FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Dries Wéry - 0009-0008-0431-963X**
Contributor name(s) (+ ORCID) & roles	**PI - Jan Michiels (https://orcid.org/0000-0001-5829-0897)**
Project number [1] & title	A novel framework for identifying antibiotic targets in *Pseudomonas aeruginosa* - 1158325N
Funder(s) GrantID [2]	1158325N
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	Antibiotics arguably constitute the most important discovery in the history of medicine. In addition to revolutionizing the treatment of life-threatening bacterial infections, these compounds have facilitated a myriad of medical procedures including cancer therapies and organ transplants. Recently, however, healthcare systems worldwide are threatened by the accelerated emergence and spread of mechanisms conferring resistance towards all antibiotics in clinical use. Furthermore, the golden era of antibiotic discovery is long gone, and new antibiotics reach the market at an alarmingly slow pace. If no action is taken, a post-antibiotic crisis is eminent in which common infections kill and routine surgical procedures come with high risks. With antibiotic resistance already claiming 1.3 million lives annually and projections indicating a potential rise to 10 million by 2050, there is an urgent need to find new antimicrobial compounds. We here propose to identify hitherto unknown genetic determinants that are essential for survival of the bacterial pathogen *Pseudomonas aeruginosa* in lab conditions as well as during infection of a host. The proposed platform uniquely integrates advanced experimental approaches including genome-scale CRISPR editing with high-throughput screening in clinically relevant conditions. Pinpointing the crucial elements of survival determinants will in the future allow for the design of new antimicrobials with unprecedented mechanisms of action.

2. 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 [3].

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*	*ONLY FOR PHYSICAL DATA*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
WP1: pooled *P. aeruginosa* PAO1 CRISPRi libraries	Full library (ca. 5000 genes) covering the entire core genome of *P. aeruginosa* and selected accessory genes present in many *P. aeruginosa* isolates.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	At least 30 cryotubes (2 ml) are required for the CRISPRi screenings and will be stored at -80°C. Additional cryotubes (2 ml) will be stored at -80°C at a different physical location in the same building to serve as back-up.
WP1: individual sgRNA strains	Strains containing prioritized sgRNAs for validation of 50 and 4 promising targets *in vitro*	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	~ 50 strains, stored in a 96 well plate at -80°C

[3] Add rows for each dataset you want to describe.

	and *in vivo* respectively.			☐ Software ☐ Other:			
WP1: CRISPRi screening data	sgRNA counts after pooled growth of *P. aeruginosa* PAO1 CRISPRi libraries *in vitro* and *in vivo*.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	sam, bam, .ab1, .fasta/fa, .qual, gb/gbk, .dna	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
WP1: network interaction data	Protein-protein and metabolic interaction networks containing the prioritized genes identified in the CRISPRi screenings.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☒ Textual ☒ Model ☐ Software ☐ Other:	SIF, XGMML, GraphML, PSI-MI Level 1 and 2.5, .json, .csv, .xlsx, GAF, GPAD, GPI, .txt, .sql, .jpg, .png, .svg, .pdf	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
WP1: results from phenotypic analyses	Microscopy, flow cytometry, FACS and omics data resulting from phenotypic analyses probing for the function of the prioritized genes.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.tiff, .png, .jpeg, .nd2, .csv, .xit, .txt, .xlsx	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
WP2: pooled barcoded *P.*	Pooled library of mutants	☒ Generate new data	☐ Digital	☐ Audiovisual		☐ < 1 GB	At least 9 cryotubes (2 ml) are required

aeruginosa PAO1 MAGESTIC libraries	resulting from saturation editing of 10 prioritized genes.	☐ Reuse existing data	☒ Physical	☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	for the MAGESTIC screenings and will be stored at -80°C. Additional cryotubes (2 ml) will be stored at -80°C at a different physical location in the same building to serve as back-up.
WP2: site-specific *P. aeruginosa* PAO1 mutants	Mutants resulting from editing of 1 exogenous gene and up to 8 genomic loci and saturation editing of 1 endogenous gene.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Cryotubes (2 ml; glycerol stocks) will be stored at -80°C.
WP2: MAGESTIC screening data	Barcode counts after pooled growth of *P. aeruginosa* PAO1 MAGESTIC libraries *in vitro* and *in vivo*.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	sam, bam, .ab1, .fasta/fa, .qual, gb/gbk, .dna	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
WP2: lists of mutations	Lists of mutations	☒ Generate new data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images	.csv, .xlsx, .txt	☒ < 1 GB ☐ < 100 GB	

