Proceedings of the LXV SIGA Annual Congress
Piacenza, 6/9 September, 2022
ISBN: **978-88-944843-3-5**

Poster Communication Abstract – 4.08

EXPLOITING RE-SEQUENCING DATA OF A THERMO-TOLERANT GENOTYPE TO ENHANCE HIGH-TEMPERATURE RESPONSE IN TOMATO

GRACI S.*, RUGGIERI V.**, BARONE A.*

*) Department of Agricultural Sciences, University of Naples Federico II, Portici, 80055 Naples, Italy
**) Biomeets Consulting, Carrer d´Àlaba, 61, Barcelona, Spain

SNPs, InDels, phylogeny, wild species, candidate genes

Global climate change and rising temperatures negatively affect tomato cultivations, comporting yield-losses production caused by alteration of metabolic activities, photosynthesis, cellular divisions, protein folding, flower abscission, poor pollen germination. In this scenario, development of heat-tolerant genotypes that can survive high-temperatures represents a priority. This work aimed at investigating the genome sequence of the thermo-tolerant E42 genotype to explore the origin of its genetic variability and to identify candidate genes involved in high-temperatures response. To these aims, whole-genome re-sequencing E42 data, available at the Department of Agricultural Sciences, were used to investigate the referred genotype. Raw reads were processed, mapped and filtered (Trimmomatic, Bowtie-2, Samtools, BCFtools, VCFtools) to obtain variants. In order to investigate the origin of the genetic variability, a phylogenetic analysis was performed using the neighbor-joining (NJ) method implemented in VCF-kit tool (http://vcf-kit.readthedocs.io/), with a dataset of genomic wild species sequences retrieved from the European Bioinformatics Institute (PRJEB5235, Sahu and Chattopadhyay 2017). Lastly, looking for heat-tolerance related genes, SnpEff v. 4.5 analysis was performed, using the SL4.0 tomato genome version and the iTAG4.1 version of the tomato annotation, in order to annotate the genes and to predict the putative effects of the SNP and InDel mutations of the E42 genotype. Results of variant calling analysis evidenced E42 polymorphisms accumulation on chromosomes 1, 4, 7 and 12. Phylogenetic analysis conducted on the 12 chromosomes of 82 accessions belonging to 13 tomato species showed that E42 clustered mostly with *S. lycopersicum* accessions, while the

cluster also involved *S. pimpinellifolium*, *S. galapagense* and *S. cheesmaniae* in the four polymorphic chromosomes. Finally, SnpEff analysis was performed to estimate the probable impact of SNP and InDel mutations on genes' expression and proteins' translation. Out of 2,394,426 mutations (2,078,012 SNPs and 316,414 InDels), E42 sequence showed 2,773 HIGH variants (937 SNPs and 1,787 InDels) and 19,066 MODERATE variants (18,347 SNPs and 638 InDels). HIGH and MODERATE variants interested 2,153 and 7,244 genes respectively, among which 1,300 and 5,075 genes mapped on chromosomes 1, 4, 7 and 12. Among these, several genes were involved in heat stress response, including genes coding for members of the FACT subunit complex SSRP1 protein family, transcription repressor OFP3, isocitrate dehydrogenase [NADP], adenine nucleotide alpha hydrolases-like superfamily protein, class I heat shock protein, DnaJ protein like, auxin efflux carrier component, flowering locus T. Further analysis will be performed on promoter regions to found regulatory elements that direct high-temperatures genes' expression and to identify sources of E42 heat stress response, also considering its genetic variability origin.