004) have recently shown

[3]*Current address: Department of Zoology, University of Wisconsin, 250 N. Mills St., Madison, WI 53706, USA.*

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

Table 1. Identities and properties of red clover cDNAs

Clone	Genbank accession	Encoded protein	Peptide length[z]	CTP[y] length[z]	Sequence identity[x] (%)	
					A. thaliana	*M. truncatula*
MM2	AY372368	Actin	377	–	84	95
RL1	AY340641	10.2 kDa Photosystem I protein (PsaH)	146	50	73	92
MM9	AY348557	10 kDa Photosystem II protein (PsbR)	136	36	72	92
CM1	AY340642	UV-repressible protein (PsbX)	118	50	61	88
PN2	AY340640	Early Light-induced protein (ELIP)	198	46	57	91
PN1	AY340639	Ferredoxin I (Fed)	152	55	64	87
JS1	AY430082	Chlorophyll a/b binding protein (Cab)	266	43	77	93
ME2	AY430081	Fructose-bisphosphate aldolase	397	46	75	92
RT2	AY434497	Chloroplastic superoxide dismutase (SOD)	202	47	73	91
RT1	AY445630	GTP-binding protein typA (GTP-BP)	676	63	74	93

[z]Length in amino acid residues.
[y]Chloroplast transit peptide.
[x]Nucleotide sequence identity of open reading frame.

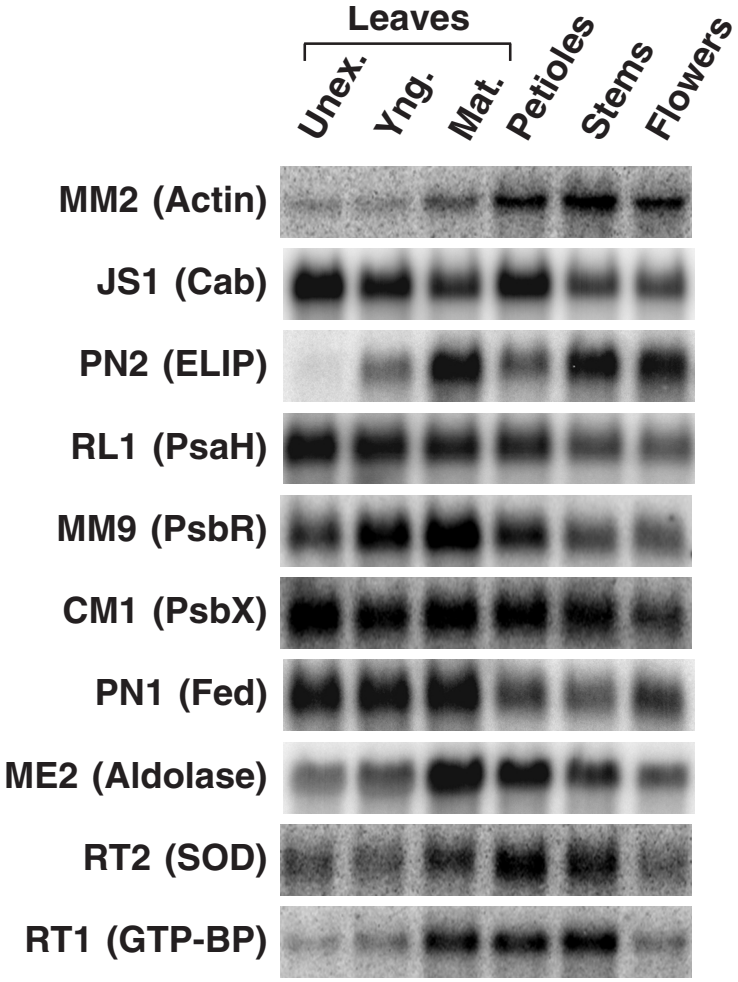

Fig. 1. Expression patterns of several red clover genes. Red clover RNA from unexpanded (Unex.), young (Yng.), and mature (Mat.) leaves; mature petioles; stems; and flowers was fractionated on denaturing agarose gels and transferred to nylon membranes. The resulting RNA blots were used in hybridization experiments with ^{32}P-labeled DNA probes corresponding to the indicated red clover clones and imaged using a phosphoimager. Grayscale ranges for blots were selected to optimize signal-to-noise ratio for each blot. The gene product name of each clone is indicated in parenthesis.

