MPMI Vol. 8, No. 5, 1995, pp. 781–783. 

*Research Note*

# Expressed Sequence Tags of Randomly Selected cDNA Clones from *Eucalyptus globulus–Pisolithus tinctorius* Ectomycorrhiza

**Denis Tagu and Francis Martin**

Equipe de Microbiologie Forestière, Centre I.N.R.A. de Nancy, 54280 Champenoux, France
Received 30 December 1994. Accepted 25 May 1995.

**Random sequencing of cDNA clones from *Eucalyptus globulus–Pisolithus tinctorius* ectomycorrhizal tissues was carried out to generate expressed sequence tags (ESTs). Database comparisons revealed that 42% of the cDNAs corresponded to previously sequenced genes. These ESTs represent efficient molecular markers to analyze changes in gene expression during the formation of the ectomycorrhizal symbiosis.**

*Additional keywords:* cytoskeleton, *Eucalyptus globulus*, hydrophobin, *Pisolithus tinctorius*, PR-proteins, proteasome.

Ectomycorrhizas are characterized structurally by the presence of a pseudoparenchymatous fungal mantle ensheathing the root and a network of intercellular hyphae characterized by labyrinthine branching. During development of the symbiosis, cell differentiation, and tissue patterning give rise to a novel spatial organization, changes in cell shape, and the generation of different cell types. Morphological differentiation is accompanied by modifications of protein biosynthesis (Burgess et al. 1995; Guttenberger and Hampp 1992; Hilbert and Martin 1988; Hilbert et al. 1991; Simoneau et al. 1993; Simoneau et al. 1994), by alterations in gene expression (Nehls and Martin 1995; Tagu et al. 1993), and by the onset of a novel metabolic organization in fungal and plant cells leading to the functioning symbiotic organ (Martin and Hilbert 1991).

Several molecular approaches, including differential hybridizations and use of heterologous probes, have been developed to identify genes which are preferentially expressed in *Eucalyptus globulus-Pisolithus tinctorius* ectomycorrhiza, and thus are targets for developmental regulation (Martin and Tagu 1995). We have recently attempted to generate a transcript catalogue of abundant and moderately expressed genes in eucalypt symbiotic tissues by systematic sequencing of cDNAs to generate expressed sequence tags (ESTs).

Extensive analyses of ESTs have been carried out first in human tissues (Adams et al. 1991, 1993) and then applied to plants (Höfte et al. 1993; Keith et al. 1993; Kurata et al. 1994; Newman et al. 1994; Park et al. 1993; Sasaki et al. 1994; Shen et al. 1994; Uchimiya et al. 1992; Umeda et al. 1994). In this paper, we report the result of partial sequencing and database comparison of cDNA clones of symbiotic tissues of the *E. globulus–P. tinctorius* ectomycorrhiza. The results show that this approach provides valuable cDNA clones which could then represent molecular markers of the symbiosis development.

Germination of half-sib seeds of *E. globulus* ssp. *bicostata* (Maid et al.) and growth of seedlings were carried out as described in Hilbert et al. (1991). The basidiomycete *Pisolithus tinctorius* Coker & Couch (isolate 441) was used for aseptic establishment of ectomycorrhizas. A directionally cloned cDNA library in the phage vector λZAPII was constructed from poly(A)+ RNA obtained from differentiating (4-day-old) eucalypt ectomycorrhizas (Tagu et al. 1993). By 4 days after inoculation of eucalypt roots by *Pisolithus*, the fungal mantle was well developed and tightly appressed to root epidermal cells. Individual recombinant plaques from the cDNA library were recovered. Inserts were obtained by PCR (Saiki et al. 1988) using universal and reverse primers (Tagu et al. 1993). cDNA larger than 500 bp were purified on a Qiagen column (Dusseldorf, Germany, QIAquick-spin PCR purification kit). The selected cDNA clones were sent to Euro Séquences Gènes Services (Montigny-le-Bretonneux, France) for automated sequencing. The coding strand was polymerized from the T3 sequencing primer (Applied Biosystems PRISM Ready Reaction Dye Primer Cycle sequencing kit). Sequencing reactions were run on an ABI 373A sequencer (Applied Biosystems, Foster City, CA). Each cDNA was sequenced at least two times. Sequence data were stored, assembled and analyzed using the SeqApp application (version 1.9; Gilbert, 1992; anonymous ftp from iubio.bio.indiana.edu). Sequences were edited to remove vector and ambiguous sequences, translated in the six reading frames and compared with the sequences deposited in the NCBI nonredundant database using the BLAST network service (cutoff PAM120; high score: 60) (Altschul et al. 1994) and the Internet link of SeqApp.

