Resolution of resistance genes within melon QTL regions by optical mapping

A Data Management Plan created using DMPonline

Creators: Dick de Ridder, Sven Warris

Affiliation: Other

Funder: ZonMw (Nederlands)

Template: Data management ZonMw-template 2019

ORCID iD: 0000-0002-4944-4310

Grant number: 435005027 (40-43500-98-5038)

Project abstract:

Plant genomes are highly complex, variable and difficult to assemble. Most assemblies in plants are performed to provide information on complex genomic regions associated with important traits, such as resistance to pests, pathogens and stress. Detecting those regions and their propagation is of high importance for plant breeding and can speed up development of new varieties. Genotyping has insufficient resolution for these complex genomic regions, but the problem may be solved by optical mapping. In this project, we will use dual-color-labelling of key genomic features in melon plants to resolve resistance genes within specific trait-related genomic (QTL) regions.

Last modified: 20-08-2020

Copyright information:

Created using DMPonline. Last modified 20 August 2020

1 of 4

Resolution of resistance genes within melon QTL regions by optical mapping

1. General features of the project and data collection

Prof. dr. ir. D. de de Ridder (Projectleader and secretary)
T: +31317484074 | E: dick.deridder@wur.nl
Wageningen University
Departement of Plant Sciences
Bioinformatics
Postbus 1
6700 AP WAGENINGEN

- The expert is connected to my department or institution (please explain his/hr expertise related to data stewardship)

Dr. Sven Warris
T: +621877849 | E: sven.warris@wur.nl
Wageningen Research
Applied Bioinformatics
Postbus 1
6700 AP WAGENINGEN
Appointed data steward and bioinformatics researcher

- Generate new data
- Use existing data (please specify)

Existing data will comprise of the melon genome, related plants and their genome annotations.

- Exclusively quantitative data

- Yes, I have permission to use the data

Resources are publicly available.

- Yes, I will collect the new data in conjunction with other researchers or research groups
- Yes, the new data will be (partly) provided by a project partner or supplier

- Yes, clear arrangements have been made regarding data management and intellectual property through a consortium agreement

- Yes (please specify)

Around 100GB (2 Bionano runs and subsequent data analyses of about 10 plant samples).

- Documentation of the research process, including documentation of all participants
- Data documentation
- (Several versions of) processed data
- Software
- Raw data

Metadata for all plants used will be collected, documented and shared. The raw Bionano data will be made available through a public repository, linking to these metadata. The resulting genome mapping data will be described in a document and also be made publicly available.
OptiTools is already OSS.

- Yes, I will make use of my institution's standard facilities for storage and backup of my data

Both local infrastructures at the guest and hotel have large storage facilities, including redundant data storage. Data will also be delivered on external HDD.

2. Legislation (including privacy)

- No, I will not be doing research involving human subjects; proceed to section 3 (Making data findable)

Question not answered.

Question not answered.

Question not answered.

3. Making data findable

- Yes, it can be found through an online (metadata) catalogue or web portal (please specify)
- Yes, it can be found through the search engine of the archive or repository in which it is stored (please specify)

Data will be deposited on EBI ENA or NCBI SRA, depending on the data type. Large Bionano data will be made available through DANS EASY. All three databases create DOIs and require metadata.

- Yes, I will use a metadata scheme specific for my field of research (please specify)

We will use the metadata scheme used by either NCBI/SRA or EBI/ENA, whichever we will choose for submission:
https://www.ncbi.nlm.nih.gov/sra/docs/submitmeta/
https://ena-docs.readthedocs.io/en/latest/submit/analyses/bionano-maps.html

- Yes, I will be using the DOI code

4. Making data accessible

- Yes, immediately

- Yes, proceed to section 5 (Making data interoperable)

Question not answered.

Question not answered.

5. Making data interoperable

- Yes (please specify)

All supplied data files, such as Bionano data and genome assembly and annotations, will be made available in by industry accepted data format. These include BNX, fasta and GFF3 files as well as a tabular format for the metadata.

Created using DMPonline. Last modified 20 August 2020

3 of 4

- Yes, metadata standard (please specify)

Plant (phenotype) data will be classified according to the Plant Trait Ontology (https://www.ebi.ac.uk/ols/ontologies/to)

- No (please explain)

No human data are being used or generated.

6. Making data reusable

- I will document the software used in the course of the project (please specify)
- I will perform quality checks on the data to ensure that they are complete, correct and consistent (please explain)
- I will document the research process (please explain)

The data measurement process will be documented in the WUR E-Labjournal. Software documentation will include a description of the approach, regular well-described commits and version controlled. Further computer analysis will be detailed in a report/publication.
The Bionano data will go through a QC at the hotel.

- Yes

Intermediate data, eg. data which can be derived from the available raw data following the described approaches and which are not the end result, will be deleted.

- Yes (please specify)

It will be around 50 GB to 100 GB.

- Yes, and this archive has a data seal of approval (please specify the archive)

Main data will be stored in DANS EASY. However, the project (meta)data, genome data and annotations will be made available through EBI/NCBI. These are the industry standards and well-funded, but not certified.

- Yes, in accordance with VNSU guidelines (please specify the number of years)

Data and software will not be removed from public repositories.

Storage and compute costs are included in the costs of the project hours and sample delivery. DANS and EBI/NCBI are free of charge.

- Yes (please elaborate)

DANS and EBI/NCBI are free of charge.