broad conservation of the genome structure of various legumes, their study did not examine conservation at the nucleotide sequence level and did not include red clover. To begin gathering information regarding red clover genes and the extent of sequence similarity they share with those of other plant species (especially *M. trucatula*), we picked and

sequenced random clones from a red clover leaf cDNA library. When we initiated the study described in this communication, only a few red clover gene sequences were available in Genbank, including three rRNA-related sequences (Bena et al. 1998) and three cDNA sequences corresponding to mRNAs [two for isoflavone synthase genes (Jung et al. 2000) and a short sequence corresponding to a lectin gene fragment (Genbank Accession Z92673)]. Our collection of gene sequences were compared to those in Genbank, TIGR, and TAIR databases. We also carried out expression analyses via RNA blotting, since such baseline expression data of a small collection of red clover genes could prove useful to other researchers conducting molecular genetic experiments in this species.

A red clover clone (Sullivan et al. 2004) from a population of cultivar WI-2 (Lot C136) (Smith and Maxwell 1980) was the source of plant tissue for this study. An amplified cDNA library prepared from unexpanded and young leaves of red clover (Sullivan et al. 2004) in the λ_{ZAP} vector was plated on XL1-Blue MRF' in the presence of IPTG (isopropyl-β-D-thiogalactoside) and X-Gal (5-bromo-4-chloro-3-indollyl β-D-galactopyranoside), individual clear plaques were picked, and selected λ clones rescued to pBluescript phagemids according to the manufacturer's protocols (Stratagene, LaJolla, CA). Phagemid DNA was prepared using miniprep kits (Promega, Madison, WI or Qiagen, Valencia, CA) and cDNA inserts were sequenced by cycle sequencing using BigDye terminator chemistry and automated DNA sequencers (Applied Biosystems, Foster City, CA). Initial sequencing of 5′ and 3′ ends of each cloned cDNA was used to identify phagemids with single, full-length cDNA inserts. Phagemids with multiple inserts and partial cDNA clones were not further characterized. For RNA blotting, RNA was isolated from unexpanded, fully expanded young (leaflets ≤ 1.5 cm), and mature leaves (leaflets ≥ 2.0 cm), mature petioles, stems, and flowers of red clover by the method of Chang et al. (1993). RNA blotting and hybridization probe preparation was performed as described elsewhere (Sullivan et al. 2004). Hybridization was carried out at 60°C in buffer containing 50% Formamide and 0.8 M NaCl and post-hybridization washes were carried out in 0.2× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.1% SDS at 60°C. Radioactive signals were detected with a Molecular Dynamics PhosphoImager Model 425 or Typhoon imaging system.

Ten full-length clones were sequenced in entirety on both strands and gene identifications were made by BLAST searches of Genbank (http://www.ncbi.nlm.nih.gov/), TAIR (The Arabidopsis Information Resource, http://www.arabidopsis.org/), and TIGR MtGI (The Institute for Genomic Research *Medicago truncatula* Gene Index, http://www.tigr.org/tdb/tgi/mtgi/) databases (Table 1). Additional sequence analyses were carried out using Wisconsin Package Version 10.3 (Accelrys Inc., San Diego, CA) and the ChloroP algorithm (Emanuelsson et al. 1999); http://www.cbs.dtu.dk/services). Of the 10 genes analyzed in this study, nine encoded chloroplast proteins with predicted chloroplast transit peptides ranging in length from 36 to 63 amino acids. The cDNAs included genes encoding actin; several proteins involved in photosynthesis including PsaH, PsbR, PsbX, early light-induced protein (ELIP), ferredoxin, chlorophyll a/b binding protein; fructose-bisphosphate aldolase; chloroplastic superoxide dismutase; and GTP-binding protein typA. The open reading frames of the red clover genes showed 57–85% (median = 73%) and 87–95% (median = 92%) nucleotide sequence identity when compared with open reading frames derived from the corresponding *A. thaliana* genes or *M. truncatula* tentative consensus sequences, respectively. Although it is not surprising that genes of the legumes *M. truncatula* and red clover share a high degree of sequence similarity, this level of sequence identity is significantly higher than what we had seen with red clover polyphenol oxidase genes (Sullivan et al. 2004), which share only 70% sequence identity with their *M. truncatula* counterparts. This relatively low level of sequence identity may reflect a functional difference of the red clover PPO enzyme, whose expression pattern and activity are significantly different from those of *Medicago* species (Sullivan et al. 2004). Additionally, it could be that the genes selected in the present study share an unusually high degree of sequence identity with their *M. truncatula* counterparts due to the functional nature of the encoded products (i.e., many are involved in photosynthesis). Nonetheless, the high level of sequence identity between red clover and *M. truncatula* genes suggests that emerging *M. truncatula* genomics tools, such as gene microarrays, will be useful in the analysis of red clover gene expression.

