

A PLANT GENOME ANNOTATION PIPELINE FOR DISEASE RESISTANCE GENES

SANSEVERINO W.*, ROMA G.***, FRUSCIANTE L.*, ERCOLANO M.R.*

*) Department of Soil, Plant, Environmental and Animal Production Sciences, School of Biotechnology, University of Naples “Federico II”, Via Università 100, 80055 Portici (Italy)

**) Tigem – Telethon institute for genetic and medicine

resistance gene, plant disease, database, bioinformatic

Plant disease resistance genes (R-Genes) are an important class of genes which are well characterized at the molecular level. These genes play a key role in the recognition of the products of avirulence (Avr) genes from pathogens and in the activation of plant defence responses. So far, in the Solanaceae family, 29 R-genes have been isolated and characterized. Using these genes as reference sequences we have built a bioinformatic pipeline for fishing R-gene in large dataset (whole genomes, “Unigene” clusters and NCBI sequence collection). Combining phylogenetic data with conserved domain dissection 400 sequences were extrapolated from NCBI dataset. Sequences have been download ordered and stored in our SRG database (<http://srg.cbm.fvg.it/index.php>). A graphic view of putative cDNA and protein sequence structure with indication of resistance gene domain location is provided to facilitate data analysis. With this procedure we were able to fish and order all Solanaceae putative resistance sequences accurately and quickly. Our future goal is to build an automatic system for fishing sequences related to disease in all the new sequenced plant genomes.