

COMBINED GENOMIC AND TRANSCRIPTOMIC APPROACH TO UNCOVER CANDIDATE GENES FOR HEAT STRESS RESPONSE OF A THERMOTOLERANT TOMATO GENOTYPE

GRACI S.*, AIESE CIGLIANO R.***, RUGGIERI V.***, BARONE A.*

*) Department of Agricultural Sciences, University of Naples Federico II, Portici, 80055 Naples, Italy

**) Sequentia Biotech, SL, Carrer del Dr. Trueta 179, 08005 Barcelona, Spain

***) Biomeets Consulting, Carrer d'Àlaba, 61, 08005 Barcelona, Spain

SNP, InDel, high temperatures, wild species, gene network

Climate change caused by a rise in temperatures under natural conditions is predicted to significantly affect tomato plant growth and development, particularly its reproductive stage and leaf traits, potentially leading to a dramatically reduction in crop productivity. In this scenario, constitution of heat-tolerant genotypes that can survive high-temperatures represents a priority. This work aimed at identifying candidate genes (CGs) involved in the heat stress (HS) response of the thermo-tolerant E42 genotype and investigate their regulatory network. Phenotypic data recorded under normal and HS conditions evidenced that E42 produced a high number of flowers and fruits, leading to high yield despite the fruit set being around 50%. Whole-genome re-sequencing data obtained from genotype E42, available at the Department of Agricultural Sciences of the University of Naples Federico II, was used to study the genetic bases of the observed phenotype. Raw reads were processed, mapped and filtered (Trimmomatic, Bowtie-2, Samtools, BCFtools, VCFtools) to obtain variants. Moreover, SnpEff v. 4.5 analysis was performed, using the SL4.0 tomato genome version and the ITAG4.1 version of the tomato annotation, to identify the associated genes and to predict the putative effects of SNPs and InDels. In addition, publicly available RNAseq data (GEO project codes: GSE152620, GSE199011, GSE148217 and GSE163914) performed on leaf, fruit and flower bud tissues of different tomato genotypes were retrieved from NCBI in order to investigate the regulatory network of candidate genes. Gene Coexpression Network (GCN) analysis was performed with BioNero package of Bioconductor

and results were visualized on Cytoscape platform. The variant calling analysis showed a high variability for chromosomes 1, 4, 7 and 12, which may have arisen from introgressions from heat tolerant *S. pimpinellifolium* wild species, according to our phylogenetic analysis. The SnpEff analysis allowed us to identify variants with HIGH (18) and MODERATE (139) impact from two lists of genes related to heat and reproduction. This analysis allowed to prioritize a subset of 35 CGs (Hsfs, Hsps, flower- and pollen-related) mapping in polymorphic regions of E42, some colocalizing with QTLs controlling flowering in tomato. To investigate the relationships between these polymorphic genes, GCN analysis was performed. Results highlighted one module involving six Hsps and two Hsfs (HsfA3 and HsfC1) that, following a gene enrichment analysis, presented Gene Ontology terms related to HS response. Among these genes, the LeHsp100 (Soly02g088610) presented the highest number of HIGH and/or MODERATE variations (four) and was reported to contribute to the acquisition of thermotolerance during the HS response. Further analysis will be performed on these CGs of E42 to validate their expression levels under normal and HS conditions and to shed light on the genetic bases of its HS responsive mechanisms.