Results of RNA blotting experiments indicate the genes show a wide range of expression patterns, both in terms of the levels of gene expression and tissues in which the genes are expressed (Fig. 1 and data not shown). Ferredoxin and several of the photosynthetic genes, including chlorophyll a/b binding protein and PsaH had especially high expression, particularly in green tissues, whereas actin and superoxide dismutase had relatively low expression levels. The high level expression of ferredoxin, chlorophyll a/b binding protein, and PsaH in unexpanded and young leaves (the source tissue for the library) was also reflected in recovery of multiple cDNA clones corresponding to these genes (data not shown). With their differing levels and patterns of expression, the isolated red clover cDNAs are proving to be useful controls in gene expression studies and other molecular genetics experiments involving red clover and *Medicago* species. For example, we have used oligonucleotide primer pairs based on the red clover actin sequence as a control in real time PCR experiments, and were able to validate the real time PCR data with the RNA blot data presented here.

We wish to thank the Edgewood College students who contributed to this study including Merici Evans Awe, Craig Machut, Meghan Magner, Patrick Newman, Jessica Sam, Ryan Turner, and Regan Larson Veith.

Bena, G., Prosperi J. M., Lejeune, B. and Oliveri, I. 1998. Evolution of annual species of the genus *Medicago*: a molecular phylogenetic approach. Mol. Phylogenet. Evol. **9**: 552–559.

Chang, S., Puryear, J. and Cairney, J. 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. **11**: 113–116.

Choi, H. K., Mun, J. H., Kim, D. J., Hongyan, Z., Baek, J. M, Mudge, J., Roe, B., Ellis, N., Doyle, J., Kiss, G. B., Young, N. D.

and Cook, D. R. 2004. Estimating genome conservation between crop and model legume species. Proc. Nat. Acad. Sci. USA **101**: 15289–15294.

Emanuelsson, O., Nielsen, H. and von Heijne, G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Prot. Sci. **8**: 978–984.

Frame, J., Charlton, J. F. L. and Laidlaw, A. S. 1998. Temperate forage legumes. CAB International, New York, NY. 327 pp.

Jung, W., Lau, S. M., Yu, O., Odell, J., Fader, G. and McGonigle, B. 2000. Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. Nature Biotechnol. **18**: 208–212.

Saloniemi, H., Wahala, K., Nykanenkurki, P., Kallela, K. and Saastamoinen, I. 1995. Phytoestrogen content and estrogenic effect of legume fodder. Proc. Soc. Exp. Biol. Med. **208**: 13–17.

Smith, R. R. and Maxwell, D. P. 1980. Registration of WI-1 and WI-2 red clover. Crop Sci. **20**: 831.

Sullivan, M. L., Hatfield, R. D., Thoma, S. L. and Samac, D. A. 2004. Cloning and characterization of red clover polyphenol oxidase cDNAs and expression of active protein in *Escherichia coli* and transgenic alfalfa. Plant Physiol. **136**: 3234–3244.