

## **P22 – FEMVitisDB: a FAIR data management system for data integration in grapevine**

Vezzulli, Silvia; Bianco, Luca; Nicolini, Daniela; Betta, Giulia; Dorigatti, Cinzia; Zulini, Luca; Donati, Claudio; Costantini, Laura; Stefanini, Marco\*; Fontana, Paolo; Micheletti, Diego  
Fondazione Edmund Mach, Research and Innovation Centre, San Michele all'Adige, Italy  
\*marco.stefanini@fmach.it

### **Abstract**

A grapevine germplasm collection including more than 2,000 unique genetic profiles, distinguished by their origin and/or distinctive phenotypic characteristics, is maintained and preserved at the Fondazione Edmund Mach (FEM, Italy). Each genotype is represented by 1-20 accessions. New acquisitions from repositories around the world and wild collections consist of released or naturalized *Vitis* hybrids as well as *Vitis* species, useful to scout new genetic resources for diversity studies and breeding purposes (wine/table grapes and rootstocks). Most accessions were genotyped with the reference set of 9 microsatellite markers to verify their trueness-to-type against international and national databases. Moreover, the entire germplasm collection has been phenotyped during three successive years for several traits related to ampelography, vine development (e.g. phenological stages, fertility), biotic stress response, berry and wine composition (e.g. chemical parameters). In parallel, crossbreeding programs have been carried out to obtain new genotypes coupling winemaking quality and resistance to different biotic stresses.

We developed the FEMVitisDB database to host the above information and that produced by the current and future grape breeding programs. This infrastructure is focused on the "FAIR" paradigm which emphasizes the importance of organizing data and metadata. The infrastructure core is a PostgreSQL database compliant with the MIAPPE (Minimum Information About a Plant Phenotyping Experiment) standard. The database can host both genotyping and phenotyping data allowing the storage of heterogeneous datasets including dense time series and high throughput genotyping data. The architecture is ontology driven (i.e. the semantics of all the terms can be specified by a suitable ontology) to provide a powerful system to integrate data originating from different experiments and platforms. Where feasible, the OIV code - as described in the *Vitis* ontology - has been applied for the standardization of the phenotype description; alternatively, the grape or plant ontologies have been used. The database is integrated with standardized RESTful Web Service API based on BrAPI (brapi.org) which provides access to the data. The data management system is completed by the BrAPIViewer, a web frontend returning all the necessary information in a user-friendly way.

**Keywords:** digitalization, data standardization, germplasm valorisation, genotyping, phenotyping