Proceedings of the 50th Italian Society of Agricultural Genetics Annual Congress
Ischia, Italy – 10/14 September, 2006
ISBN **88-900622-7-4**

Poster Abstract – D.53

AN INTEGRATED COMPUTATIONAL PLATFORM FOR TOMATO GENOMICS

M. L. CHIUSANO*, N. D'AGOSTINO*, A. TRAINI*, L. FRUSCIANTE**

*) Department of Structural and Functional Biology, University 'Federico II', 80134 Naples, Italy
**) Department of Soil, Plant and Environmental Sciences, University 'Federico II', 80055 Portici (NA), Italy

Solanum lycopersicum, EST data-banks, experimantal genome annotation, Genome Browser

Motivation: We present here our effort as partners of the Solanaceae (SOL) Genomics Network. The long term goal of the Consortium is to build a network of resources and information dedicated to the biology of the Solanaceae family which includes many species of relevant agricultural interest.

In the frame of the International Tomato Sequencing Project (http://www.sgn.cornell.edu/), to efficiently exploit genomics data generated within the Consortium, the Bioinformatics Committee coordinates data management and integration and will offer analysis tools in a distributed platform to support the research of all the SOL partners.

As participants to the Bioinformatics unit of the Network, in the first year project we set up a workbench to support the experimental annotation of the *Solanum lycopersicum* (tomato) genome as well as the analysis of sequence collections from other Solanaceae species.

Methods: Using an automated approach (D’Agostino *et al.*, BMC Bioinformatics 2005) we built and maintain two repositories of ESTs from tomato and potato species downloaded from dbEST. The repositories provide a complete set of computationally defined transcript indices representing unique sequences, singletons or tentative consensus (TC), derived from EST clustering analysis. We based the annotation of the expressed sequences on the use of controlled vocabularies such as the Gene Ontologies (GO; The Gene Ontology Consortium 2000) and the Enzyme Commission numbers (EC; Bairoch, 2000) and implemented the ‘on the fly’ mapping of the expressed sequences onto known metabolic pathways from KEGG (Kaneisha et al., 2004). We included in the annotation pipeline the analysis of non coding RNAs and we provided, for each tomato EST, a link to the identification numbers of TOM1, the reference cDNA microarray for Tomato (TED; http://ted.bti.cornell.edu/).

The preliminary experimental annotation of the BAC sequences available from the SOL Genomics Network was based on the experimental data from tomato and potato ESTs and TCs. The annotated BACs are available by the Generic Genome Browser (GBrowse) interface to allow selection for reference *gene models* to train predictive methods. The two EST data-banks and the Generic Genome Browser are cross-referenced to provide an integrated platform.

Results: The platform was built to provide an Italian resource for the genomics of Solanaceae family and for the International Tomato Genome Sequencing Project. The web based interface of the platform can be accessed browsing genome data in the form of BAC sequences or querying the EST repositories. The platform is useful to support the experimental annotation of the genomic data, indeed, EST collections are a quick route for discovering new genes and for confirming coding regions in genomic sequences. Moreover, we believe that the possibility to investigate on specific expression patterns as well as on coding or non coding gene families from the annotated EST databanks provide a relevant support for the investigation and the comprehension of genome organization and functionalities. We may well hope that our effort represent a framework for a useful organization of structure genomics data and for meaningful functional analyses based on comparative approaches with other model plant species to provide a reference within the Solanaceae community as well as for other similar efforts.

Acknowledgements:This work is supported by the Agronanotech Project (Ministry of Agriculture, Italy) and by the EU-SOL Project (VI Frame Programme, European Union).

URL: http://biosrv.cab.unina.it