	binarily classified as “tolerated” vs. “not tolerated” by *P. aeruginosa* PAO1 during *in vitro* and/or *in vivo* growth.	☐ Reuse existing data		☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
All WP: primers	Primers will be ordered at IDT.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Vials stored at -20°C.
All WP: purified plasmids	Plasmids created during optimization of the CRISPRi and CRISPR editing systems.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Vials stored at -20°C.
All WP: strains other than those generated for genome-wide CRISPRi and	Strains created during optimization of the CRISPRi and CRISPR editing systems.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Cryotubes (2 ml; glycerol stocks) will be stored at -80°C.

gene-specific MAGESTIC screenings				☐ Software ☐ Other:			
All WP: sequences and sequencing data	Plasmid design and validation of plasmid construction and transformation/ electroporation. Validation of genome editing.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	sam, bam, .fasta/fa, .fastq, .qual, gb/gbk, .dna, .tsv	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
All WP: scripts	Scripts for sgRNA design, analysis of RNAseq data, …	☒ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☒ Model ☐ Software ☐ Other:	.r, .py	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
All WP: digital images	Gel scans (f.e. from PCRs), figures, graphs, illustrations, …	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.tiff, .jpeg,.svg, .pdf	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
All WP: manuscripts	Manuscripts	☒ Generate new data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images	.doc, .pdf	☒ < 1 GB ☐ < 100 GB	/

		☐ Reuse existing data		☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
All WP: research documentation	Lab protocols and notebooks	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.doc, .one, .onetoc2, .eln	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

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.	**Scripts:** We have *in-house* sgRNA design and RNAseq data analysis scripts that have not been published yet. For the analysis of MAGESTIC screening data, we will use and modify *in-house* scripts for deep mutational scanning of essential bacterial proteins (*Nat Commun* (2023) 14: 241.). **Protocols:** We have *in-house* optimized protocols for the generation of CRISPRi and MAGESTIC libraries from oligo arrays, but these are not published yet. Which protocols will be used in T1.5 (Larger genetic networks and underlying molecular mechanisms) depends largely on the output generated in T1.2, T1.3 and T1.4. Nonetheless, we can rely on *in-house* protocols for some assays (e.g. *Front Microbiol* (2017) 8: 1193.).

	The validity of our CRISPRi screenings will be verified based on transposon-based records of the essential *P. aeruginosa* genome (*Proc Natl Acad Sci U S A* (2019) 116: 10072-80.).
Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☒ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: WP1 and WP2 both comprise mice experiments. I already obtained the FELASA B certificate required for these experiments, but approval by the institutional Ethical Committee for Animal Testing will only be sought when the above-mentioned CRISPRi and CRISPR editing approaches are optimized for *in vivo* testing. Evidently, we will not proceed with the mice experiments until we receive approval from the Ethical Committee. There are no other ethical issues concerning the proposed work.
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☐ Yes (provide PRET G-number or EC S-number below) ☒ No Additional information: There are no privacy issues concerning the proposed work.
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 ☐ No If yes, please comment: Potential tech transfer will be discussed with the KU Leuven Research & Development – Tech Transfer Office. The newly developed tools (high-quality CRISPRi libraries and deep mutational scanning in *P. aeruginosa*) as well as lists of prioritized drug targets along with predictions on druggability of these targets have valorization potential.
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 to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain: Existing agreements between VIB and KU Leuven do not restrict publication of data.

[4] See Glossary Flemish Standard Data Management Plan

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 to what data they relate and which restrictions will be asserted.	☒ Yes ☐ No If yes, please explain: Ownership of the generated data belongs to KU Leuven; copyright of the data belongs to Jan Michiels and Sibylle Vonesch.

3. 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). *RDM guidance on documentation and metadata*.	**Biological material:** Cryotubes and multi-well plates will be labelled with a reference number that links to an entry in our Microsoft Access Database which is hosted on a central server and accessible to all people involved in the project. All relevant information on the specific libraries, mutants and other strains will be included in this database. This includes a library/mutant/strain identifier, a thorough description of how the libraries/mutants/strains were constructed and a link to whole genome sequence if applicable. **Experimental results:** Data will be generated following standardized protocols which are stored in a central OneNote notebook. Furthermore, an E-Notebook (ELN) will be used to register day-by-day activities. Raw data, history and context of experiments, protocols and analysed data will be uploaded to this E-Notebook and backed up in the cloud. After publication or upon submission of manuscripts for publication, all datasets described in the publication will be deposited in dedicated data repositories (see below). **Scripts:** All scripts for designing sgRNAs, analysis of RNAseq data, … will be properly annotated so that the code is understandable and can be used to re-generate the results. After publication or upon submission of manuscripts for publication, all scripts will be uploaded in dedicated data repositories (see below), including a readme file explaining what each script is exactly used for.
Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: Various data types come with their own metadata containing technical information about settings, machine types, pixel density, resolution, channels... Examples of these include .fastq NGS files containing standard metadata on sequencing technique, or .nd2 following the

metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	Nikon metadata standards. Throughout the project, these data files will be preserved with their original metadata. For .txt, .csv, .xlsx files containing tabular information, extra tabs or a head text section will be used to explain the data, the meaning of the columns etc. For others lacking a formally acknowledged metadata standard, Dublin Core Metadata will be used and a readme file will be saved in the same directory of the datafiles to explain all the various data files and give a broad overview of the analysis steps. Moreover, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards that are more specific to our data. After publication or upon submission of manuscripts for publication, all datasets described in the publication will be deposited in data repositories (see below). Depending on the repository that is used, the metadata standard used by that specific repository will be filled in. If no, please specify (where appropriate per dataset or data type) which metadata will be created: /

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☒ Shared network drive (J-drive) ☒ Personal network drive (I-drive) ☒ OneDrive (KU Leuven) ☒ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☐ Other:
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)

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 ☐ No If no, please specify: KU Leuven provides storage space on their internal server that is maintained by its IT service. The Michiels lab has two drives on this internal KU Leuven server: the J-drive (shared drive) for daily use and the K-drive for long-term data storage. When necessary, data storage capacity can and will be increased.
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	**Biological material:** -80°C freezers are present in the Michiels lab. A back-up of selected strains will be stored in -80°C freezers present at a different physical location (Kevin Verstrepen lab, Heverlee). Unauthorized people do not have access to the strains on neither one of these locations. **Experimental results:** Data are stored on secure university servers with built-in back-up and versioning. These serves are secured by two factor authorization and passwords that are frequently changed. Data collected and stored in the ELN are also protected via two factor authorization. Password-protected hard drives equipped with anti-virus programs will be used as back-up.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	**Biological material:** Costs associated with storing strains in -80°C freezers (at two physically different locations) are covered by general lab expenses. **Experimental results:** Costs associated with large volume storage of experimental data are covered by general lab expenses.

5. 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	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)

(e.g. legal or contractual restrictions, storage/budget issues, institutional policies...). *Guidance on data preservation*	
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☒ Large Volume Storage (long-term for large volumes) ☐ Shared network drive (J-drive) ☒ Other (specify): Data (biological and experimental) linked to published research will be deposited in public and free data repositories.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Costs for storage of research data in public data repositories (only for published research) and on the KU Leuven long-term, large-volume central servers is limited and covered by general lab expenses. Biological material can generally be deposited in public repositories without costs.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☒ Other, please specify: All published data will be made available at the time of publication. However, in case we identify valuable IP, we will first protect commercial exploitation, either through patenting or via an MTA that restricts the material from commercial use. This will be done after consulting with KU Leuven LRD. Unpublished, essential data will be available to (future) lab members via internal IT provisions.

If access is restricted, please specify who will be able to access the data and under what conditions.	All published data will be made available at the time of publication. However, if we identify valuable IP, we will first protect commercial exploitation through patenting or via an MTA that restricts the material from commercial use. This will be done after consulting with KU Leuven LRD. Unpublished, essential data will be available to (future) lab members via internal IT provisions.
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 per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ KU Leuven RDR ☐ Other data repository (specify) ☒ Other (specify) We aim to publish our research in renowned journals that require full disclosure upon publishing. Data will be made available in the main text, in the supplementary material or in a data repository if requested by the journal and following given deposit advice given by the journal.
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)

Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.	
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	**Biological material:** Shipment is generally paid by requesting parties. **Experimental data:** Deposit in an online data repository is free of charge. Publication costs depend on the publishing journal and will be covered by the project's budget.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Dries Wéry
Who will manage data storage and backup during the research project?	Dries Wéry
Who will manage data preservation and sharing?	Jan Michiels (PI) (Centre of Microbial and Plant Genetics, KU Leuven)
Who will update and implement this DMP?	Dries Wéry and Jan Michiels (PI) (Centre of Microbial and Plant Genetics, KU Leuven)