ISOL@ (the Italian SOLAnaceae genomics resource) evolution: novel insights into the structure and the function of the tomato genome

Chiusano ML, D'Agostino N, Di Filippo M, Maselli V, Traini A, Frusciante L

Dept. of Soil, Plant, Environmental and Animal Production Sciences,
University of Naples "Federico II"

Motivation

The 950Mb tomato genome is structured into distal gene-rich euchromatin and gene-poor peri-centromeric heterochromatin. The strategy adopted by the International Tomato Genome Sequencing Consortium is to initially sequence, using a BAC-by-BAC strategy, the euchromatic portion of the genome which corresponds to one quarter of the tomato genomic sequence (220Mb) and where the majority of genes (~90%) resides. To render the emerging tomato genomic sequences immediately usable by the community, while the international Tomato Annotation Group (iTAG) is still setting up an official annotation, we designed and implemented ISOL@, an Italian SOLAnaceae genomics resource, accessible at:
 http://biosrv.cab.unina.it/isola/.

Methods

ISOL@is conceived as a multi-level computational environment accessible through different gateways. The 'transcriptome' gateway provides an access point to explore publicly available EST collections from Solanaceae species and to inspect the corresponding virtual transcripts generated by assembling ESTs into tentative consensus sequences. The 'genome' gateway allows to browse the EST-based annotation of the tomato BAC sequences. The resulting identification of the 'expressed' loci is exploited for the definition of gene models and for the identification of putative alternative transcripts. Annotations of repeat elements, based on the Plant Repeats database at Michigan State University, are included as well.

Results

To date ISOL@ is set up to provide a preliminary view on the draft and incomplete tomato genome. We measured the repeat content and gene distribution for each BAC, providing a quick and an high resolution preview of the genome composition. In fact, despite the generally accepted idea that the euchromatic regions were gene rich, we found a number of BACs showing low gene and high repeat content outside the peri-centromeric regions. ISOL@ is expanding its facilities, through the addition of novel accessory modules, in order to provide further insights on the tomato genome structure, function and evolution. The flexibile and multi-level structure of the platform, indeed, permitted to develope tools to drive comparative genomics analysis with reference plant species and to link expression data, as revealed by real-time RT-PCR analyses, to transcriptome data (this topic is in collaboration with the Plant Research International in Wageningen, Netherlands). Finally, since proteomics and metabolomics data are the next waves of data becoming available in Solanaceae genomics research, they will be easily integrated into the ISOL@ context through pre-defined entry points.

Contact : chiusano@unina.it