Approximately 200 plaques from the cDNA library were selected randomly for PCR amplification. After elimination of inserts smaller than 500 bp, about 100 clones were partially sequenced. Six were misoriented and eliminated from the analysis.

Corresponding author: D. Tagu; E-mail: tagu@nancy.inra.fr

The GenBank accession numbers for the ESTs are L38735 through L38792.

Sequencing of the selected clones resulted in 55 informative sequence tags. To characterize the ESTs generated in this pilot experiment, comparison of all six possible translational reading frames of each EST was performed. Database comparisons of the 55 ESTs revealed that 42% cDNAs corresponded to previously sequenced genes. These cDNAs were provisionally identified based on the following criteria: (i) amino acid similarity higher than 50% over the entire sequence, (ii) an unbroken reading frame, and (iii) the presence of appropriate conserved residues. The remaining clones showed little or no similarity to genes in the databases and may represent novel genes. The number of database matches was comparable to that observed in *Caenorhabditis elegans* (Waterston et al. 1992) or *Arabidopsis thaliana* (Höfte et al. 1993), but was high compared to the 8% of identified EST from the rice or the *Brassica napus* libraries (Uchimiya et al. 1992; Umeda et al. 1994; Park et al. 1993). This discrepancy could be explained by the high cutoff score used for the database comparisons in the latter studies (Höfte et al. 1993).

As shown in Table 1, 9 sequences matched previously reported fungal genes, and 7 had similarities to plant genes. The relative abundance of fungal proteins possibly reflects the large number of fungal RNAs present in symbiotic tissues (Tagu et al. 1993). Many of the database-matched ESTs were similar to known housekeeping genes. Three cDNA clones were related to stuctural proteins. The EST32 and EST141 matched sequences encoding hydrophobins, a family of morphogenetic fungal cell wall proteins (Wessels 1993, 1994). Further studies demonstrated that these hydrophobin transcripts are very abundant in eucalypt mycorrhiza when hypae are aggregating around the root (Tagu, unpublished). Several defense- or infection-related proteins were identified as well as polypeptides involved in the assembly and turnover of proteins. Root colonization by the ectomycorrhizal hyphae could thus possibly involve degradative enzymes and defense reaction of the host-plant (Albrecht et al. 1994a, 1994b).

Further characterization of these ESTs is currently underway to (i) assign the tissue origin of the transcripts, (ii) determine whether they correspond to symbiosis-regulated mRNAs, and (iii) further characterize the genes by full-length sequencing.

Generating a transcript list through cataloguing ESTs may well provide molecular markers that are useful for deciphering the modifications of genetic programs induced by symbiosis development (Martin and Tagu 1995).

## ACKNOWLEDGMENTS

This work was supported by grants from the Eureka-Eurosilva Cooperation Programme on Tree Physiology and the I.N.R.A. (AIP development). We greatly appreciated the technical assistance of Stéphanie Blanchard. We thank B. Dell, U. Nehls, and P. Murphy for helpful discussions.

## LITERATURE CITED

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie, W. R., and Venter, J. C. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. Science 252:1651-1656.

Adams, M. D., Kerlavage, A. R., Fields, C., and Venter, J. C. 1993. 3400 new expressed sequence tags identify diversity of transcripts in human brain. Nature Genet. 4:256-267.

Albrecht, C., Asselin, A., Piché, Y., and Lapeyrie, F. 1994a. Chitinase activities are induced in *Eucalyptus globulus* roots by ectomycorrhizal or pathogenic fungi, during early colonization. Physiol. Plant. 91:104-110.

Albrecht, C., Burgess, T., Dell, B., and Lapeyrie, F. 1994b. Chitinase and peroxidase activities are induced in eucalyptus roots according to ag-

**Table 1.** The list of putatively identified ESTs from a cDNA library of *Eucalyptus globulus–Pisolithus tinctorius* ectomycorrhiza[a]

