Sequencing of oil palm (*Elais quineensis* jacq) root cDNA for development of molecular markers for breeding and disease resistance

[1]Tan, SH, [2]Abdullah, M., [1]Ng, W.H.

[1]Institute Bioscience,
Universiti Putra Malaysia, 43400 UPM Serdang Selangor,
Malaysia; [2]

Malaysia Palm Oil Board (MPOB),
No.6, Institute Persiaran, Bandar Baru Bangi,
43000 Kajang, Selangor, Malaysia.

Key words: Oil palm, Expresses Sequence Tags, roots

Introduction

Elaeis guineensis, originated from West and Central Africa, cultivated as an ornamental tree in early decades has become an essential commodity crops in Malaysia. For many years now, palm oil has been the leading vegetable oil traded in the world market. In botany, roots, the descending axis of a plant, normally underground, function to absorb water and dissolved minerals from the soil, to anchor the plant, and often to store food. There are two main types of root system: the tap-root system, in which there is a main primary root larger than the other branching roots; and the diffuse (or fibrous) root system, in which there are many slender roots with numerous smaller root branches. Roots are outgrowths of trichoblasts, and elongated by tip growth, a mechanism that is used in several important cellular systems including pollen tubes, fungal hyphae and Rhizobium-induced infection threads. The patterning, differentiation, and growth of root have been elucidated by genetic and cell biological studies, but many aspects of root function still remain unknown.

Single-pass partial sequencing of clones from cDNA library of a specific tissue to generate expressed sequence tags (ESTs) has been proven to be a rapid and efficient means of discovering genes on a large scale

Okubo et al, 1992). It also provides both quantitative and qualitative information regarding gene expression in a variety of tissues and cells (Adam et al., 1991). It may be a cost-efficient way to establish a useful approach for gene identification. ESTs generated are compared with databases of identified genes and then as a guide to assign putative identifications to the cDNAs. These methods have accelerated research by providing genetic material for further investigation.

Materials and Methods

cDNA Library Construction

A total of 5 µg of poly-A^+ mRNA, isolated from 6-months-old oil palm root system, was used to construct a cDNA library using Stratagene ZAP-cDNA synthesis kit according to the protocol described. First-strand cDNA synthesis is primed with an oligo (dT) linker primer with a XhoI cloning site and is transcribed using MMLV-RT and 5-methyl dCTP. The 5' end of each cDNA was ligated to an adapter with an EcoRI-compatible overhang. cDNA was ligated unidirectionally into the EcoRI and XhoI sites of Uni-ZAP XR vector, packaged in vitro and amplified. The phage library was randomly selected and converted to plasmid form by single-clone excision according to the procedure described by Stratagene. Lambda phage were co-infected with ExAssist helper phage into E.coli strain XL1-Blue MRF' and the bacteria were grown for 2.5 h. The culture supernatant containing single stranded phagemid form was used to infect E.coli strain SOLR and then streaked on the ampicillin-LB agar plate to obtain single colonies. Bacteria were cultured overnight and used directly for double stranded plasmid DNA preparation using alkaline lysis method, Sequencing reactions were performed by using a DYEnamic™ ET terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech) with the T3 primer. Reaction products were electrophoresed on an automated DNA sequencer (ABI PRISM 377, Applied Biosystems and Megabace 1000, Amersham Pharmacia Biotech). Analysis of nucleotide and protein sequences were searched against GenBank and dbEST databases for homology comparison. by using BLAST and FASTA program (www.ingene.upm.edu.my/blast.html)

Results and Discussion

cDNA library was constructed from the mRNA of oil palm root system using oligo (dT) as a primer for the first-strand cDNA synthesis. In this study, two kinds of cDNA libraries were constructed, where one library was the cDNA insert

with size above 1000bp and the other with size between 500bp and 1000bp. Both of these libraries had primary recombinant clones of 10^6.

In this study, a total of 1400 cDNA clones were randomly selected and sent for nucleotide sequencing. Of these 1400 clones selected, 1121 clones produced some length of readable sequence, whereas 279 clones did not. From these readable sequences, 868 clones (77.4%) were found to be homologous to the sequences exist in NCBI database, 112 clones (10%) were the unknown genes, 100 clones (8.9%) did not show any homologous, and 41 clones (3.7%) were found no insert in comparison to the database. From all the clones sequenced, there are 10.37% of expression was accounted for the unknown genes and 9.26% was no homology. Although the functions of these genes are not yet known, they may be allowed to assign their functionality once the functional information is gained from any species

In short, the clones sequenced ESTs were categorized into 16 categories. These includes:

Cell wall structure or metabolism
Cytoskeleton
Membrane-associated genes
Signal transduction
Genes involved in cell division cycle
Stress- and defense-induced genes
Vesicular trafficking, protein sorting and secretion
Chromatin and DNA metanolism
Gene expression and RNA metabolism
Secondary and hormone metabolism
Primary metabolism
Protein synthesis and processing
Hypothetical protein
Miscellaneous
Unknown protein; No homology

Conclusions

Since the root mRNA was prepared from the actively growing and differentiating cells, most of the ESTs expression appeared to be the genes that involved in primary metabolism and protein synthesis and processing are expected, i.e. about 12.50% and 11.48% respectively. Some of the genes were found to be interested in primary metabolism, these included FKF1-like protein, grpE like protein, and thioredoxin h. In the category under stress- and defense-related function, some of the genes were found to be interested, such as wound-induced protein, hypersensitive-induced response protein, drought-induced protein, and Chaperone protein DNA K, which is a heat shock protein. Although this category was existed in a small percentage (3.06%), it should be given appropriate study on the possible involvement, or suppression, of host defense. Besides these, several other sequences are especially interesting given their homology to some genes with known function in other species. These clones included shaggy-related protein kinase, MAP kinase, Mei 2-like protein, AP2 protein, NDX1 homeobox protein, OsNAC4 protein, knotted1-type homeobox protein, scarecrow-like 6 protein, Skp1 protein, metallothionein-like protein, and others.

Benefits from the study

This project has enabled large scale isolation of genes that are functioning in the oil palm root system. Further characterization of these genes using functional genomic studies will enable the manipulation for increase in nutrient adsorption in oil palm or increase in disease resistance to ganoderma

Literature cited in the text

1. Okubo K., Hori N., Matoba R., Niiyama T., Fukushima A., Kojima Y., Matsubara K. 1992. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nat Genet* 2: 173-179

2. Adams MD., Kelly JM., Gocayne JD., Dubnick M., Polymeropoulos M., Xiao H., Merril C., Wu A., Olde B., Moreno R. 1991. Complementary DNA sequencing expressed sequence tags and the human genome project. *Science* 252: 1651-1656

Project Publications in Refereed Journals:

Nil

IRPA Project number: 01-02-04-0454
UPM Research Cluster : AFF
Project Leader Dr. Tan Siang Hee