Data integration to maximise the power of omics for grapevine improvement

*Introduction to **Integrape**
data management guidelines*

*Stefania **SAVOI***

l'institut Agro | SupAgro Montpellier
agriculture · alimentation · environnement

COST

(European Cooperation in Science and Technology)
is a funding organisation for research and innovation networks

COST Actions

are network dedicated to scientific collaboration,
complementing national research funds

https://www.integrape.eu/

AIMS to develop minimal data standards and good practices in order to integrate data repositories and improve interoperability between datasets

COST Action
CA17111
INTEGRAPE

o 4 YEARS: Sept 2018 - Sept 2022

o 22 European countries

o 3 COST Near Neighbour countries

o 2 COST International partner countries

o Chair Prof. Mario Pezzotti

o Vice-Chair Prof. Anne-Francoise Adam-Blondon

o 4 working groups

Working Groups

Working Group 1

Data interoperability and definition of minimal contextual data standards

This Working Group will focus on contributing to community-wide standards for metadata along with guidelines for their proper use.

WG1 Leader: Reinhard TOEPFER

WG1 Vice-Leader: Camille Rustenholz

Prof Reinhard TOEPFER

Julius Kühn-Institute
Institute for Grapevine
Breeding
Geilweilerhof
Germany
Siebeldingen

Working Group 1 Composition

Working Group 2

Interoperability of infrastructures and web services

In this Working group we will develop a bioinformatics infrastructure to facilitate the discovery, acquisition, storage, processing and integration of diverse grapevine datasets, and this will be managed by distinct nodes.

WG2 Leader: Paul KERSEY

WG2 Vice-Leader: Daniela Holtgräwe

Dr Paul KERSEY

Royal Botanic Gardens
Kew
Richmond
Surrey
United Kingdom

Working Group 2 Composition

Working Group 3

Data analysis and best practices

Guidelines for best practice focusing on the promotion of standard protocols for omics data processing and analysis will be developed and recommended. These guidelines will help researchers to produce high-quality data and carry out informative analysis in the context of a standard framework, which will further contribute to data integration. Additionally, we will build a centralised framework for the standard annotation of grapevine genes. Such a framework will facilitate the integration of different omics data types by relying on standard formats and ontologies.

WG3 Leader: Johan TRYGG

WG3 Vice-Leader: Jerome Grimplet

Prof Johan TRYGG

Computational Life Science
Cluster (CLiC)
Department of Chemistry
Umeå University
Linneaus vag 10
Umeå
Sweden

Working Group 3 Composition

Working Group 4

Dissemination and user community assessment of guidelines and recommendations

A focus of this Working Group will be training and dissemination of the work done by the other WG possibly in joint activities with dissemination and training activities of other large agnostic initiatives. We will disseminate the recommendations made by other working Groups and train the grapevine community to follow these recommendations.
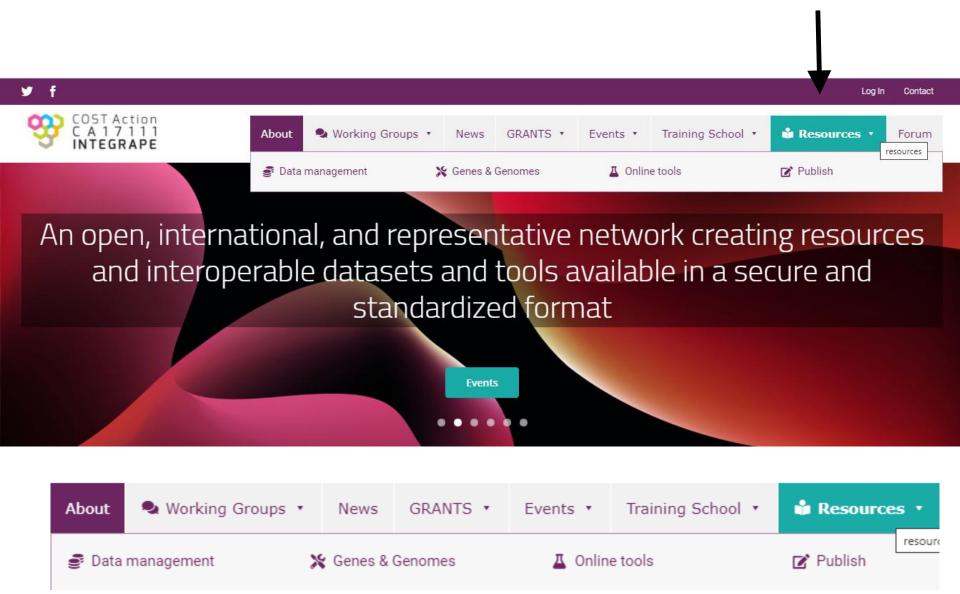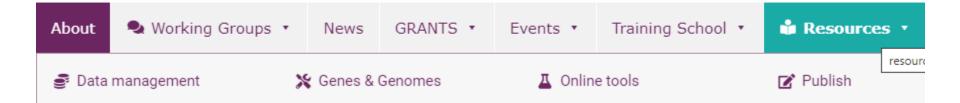
WG4 Leader: Dragan NIKOLIC

