Proceedings of the LX SIGA Annual Congress
Catania, Italy – 13/16 September, 2016
ISBN **978-88-904570-6-7**

Poster Communication Abstract – 7.07

WHOLE-GENOME RE-SEQUENCING OF TWO TOMATO LANDRACES REVEALS SEQUENCE VARIATIONS UNDERPINNING KEY ECONOMICALLY IMPORTANT TRAITS

TRANCHIDA-LOMBARDO V.*, AIESE-CIGLIANO R.**, ANZAR I.**, LANDI S.*, PALOMBIERI S.*, COLANTUONO C.***, BOSTAN H.***, TERMOLINO P.*, AVERSANO R.***, BATELLI G.*, CAMMARERI M.*, CARPUTO D.***, CHIUSANO M.L.***, CONICELLA C.*, CONSIGLIO F.*, D'AGOSTINO N.****, DE PALMA M.*, DI MATTEO A.***, GRANDILLO S.*, TUCCI, M.*, SANSEVERINO W.**, GRILLO S.*

*) Institute of Biosciences and Bioresources (CNR-IBBR), Via Università 133, 80055 Portici (Italy)
**) Sequentia Biotech, Calle Baldiri-Reixac 4-8, Barcelona, 08028 (Spain)
***) Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, Portici, 80055 (Italy)
****) Consiglio per la Ricerca in Agricoltura e l'analisi dell'Economia Agraria, Centro di Ricerca per l'Orticultura (CREA-ORT), Via Cavalleggeri 25, 84098 Pontecagnano Faiano (Italy)

tomato landraces, Next generation sequencing, quality, stress, shelf life

In the post-genomic era, one of the major challenges is the identification of alleles directly responsible for phenotype variation among different genotypes within the same species. Tomato is a model crop for understanding the development and ripening of climacteric fleshy fruits, and it is also known to be an important source of health-promoting compounds. In addition, cultivated tomato germplasm shows a high phenotypic variation despite its very low genetic diversity. Toward the identification of sequence variations responsible for stress tolerance, high fruit quality and long shelf life, we re-sequenced the genomes of two traditional landraces grown in the Campania region (Southern Italy). Crovarese, belonging to the Corbarino type (COR), and Lucariello (LUC) are typically grown under low water regimes and produce highly appreciated fruits, which can be stored up to 4-8 months. We generated 65.8M and 56.4M of paired-end 30-150 bp reads with an average insert size of 380 bp (± 52bp) and 364 bp (± 49bp) for COR and LUC, respectively. A reference-guided assembly was performed using 'Heinz 1706' as a reference genome. We estimated a mean coverage depth of ~15X for COR and 13X for LUC. Comparing the genomes of COR and LUC with that of 'Heinz 1706' we found a similar distribution of SNPs (68.8% vs. 69.9%, respectively), small deletions (8.9% vs. 8.6%) and small insertions (22.1% vs. 21.3%). Through a *de novo* assembly of the unmapped reads we identified 29 and 36 new contigs in COR and LUC, respectively. The new contigs could be assigned to the chromosomes thanks to the use of a split-read approach. On average, the contigs inserted in COR were 654bp, whereas those inserted in LUC were 616bp. Using custom RNA-seq data, a total of 43054 and 44576 gene loci were annotated in COR and LUC, corresponding to 62369 and 65094 transcripts, respectively. Among the genes showing a similar structure in COR and LUC compared to 'Heinz 1706', we identified ~2000 and 1700 SNPs causing potentially disruptive effects on the function of 1371 and 1201 genes in COR and LUC, respectively. Interesting GO categories highly represented in genes affected by sequence changes were identified. Major variations were present in stress-responsive genes as well as in fruit quality and development-related genes. From a practical perspective, the identified SNPs and

InDels are candidate polymorphisms to track DNA variations associated to key traits of economic interest.

This work was funded by the Italian Ministry of Education, University and Research, Project PON02_00395_3082360 - GENOPOMpro