| Clone # | cDNA size (bp) | Putative identification | Species | I | S | L (aa) | Access. no. |
|---|---|---|---|---|---|---|---|
| EST 7 | 500 | Saposin | *Mus musculus* | 31% | 53% | 73 | L38775 |
| EST 31 | 850 | Elongation factor 1γ | *Xenopus laevis* | 61% | 73% | 61 | L38762 |
| EST 32 | 400 | Hydrophobin | *Schizophyllum commune* | 68% | 84% | 25 | L38763 |
| EST 35 | 1,200 | Pyridoxamine phosphate oxydase | *Saccharomyces cerevisiae* | 68% | 78% | 35 | L38764 |
| EST 44 | 1,200 | Homoserine kinase | *Saccharomyces cerevisiae* | 58% | 74% | 95 | L38767 |
| EST45 | 1,100 | Metalloprotease | *Aspergillus fumigatus* | 50 | 65% | 64 | L38736 |
| EST 46 | 650 | Ubiquitin-conjugating enzyme E2 | *Saccharomyces cerevisiae* | 42% | 55% | 80 | L38768 |
| EST 54 | 850 | Glyoxylate pathway regulator | *Yarrowia lipolytica* | 49% | 64% | 50 | L38770 |
| EST 57 | 700 | Sphingomyelinase | *Clostridium perfringens* | 61% | 78% | 71 | L38772 |
| EST 60 | 900 | Proteasome component C2 | *Homo sapiens* | 63% | 84% | 123 | L38773 |
| EST75 | 350 | Unknown ORF | *Saccharomyces cerevisiae* | 45% | 71% | 43 | L38780 |
| EST 78 | 1,200 | Aconitase | *Bos taurus* | 32% | 50% | 62 | L38782 |
| EST 82 | 900 | Asparaginyl endopeptidase | *Canavalia ensiformis* | 38% | 63% | 60 | L38785 |
| EST 84 | 1,200 | Alternative oxydase | *Hansenula anomala* | 53% | 66% | 45 | L38786 |
| EST 91 | 450 | Transposase | *Lactococcus lactis* | 56% | 70% | 41 | L38789 |
| EST 94 | 600 | Cylicin | *Bos taurus* | 20% | 44% | 90 | L38790 |
| EST 141 | 400 | Hydrophobin | *Schizophyllum commune* | 39% | 54% | 87 | L38747 |
| EST 144 | 650 | Methylcrotonyl-CoA carboxylase | *Arabidopsis thaliana* | 56% | 69% | 80 | L38748 |
| EST 149 | 400 | Proteinase inhibitor | *Glycine max* | 86% | 90% | 30 | L38751 |
| EST 155 | 850 | Enoyl-acyl carrier protein reductase | *Brassica napus* | 81% | 89% | 117 | L38755 |
| EST 158 | 650 | Ubiquitin-conjugating enzyme E2 | *Arabidopsis thaliana* | 48% | 63% | 41 | L38756 |
| EST164 | 650 | Unknown ORF | *Homo sapiens* | 60% | 69% | 43 | L38757 |
| EST 167 | 500 | PR-protein STH-21 | *Solanum tuberosum* | 66% | 80% | 21 | L38758 |
| EST 173 | 600 | Dehydroquinate dehydratase / Shikimate dehydrogenase | *Nicotiana tabacum* | 58% | 76% | 77 | L38759 |

[a] Columns refer respectively to (1) the EST number, (2) the length of the cDNA clone, (3) the identity of the protein with the highest similarity score, (4) the species, (5)% identity, (6)% similarity, (7) the overlap in amino acids and (8) the accession number identifier in dbEST (NCBI). aa: amino acids. Access. no: accession number in the database. I: identity. L: length. S: similarity.

gressiveness of Australian ectomycorrhizal strains of *Pisolithus* sp. New Phytol. 127:217-222.

Altschul, S. F., Boguski, M. S., Gish, W., and Wootton J. C. 1994. Issues in searching molecular sequence databases. Nature Genet. 6:119-129.

Burgess, T., Laurent, P., Dell, B., Malajczuk, N., and Martin, F. 1995. Effect of the fungal isolate aggressivity on the biosynthesis of symbiosis-related polypeptides in differentiating eucalypt ectomycorrhiza. Planta 195:408-417.

Guttenberger, M., and Hampp, R. 1992. Ectomycorrhizins—Symbiosis-specific or artifactual polypeptides from ectomycorrhizas? Planta 188:129-136.

Hilbert, J. L., and Martin, F. 1988. Regulation of gene expression in ectomycorrhizas. I. Protein changes and the presence of ectomycorrhiza-specific polypeptides in the *Pisolithus-Eucalyptus* symbiosis. New Phytol. 110:339-346.

Hilbert, J. L., Costa, G., and Martin, F. 1991. Ectomycorrhizin synthesis and polypeptide changes during the early stage of eucalypt mycorrhiza development. Plant Physiol. 97:977-984.

Höfte, H., Desprez, T., Amselem, J., Chiapello, H., Caboche, M., Moisan, A., Jourjon, M. F., Charpenteau, J. L., Berthomieu, P., Guerrier, D., Giraudat, J., Quigley, F., Thomas, F., Yu, D. Y., Mache, R., Raynal, M., Cooke, R., Grellet, F., Delseny, M., Parmentier, Y., De marcillac, G., Gigot, C., Fleck, J., Phillips, G., Axelos, M., Bardet, C., Tremousaygue, D., and Lescure, B. 1993. An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. Plant J. 4:1051-1061.

