

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

## Global gene expression patterns in rice

ArticleInfo		
ArticleID	:	3544
ArticleDOI	:	10.1186/gb-2000-1-1-reports010
ArticleCitationID	:	reports010
ArticleSequenceNumber	:	35
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 1999-11-29 Received : 1999-11-29 OnlineDate : 2000-4-27
ArticleCopyright	:	BioMed Central Ltd2000
ArticleGrants	:	
ArticleContext	:	130591111

Philippe Reymond

## Abstract

Correlation-based clustering of plant expressed sequence tags allows conclusions to be drawn about the relationship between genes based on their expression patterns in different tissues.

## Significance and context

Growth of higher plants is accompanied by the development of specific organs, tissues and cell-types. It is increasingly clear that developmental changes are mirrored by global changes in gene expression and that valuable information will be gained when all these molecular changes can be monitored simultaneously. In many cases, the function of a large number of unknown genes might be inferred from their expression profiles. New technologies such as cDNA microarrays or oligonucleotide chips enable analysis of the abundance of thousands of transcripts. A similar but '*in silico*' approach takes advantage of large-scale, single-pass partial sequencing of cDNA clones (expressed sequence tags, ESTs) from a large numbers of libraries. This approach assumes that the representation of a cDNA in a database is proportional to the abundance of the cognate transcript in the tissue or cell type used to make the library. Using sequence information from rice ESTs, the authors of this paper present a rigorous statistical method that enables both the association of genes on the basis of their tissue-dependent expression patterns and the association of plant tissues via their common patterns of gene expression.

## Key results

The authors used 10 rice cDNA libraries represented in [dbEST: database of expressed sequence tags](#). Each library contained at least 890 ESTs and was, in most cases, prepared from a different tissue or developmental stage. ESTs were organized into clusters and contig sequences, and expression profiles (EST counts) were derived for each of 707 contigs containing five or more constituent ESTs. In order to identify genes exhibiting a similar expression pattern, a statistical method (Pearson correlation coefficient) was used to calculate similarity between pairs of genes. These pairs of contigs were then organized into mutually matching clusters. The authors show, for example, that genes encoding storage proteins are clustered together and are predominantly found in libraries prepared from immature seed and panicle at ripening stage. The method is also successfully used to assess pairwise similarity between whole cDNA libraries and shows that two tissues expressing a similar complement of genes are clustered together. Finally, a two-dimensional graphical representation of expression measurements is

presented which allows a rapid visualization of clusters of genes obeying similar expression patterns in different conditions (different libraries).

## Links

A method for EST quality control and generation of contigs can be found at the [Structural and genetic information](#) server.

## Reporter's comments

Convincing evidence is provided that a rigorous statistical analysis of EST libraries allows fine-scale identification of sequences with correlated expression profiles. The application of this approach to a large collection of cDNA libraries prepared from different organisms at different developmental stages will certainly provide a valuable alternative to cDNA microarray studies in generating gene expression data. A limitation of such a technique is the need for standardization of the preparation of cDNA libraries to ensure that EST frequency tightly correlates with transcript abundance. As the method relies on the availability of sequence information from EST libraries it will also require such large-scale programs to be continued. An interesting use of the protocol presented in this paper could be to compare the cDNA libraries prepared with tissues or cell-types from distantly related species, something that is not currently feasible with cDNA microarrays because of the lack of sequence homology.

## Table of links

*Genome Research*

[dbEST: database of expressed sequence tags](#)

[Structural and genetic information](#)

## Access

The clustered correlation map and associated results presented in the paper are available from the authors.

## References

1. Ewing RM, Kahla AB, Poirot O, Lopez F, Audic S, Claverie JM: Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res.* 1999, 9: 950-959. 1088-9051