FWO DMP Template - Flemish Standard Data Management Plan

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	**Jolien Vreys** **https://orcid.org/0000-0001-5310-336X**
Contributor name(s) (+ ORCID) & roles	**Promotor: Patrick Van Dijck** **https://orcid.org/0000-0002-1542-897X**
Project number[1] & title	**3E211159** **Why is *Candida glabrata* specialized for solely glucose utilization with so many other sugars available in our body?**
Funder(s) GrantID[2]	**11L0423N**
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible: **https://ror.org/05f950310**

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. 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.

[3] Research Organization Registry Community. https://ror.org/

Please provide a short project description	The human health is threatened by the emergence of drug resistant fungal species, like *Candida glabrata*, indicating the need for new antifungal drugs and targets. The central metabolism appears to be tied to the pathogenicity of fungi. This also seems to be the case for *C. glabrata*'s sugar metabolism. Nevertheless, *C. glabrata* is an odd fungus concerning its sugar metabolism. It can only use glucose and trehalose as fermentable carbon sources despite the presence of gene orthologs for transport and metabolization of other sugars. Interestingly, a high glucose import rate was observed in *C. glabrata*, while there are only 11 hexose transporters encoded. Did *C. glabrata* become specialized for glucose transport upon its adaptation to the human host? To investigate this, we will determine the substrate specificity and kinetics of the hexose transporters and compare sugar transport of clinical and environmental isolates. Once sugar is imported, glycolysis takes place. We found five sugar kinases, catalyzing sugar phosphorylation, in *C. glabrata*. This is surprising as this is more than found in related fungi and they presumably only have glucose as substrate. To gain more knowledge about this, we will determine whether the sugar kinases are able to phosphorylate other sugars and if they have additional roles. Finally, as several virulence-related pathways require sugar phosphorylation, we will look into the effect of this phosphorylation on *in vitro*, *ex vivo* and *in vivo* virulence.

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[4].

				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
		☐ Generate new data ☐ Reuse existing data	☐ Digital ☐ Physical	☐ Observational ☐ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Growth curves	Multiskan	Generate new data	Digital	Experimental	.xlsx and .pzfx	<100 GB	
Fluorescence measurement and absorbance	Synergy H1	Generate new data	Digital	Experimental	.xlsx and .pzfx	<100 GB	

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

Flow cytometry	Guava	Generate new data	Digital	Experimental	.fcs and .pzfx	<100 GB	
Digital images	Microscopy images, gel scans, plate images, graphs, illustrations, figures	Generate new data	Digital	Experimental	.jpeg, .tif, .ai and .pdf	<100 GB	
Sequences	CLC	Generate new data	Digital	Experimental	.dna	<100 GB	
qPCR	Expression levels	Generate new data	Digital	Experimental	.xlsx and .pzfx	<100 GB	
Transport experiments	Scintillation counting	Generate new data	Digital	Experimental	.xlsx and .pzfx	<100 GB	
Strains	Deletion strains, fluorescence-tagged strains, HXT-null strain, HXT-complemented strains, environmental isolates, clinical isolates	Generate new data, reuse existing data	Physical				<500 strains
Plasmids	Deletion cassettes, tagging cassette, pLS10	Generate new data, reuse existing data	Physical				<100 plasmids

Guidance: *Data can be digital or physical (for example biobank, biological samples, …). Data type: Data are often grouped by type (observational, experimental etc.), format and/or collection/generation method.* *Examples of data types: observational (e.g. survey results, sensor readings, sensory observations); experimental (e.g. microscopy, spectroscopy, chromatograms, gene sequences); compiled/aggregated data[5] (e.g. text & data mining, derived variables, 3D modelling); simulation data (e.g. climate models); software, etc.* *Examples of data formats: tabular data (.por,. spss, structured text or mark-up file XML, .tab, .csv), textual data (.rtf, .xml, .txt), geospatial data (.dwg,. GML, ..), image data, audio data, video data, documentation & computational script.* *digital data volume: Please estimate the upper limit of the volume of the data per dataset or data type.* *physical volume: Please estimate the physical volume of the research materials (for example the number of relevant biological samples that need to be stored and preserved during the project and/or after).*	
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.	**Strains and plasmids generated by other research groups and other lab members will be used to generate deletion mutants, fluorescent-tagged strains and complementation strains.** • **pYC44 and pYC56 plasmid: *doi: 10.1016/j.fgb.2015.04.020.*** • **pLS10 plasmid: doi: 10.1080/21505594.2020.1868825** • ***hxt*-null mutant: doi: 10.1093/femsyr/foy107** **The environmental strains and clinical strains will be obtained from several other labs that give us permission to use them.**
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, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☒ Yes, animal data ☐ Yes, dual use ☐ No If yes, please describe: **For animal experiment, we will request approval from the ethical committee.**

