



COST Action
CA 17111
INTEGRAPE

1st Annual Meeting

INTEGRAPE 2019

**Data Integration as a key step for future
grapevine**

Book of Abstracts



Chania, Greece
25 - 28 March, 2019



Funded by the Horizon 2020
Framework Programme
of the European Union



NES2RA: A TOOL FOR GRAPEVINE TRANSCRIPTOMIC DATA MINING

PILATI STEFANIA¹, MALACARNE GIULIA¹, VALENTINI SAMUEL², MORETTO MARCO³, ASNICAR FRANCESCO², MASERA LUCA², SONEGO PAOLO³, CAVECCHIA VALTER⁴, BLANZIERI ENRICO², MOSER CLAUDIO¹

¹FONDAZIONE EDMUND MACH, Department of Genomics and Biology of Fruit Crops, Research and Innovation Centre, SAN MICHELE ALL'ADIGE, Italy, ²UNIVERSITY OF TRENTO, Department of Information Engineering and Computer Science, TRENTO, Italy, ³FONDAZIONE EDMUND MACH, Unit of Computational Biology, Research and Innovation Centre, SAN MICHELE ALL'ADIGE, Italy, ⁴CNR, Institute of Materials for Electronics and Magnetism, TRENTO, Italy

The development of “omics” technologies to study gene expression has revolutionized our perspective from the single gene to the gene network level. However, the complexity of the system biology approach requires appropriate mathematical, computational and statistical tools to analyze data and extract information. Grapevine transcriptomic data are currently collected in two databases: the VITis Co-expression DataBase (VTCdb, Wong et al., 2013) dedicated to data obtained with microarray technology and the Vitis Expression Studies Platform Using COLOMBOS Compendia Instances (VESPUCCI, Moretto et al., 2016) including data from both microarrays and RNAseq experiments. Here, we present the application of the algorithm of Network Expansion by Subsetting and Ranking Aggregation (NES2RA, Asnicar et al., 2016) to expand Local Gene Networks (LGN) in grapevine using transcriptomic data stored in the VESPUCCI compendium. NES2RA is based on the PC-algorithm (Spirtes and Glymour, 1991), a gaussian graphical model (GGM) that finds causal relationships from observational data. It is based on a systematic test for conditional independence to retain significant relations between pairs of genes. It starts from a fully connected network and removes interactions between genes, whenever it finds a set of genes that supports that interaction (i.e., separation set). Due to the computational power requirement of NES2RA algorithm, it has been running as part of the gene@home project, a distributed computation project which relies on thousands of volunteers' computers by means of the TN-Grid, an infrastructure based on BOINC system (Asnicar et al., 2015). NES2RA has been used to expand four LGNs related to the grapevine response to climate changes (Malacarne et al., 2018). The obtained expansion gene lists have been analyzed by means of statistical tools - such as gene annotation and functional categories enrichment to assess the functional coherence between LGNs and expansion gene lists and promoter analysis to test co-regulation among output genes - and compared with experimental results, when available, and literature. These analyses produced promising results in support of the meaningfulness of this approach. Moreover, the LGNs expansions can be visualized as networks, thus providing the biologist with a prompt information about the significant relationships retained by NES2RA, highlighting positive or negative correlations within gene pairs. We are currently developing NES2RA algorithm to make it available as a web tool to be used in real time and exploring new applications.