Biochemical identification of virus susceptibility factors for crop improvement and sustainable cultivation

A Data Management Plan created using DMPonline.be

Creator: Roosje Van Ende

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: S000624N

ID: 206301

Start date: 10-01-2023

End date: 10-01-2027

Project abstract:

Because they have been implicated in the decline of bees, the EU decided to ban the use of neonicotinoid insecticides in 2018. A typical use of these products was the pre-treatment of field crop seeds to protect the emerging seedling from aphid feeding, which transmits viruses to the plant, with severe potential yield losses as a consequence. Climate change exacerbates this problem, because even small increases in temperature increase the number of aphids spreading the disease. As a consequence, substantial yield losses are predicted in the EU. Growing resistant varieties is the most sustainable way to avoid such yield losses. Here we propose to apply a novel biochemical approach to identify virus susceptibility factors. Susceptibility factors are those plant factors that are hijacked by the virus to infect, replicate and spread in the plant. By inactivating or modifying these factors, the plant can be brought outside of the host range of the virus. Because most plant viruses consist of an RNA genome, plant susceptibility factors often have the capacity to physically interact with RNA. Therefore, we propose to identify the RNA binding proteins that physically interact with the viral genome in vivo. Identifying such proteins based on their capability to interact with RNA has become possible through recent developments in the RNA-protein field and in the Geuten group these methods have been combined into new purification strategies and applied to plants. The identification of virus susceptibility genes is of general interest to sustainable agriculture.

Last modified: 11-04-2024

Created using DMPonline.be. Last modified 11 April 2024

1 of 7

Biochemical identification of virus susceptibility factors for crop improvement and sustainable cultivation
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Only for digital data Digital Data Type	Only for digital data Digital Data format	Only for digital data Digital data volume (MB/GB/TB)	Only for physical data Physical volume
N. benthamiana CRISPR lines genotyping data	Sequencing data from LGC	New data	Digital	Experimental and software (ICE synthego)	.ab1	<100MB	
N. benthamiana CRISPR lines phenotyping for viral resistance	Observation of mutiple phenotypic properties of TRV-infected S-gene KO lines	New data	Digital	Observational	.xlsx	<100MB	
N. benthamiana CRISPR lines experimental data	RT-qPCR of TRV-infected S-gene KO lines	New data	Digital	Experimental	.xlsx	<100MB	
Determine AA binding site S gene	Y3H and RIP-qPCR	New data	Digital and physical (cloning)	Experimental	.xlsx	<100MB	1 box in -80°C
N. benthamiana S-factor interactome	Proximity dependent labeling and MS data	New data	Digital	Software (Maxquant)	.raw	<100GB	
Confirm BYDV infections	RT-qPCR and DAS-ELISA	New data	Digital	Experimental	.xlsx	<100MB	
Quality control BYDV SAPS-capture	RT-qPCR, Silver stain, Nanodrop, Bioanalyzer, Gel elektrophoresis	New data	Digital	Experimental	.jpg .xlsx .txt	<100MB	
SAPS-capture BYDV	MS data RBPs interacting with the virus	New data	Digital	Software (Maxquant)	.raw	<100GB	
Target prioritisation in wheat/barley protoplasts	CRISPi or amiRNA	New data	Physical (cloning)	Experimental			1/2 box in -80°C
Wheat/barley CRISPR lines genotyping data	designing homeolog specific primer pairs and indexed pooled amplicon sequencing	New data	Digital	Experimental and software (ICE synthego)	.ab1	<100MB	

Wheat/barleyCRISPR lines phenotyping for viral resistance	Observation of mutiple phenotypic properties of BYDV-infected S-gene KO lines	New data	Digital	Observational	.xlsx	<100MB	
Wheat/barley CRISPR lines experimental data	RT-qPCR and DAS-ELISA of TRV-infected S-gene KO lines	New Data	Digital	Experimental	.xlsx	<100MB	
Quality control BYV SAPS-capture	RT-qPCR, Silver stain, Nanodrop, Bioanalyzer, Gel elektrophoresis	New data	Digital	Experimental	.jpg .xlsx .txt	<100MB	
SAPS-capture BYV	MS data RBPs interacting with the virus	New data	Digital	Software (Maxquant)	.raw	<100GB	
Target prioritisation in wheat/barley protoplasts	CRISPi or amiRNA	New data	Physical (cloning)	Experimental			1/2 box in -80°C
Sugar beet CRISPR lines genotyping data	SES Vanderhave	New Data	Digital	Experimental and software		<100MB	
Statistical analysis	Statistical analysis of RT-qPCR/DAS-ELISA/phenotypic quantifications etc.	New Data	Digital	Software (Graphpad prism)	.pzfx	<100MB	
Lab notes	Lab notes	New Data	Physical	Observational			1book/year

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Use data from previous project (phD thesis Roosje Van Ende) as a start for this project (list of potential S gene candidates in N.benthamiana).