WG4 Vice-Leader: Panagiotis Kalaitzis

Prof Dragan NIKOLIĆ

University of Belgrade
Faculty of Agriculture
Nemanjina 6
Serbia
11080
Belgrade-Zemun
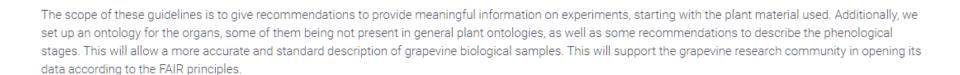
Working Group 4 Composition

COST Action
CA17111
INTEGRAPE

About | 💬 Working Groups ▾ | News | GRANTS ▾ | Events ▾ | Training School ▾ | 📖 Resources ▾ | Forum

resources

🪙 Data management | ✂ Genes & Genomes | ⚗ Online tools | 📝 Publish

An open, international, and representative network creating resources and interoperable datasets and tools available in a secure and standardized format

Events

● ● ● ● ● ● ●

📚 Data management   🛠 Genes & Genomes   ⚗ Online tools   📝 Publish

Data management

Guidelines for Data Management

The scope of these guidelines is to give recommendations to provide meaningful information on experiments, starting with the plant material used. Additionally, we set up an ontology for the organs, some of them being not present in general plant ontologies, as well as some recommendations to describe the phenological stages. This will allow a more accurate and standard description of grapevine biological samples. This will support the grapevine research community in opening its data according to the FAIR principles.

📝 *How to describe an experiment*

🧬 *How to submit sequence data to ENA*

🗄 *How to submit metabolomic data to MetaboLights*

🔗 *How to standardize JBrowse's tracks*

📘 *Apollo Manual Curation Guide for the PN40024.v4 assembly (under construction)*

How to describe an experiment

How to describe a grapevine experiment

The scope of these guidelines is to give recommendations about standard metadata for experimentation and sampling:
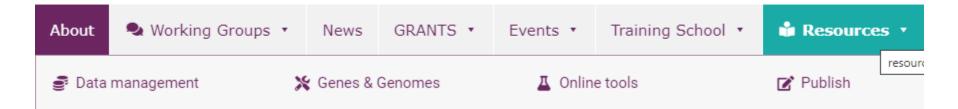
Metadata about the experiment set up ⌄

Identification of the plant material ⌄

Standard vocabulary for organs or plant anatomical entities ⌄

Standard description of development stages ⌄

How to submit metabolomic data to MetaboLights

Submission Process ⌄

Study Description ⌄

Protocols ⌄

a. Sample collection ⌄

b. Extraction (Sample Preparation) ⌄

c. Chromatography ⌄

d. Mass Spectrometry ⌄

e. Data Transformation ⌄

f. Metabolite Identification ⌄

Samples, Assays and Metabolites ⌄

Files ⌄

Validation Rules ⌄

Available Datasets ⌄

About	💬 Working Groups ▾	News	GRANTS ▾	Events ▾	Training School ▾	📖 Resources ▾

resourc

💰 Data management   ✗ Genes & Genomes   ⚗ Online tools   ✎ Publish

Genes & Genomes

Genes & Genomes

The INTEGRAPE Cost Action has created this centralised catalogue of genome accessions/annotations, characterised genes and surveyed gene families, as well as some recommendations for naming genes and gene families, according to the Super-Nomenclature Committee for Grape Gene Annotation.

The scope of these catalogue is to offer a unique repository with more accurate and standard description of grapevine genes and genomes. This will support the grapevine research community in opening its data according to the FAIR principles.

✎ *Genome accessions*

⧖ *Reference Gene Catalogue & Nomenclature recommendations*

Genome accessions

PN40024 40X New Assembly - Under Construction

PN40024 12X.2 Assembly (URGI)

Other wine grape cultivar and Vitis species assemblies/annotations (GrapeGenomics – University of California Davis)

Vitis PanGenome browser (under construction)

PN40024 12X.0 Assembly (CRIBI)- Obsolete

Reference Gene Catalogue & Nomenclature recommendations

Grapevine Gene Nomenclature

Grapevine Reference Gene Catalogue

New Gene Submission Form

Online Tools for the Analysis of Grapevine Data

The scope of this repository is to serve the grapevine community with a list of tools for genomics/transcriptomics analyses, including the visualization of gene expression data across publicly available data, generation of gene co-expression networks, etc.

- VESPUCCI – Vitis Expression Study Platform Using Command Compendium Instance ⌄
- GREAT – GRape Expression ATlas (under construction) ⌄
- VTC-Agg. – Aggregate and correlation rank grapevine gene co-expression networks ⌄
- OneGenE – One Gene network Expansion (under construction) ⌄
- Vitis Visualization (VitViz) ⌄
- Grape eFP Browser ⌄
- VitisCyc ⌄
- Mapman Tools adapted for Grapevine (under construction) ⌄
- Gramene database ⌄
- Vitis PanGenome browser (under construction) ⌄

*If you wish do add your own tool or analysis resource send a mail to **tomas.matus (at) uv.es** with your request.*

https://www.integrape.eu/

Thanks for your attention