Proceedings of the XLIX Italian Society of Agricultural Genetics Annual Congress
Potenza, Italy – 12/15 September, 2005
ISBN **88-900622-6-6**

Oral Communication Abstract - 6.03

ANALYSIS OF THE "METABOLIC" GENOME' OF ARABIDOPSIS USING MICOARRAY DATA: IMPLICATIONS FOR PLANT BIOTECHNOLOGY

P. MORANDINI*, L. MIZZI**, M. SOLOVIEV***, G. MARINO*, M. VESCOVI****,
M. POOLMAN*****, D. A. FELL*****

*) Dipartimento di Biologia, Università di Milano (Italy) - piero.morandini@unimi.it
**) Dipartimento di Scienze Biomolecolari e Biotecnologie, Università di Milano (Italy)
***) School of Biological Sciences, Royal Holloway, University of London, UK
****) Research Systems Italy Srl, Agrate Brianza, Milano (Italy)
*****) School of Biological & Molecular, Sciences, Oxford Brookes University, Oxford, UK

Each microarray experiment is capable of producing measurements of transcript level for many thousands of genes. The analysis of such large amount of data is usually restricted to the comparison among experiments (time series or treated vs control). The reason for this is twofold: such results are easy to produce with little computation (i.e. identify the genes showing the largest variation among slides) and easy to grasp (induction or repression of single genes). In our approach we analyzed the genes coding for metabolic enzymes of *Arabidopsis thaliana* using two-gene scatter plots over all the publicly available data from a microarray database. The Pearson Correlation coefficient of each scatter plot was used as a measure of the co-regulation for each pair of transcripts. Software tools were constructed to calculate the Pearson coefficient i) of a gene against all the other genes present on the microarray chip and ii) of all possible gene pairs from a given list generating matrices of correlation coefficient. A graphical representation of such matrices was devised to allow a bird's eye view on the entire metabolic genome of Arabidopsis.

Analysis of the matrices reveals that, at the genomic level, the phenomenon of parallel activation is rather common in major pathways of plant metabolism, at least at the transcript level, thus suggesting that the best approach for metabolic engineering of endogenous pathways is the coordinate induction of enzymes (the so called 'universal method', see Morandini and Salamini, 2003 Trends Plant Sci. 8:70-5). It is mandatory to stress that the expression level for a gene does not coincide with the protein level, but can be used, in most cases, at least as a gross index for the activity.

Many predictions on the role of different enzymes isoforms in the carotenoid and sterol pathways, on the linking of metabolic pathways and on the function of several genes will be presented.