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- Yes

During the timeframe of the project, we will work for Flanders based plant breeding companies (BASF and SES VanderHave) and through the established partnerships and know-how we will also aim for the set-up of a technology platform. The spin-off company would have a mixed business model, both creating on the one hand own IP through in house R&D activities and licensing it out to breeding companies and on the other hand demand driven research services and R&D partnerships with breeding companies to develop resistance traits upon request in their elite germ plasm. The application of the proof-of-concept for TRV in N. benthamiana to economically important viruses: BYDV and BYV will

generate IP (datasets: SAPS-capture BYDV and SAPS-capture BYV and validation of these results)

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

The data is sorted in folders named according to the type of data (eg. MS, RT-qPCR etc.) labeled with the date and short title explaining the experiment. Detailed information about the experimental set-up, methodology, observations, variable-level detail, results, conclusions etc. can be found in the lab notes (at the date the data is labeled with). Key experiments/data and protocols are also stored on 'teams' in the same manner, but with a more elaborate explanation of the experiment. Physical samples for long-time storage are stored in the -80°C freezer, an online database shared on 'Teams' provides information about these samples.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

For the MS data, PDBx/mmCIF will be used. This is the standard archive format used by the Protein Data Bank (PDB). It provides both metadata and data according to properties defined in the PDB Exchange Dictionary and the Macromolecular Crystallographic Information Framework (mmCIF).

3. Data storage & back-up during the research project

Where will the data be stored?

All data is stored on the cloud service of KULeuven, OneDrive for Business cloud system (2 TB of storage capacity). Key experiments/datasets/protocols are also stored/shared on the Lab's teams channel. Copies are kept on personal devices provided by KULeuven and running on LUNA software. The generated MS data will be deposited in a corresponding database, e.g. PRIDE. Physical samples for long-time storage are stored in the -80°C freezer, an online database shared on 'Teams' provides information about these samples.

How will the data be backed up?

The data will be stored at OneDrive for Business cloud system (2 TB of storage capacity), a service provided by the university, which automatically synchronizes the data everyday. A back-up of key datasets is available at 'teams'. For large datasets such as mass

spectrometry data (> 1 GB), the lab's hard drive could also be used as backup instead of cloud storage.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

OneDrive for Business cloud system (2 TB of storage capacity). All data within this project requires only limited storage capacity.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

All data stored in OneDrive for Business cloud service and Teams can only be accessed by a two-factor authentication system.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

OneDrive for Business is provided by KULeuven. The cost of Teams is covered by the lab.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data will be retained at least for the expected 5 year period after the end of the project.

Where will these data be archived (stored and curated for the long-term)?

For long-term storage, the data is uploaded the lab's hard drive. Shared data stored on 'Teams' during the project remains available to access after the project.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The costs are covered by the lab.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- No (closed access)

The project aims to generate intellectual property (IP) for licensing to breeding companies. While the protection of IP is prioritized, the initial objective of the project does not include immediate publication of results.

If access is restricted, please specify who will be able to access the data and under what conditions.

Apart from members within our research group, access to the data is only granted to external parties upon signing a non-disclosure agreement (NDA) or when licensing the intellectual property (IP).

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

The raw data of the mass spectrometry will be submitted to PRIDE database, which is necessary for patent application. Patents will be accessible in the patent database (eg. EPO,WIPO etc.).

When will the data be made available?

No research data will be made available unless through published patents.

Which data usage licenses are you going to provide? If none, please explain why.

The goal of the project is to generate IP, which can be licensed by breeding companies.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

Mass spectrometry data deposited to PRIDE is identifyable through its PRIDE dataset identifyer "PXDXXXXX"n, which will be available in the accompanying patent.

What are the expected costs for data sharing? How will these costs be covered?

Question not answered.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

All people involved in the project are responsible for their own (meta)data management.

Who will manage data storage and backup during the research project?

All people involved in the project are responsible for their own backup (meta)data management.

Who will manage data preservation and sharing?

Prof. dr. Koen Geuten is responsible for ensuring data preservation and reuse

Who will update and implement this DMP?

Prof. dr. Koen Geuten is responsible for updating and implementing this DMP.

Created using DMPonline.be. Last modified 11 April 2024

7 of 7