Keith, C. S., Hoang, D. O., Barrett, B. M., Feigelman, B., Nelson, M. C., Thai, H., and Baysdorfer, C. 1993. Partial sequence analysis of 130 randomly selected maize cDNA clones. Plant Physiol. 101:329-332.

Kurata, N., Nagamura, Y., Yamamoto, K., Harushima, Y., Sue, N., Wu, J., Antonio, B. A., Shomura, A., Shimizu, T., Lin, S-Y., Inoue, T., Fukuda, A., Shimano, T., Kuboki, Y., Toyama, T., Miyamoto, Y., Kirihara, T., Hayasaka, K., Miyao, A., Monna, L., Zhong, H. S., Tamura, Y., Wang, Z-X., Momma, T., Umehara, Y., Yano, M., Sasaki, T., and Minobe, Y. 1994. A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genet. 8:365-372.

Martin, F. M., and Hilbert, J. L. 1991. Morphological, biochemical and molecular changes during ectomycorrhiza development. Experientia 47:321-331.

Martin, F., and Tagu, D. 1995. Ectomycorrhiza development: a molecular perspective. Pages 29-58 in: Mycorrhiza: Structure, Function, Molecular Biology andBiotechnology. A. K. Varma and B. Hock, eds. Springer-Verlag, Berlin.

Nehls, U., and Martin, F. 1995. Identification of differential expressed genes during ectomycorrhiza formation. In: Biotechnology of Ectomycorrhizae: Molecular Approaches. V. Stocchi, eds. Plenum Publishing Corporation, New York. In press.

Newman,T., De Bruijn, F. J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C. 1994. Genes galore: A summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDAN clones. Plant Physiol. 106:1241-1255.

Park, Y. S., Kwak, J. M., Kwon, O. Y., Kim, Y. S., Lee, D. S., Cho, M. J., Lee, H. H., and Nam, H. G. 1993. Generation of expressed sequence tags of random root cDNA clones of *Brassica napus* by single-run partial sequencing. Plant Physiol. 103:359-370.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487-491.

Sasaki, T., Song, J., Koga-Ban, Y., Matsui, E., Fang, F., Higo, H., Nagasaki, H., Hori, M., Miya, M., Murayama-Kayano, E., Takiguchi, T., Takasuga, A., Niki, T., Ishimaru, K., Ikeda, H., Yamamoto, Y., Mukai, Y., Ohta, I., Miyadera, N., Havukkala, I., and Minobe, Y. 1994. Toward cataloguing all rice genes: Large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J. 6:615-624.

Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, B., McCreery, T., Helentjaris, T., Baysdorfer, C., Almira, E., Ferl, R. J., Habben, J. E., and Larkins, B. 1994. Partial sequencing and mapping of clones from two cDNA maize cDNA libraries. Plant Mol. Biol. 26:1085-1101.

Simoneau, P., Juge, C., Dupuis, J. Y., Viemont, J. D., Moreau, C., and Strullu, D. G. 1994. Protein biosynthesis changes during mycorrhiza formation in roots of micropropagated birch. Acta Bot. Gallica 141:429-435.

Simoneau, P., Viémont, J. D., Moreau, J. C., and Strullu, D. G. 1993. Symbiosis-related polypeptides associated with the early stages of ectomycorrhiza organogenesis in birch (*Betula pendula* Roth). New Phytol. 124:495-504.

Tagu, D., Python, M., Crétin, C., and Martin, F. 1993. Cloning symbiosis-related cDNAs from eucalypt ectomycorrhiza by PCR-assisted differential screening. New Phytol. 125:339-343.

Uchimiya, H., Kidou, S., Shimazaki, T., Aotsuka, S., Takamatsu, S., Nishi, R., Hashimoto, H.,Matsubayashi, Y., Kidou, N., Umeda, M., and Kato, A. 1992. Random sequencing of cDNA libraries reveals a variety of expressed genes in cultured cells of rice (*Oryza sativa* L.). Plant J. 2:1005-1009.

Umeda, M., Hara, C., Matsubayashi, Y., Li, H. H., Liu, Q., Tadokoro, F., Aotsuka, S. and Uchimiya, H. 1994. Expressed sequence tags from cultured cells of rice (*Oryza sativa* L.) under stressed conditions: Analysis of transcripts of genes engaged in ATP-generating pathways. Plant Mol. Biol. 25:469-478.

Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Showkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierry-Mieg, J., and Sulston, J. 1992. A survey of expressed genes in *Caenorhabditis elegans*. Nature Genet. 1:114-123.

Wessels, J. G. H. 1993. Tansley Review N°45. Wall growth, protein excretion and morphogenesis in fungi. New Phytol. 123:397-413.

Wessels, J. G. H. 1994. Developmental regulation of fungal cell wall formation. Annu. Rev. Phytopathol. 32:413-437.