

## OLIVE FRUIT TRANSCRIPTOME ANALYSIS THROUGH 454 PYROSEQUENCING

ALAGNA F.\*, BALDONI L.\*, PERROTTA G.\*\*, TORCHIA L.\*\*, GIULIANO G.\*\*\*,  
D'AGOSTINO N.\*\*\*\*, CHIUSANO M.L.\*\*\*\*

\*) CNR – Institute of Plant Genetics, Via Madonna Alta 130, 06128 Perugia

\*\*) ENEA, Research Center TRISAIA, S.S. 106 Ionica, 75026 Rotondella (Matera)

\*\*\*) ENEA, Research Center CASACCIA, S.M. Galeria 00163, Roma

\*\*\*\*) Dept. of Soil, Plant, Environmental and Animal Production Sciences, University of Naples 'Federico II', Via Università 100, 80055 Portici (Naples)

*454 sequencing, EST, transcriptome, Olea europaea, secoiridoids*

The application of a massively parallel pyro-sequencing technology (by 454 Life Sciences Corporation) provides hints to improve the characterization of the olive fruit mesocarp transcriptome, with a focus on terpenoid metabolism, due to the importance of these compounds on the mechanism of pathogen resistance and their effects on human health.

Four different cDNA libraries were prepared applying the SMART technology (Clontech) from fruits of a cultivar characterized by a high oleuropein (main terpene secoiridoid) content and an oleuropein-lacking natural variant, respectively at the beginning and at the end of fruit development.

Five µg of each purified cDNA library were provided for 454 sequencing.

A total of about 260,000 reads were generated for all of the samples with a total output of 58 Mb. Between 50,000 and 77,000 reads were obtained for each sample. The average read lengths were between 217 and 224 bp. Quality filtered sequences from whole shotgun sequencing were *de novo* assembled, obtaining contigs up to 3,920 bp long.

The ParPEST (Parallel Processing of ESTs) pipeline has been tweaked and redesigned in order to manage assembled 454 data. EST assemblies will be annotated and classified according to their biological functions. In addition, we are going to evaluate and compare gene expression patterns as they reflect the composition of the EST libraries prepared from different genotypes and developmental stages. Raw data processing is in progress and results will be adequately reported and organized into a MySQL relational database. Finally, an user-friendly web interface will be created to allow data browsing.