[5] These data are generated by combining multiple existing datasets.

Will you process personal data[6]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:
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:
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:
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:

[6] See Glossary Flemish Standard Data Management Plan

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).	***All experiments are placed in folders (stored on the personal folder in the J-drive) stating the used protocol, the purpose and lay-out of the experiment, the raw data, analysis and conclusions. The standard operating protocols are collected in shared database on the J-drive.***
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 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.*	☐ Yes ☒ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: If no, please specify (where appropriate per dataset or data type) which metadata will be created: **All data captured by measurements of a physicochemical property in a batch mode will be manually curated to create meaningful metadata. The processing of the raw data is carried out in Graphpad prism and by creating graphs the data become meaningful to others.**

4. Data Storage & Back-up during the Research Project

Where will the data be stored?	***The data generated during this research will be preserved in several manners. First, the data on the applicant's computer are backed up daily, on an external hard drive. Secondly, the data is transferred regularly to storage repositories maintained by theKU Leuven (hard drive and Box system). After the research, all data is maintained at these storage repositories at KU Leuven. Our lab uses four different drives: a shared drive, a personal drive, a large volume storage drive and lastly a drive used to archive results and presentations.***
How will the data be backed up? *What storage and backup procedures will be in place to prevent data loss? Describe the locations, storage media and procedures that will be used for storing and backing up digital and non-digital data during research.*[7] *Refer to institution-specific policies regarding backup procedures when appropriate.*	**The central server of the KU Leuven has automatic back-up procedures.**
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 yes, please specify concisely: **The servers of the KU Leuven, where the data is stored, has no limit on data storage.** If no, please specify:

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

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.* [7]	***The KU Leuven server is a secure environment for data saving. The data is collected in folders only accessible for people working on this research. Moreover, the work laptop is protected by Windows defender and is managed by KU Leuven.***
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	***The cost of the J-drive is €519/TB/year and will be covered by the host lab.***

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 (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).	***All data will be stored in the servers from the KUL and on a hard drive.***
Where will these data be archived (stored and curated for the long-term)?	***The data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.***

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	***The costs are €113,84/TB/year and will be covered by the host lab.***

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, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☐ Other, please specify:
If access is restricted, please specify who will be able to access the data and under what conditions.	**Only people working on the project will have access to the folders containing the data. After publication, data be available upon request.**
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.	***All data will be saved in the servers from the KUL and be available from there.***

When will the data be made available? *This could be a specific date (dd/mm/yyyy) or an indication such as 'upon publication of research results'.*	***The data will be available upon request after publication.***
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.* *Example Answer: E.g. "Data from the project that can be shared will be made available under a Creative Commons Attribution license (CC-BY 4.0), so that users have to give credit to the original data creators."* [8]	**Currently, we will provide none. Data will be available on request. This might change depending on the results of the research.**
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 ☒ No If yes:
What are the expected costs for data sharing? How will these costs be covered?	***Since the data is shared upon request, there are no expected costs.***

[8] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	***Jolien Vreys will be the main responsible for data documentation & metadata. Prof. Patrick Van Dijck is co-responsible for the data storage and backup of the server.***
Who will manage data storage and backup during the research project?	***Jolien Vreys will be the main responsible and Prof. Patrick Van Dijck will be co-responsible for the data storage and backup of the server.***
Who will manage data preservation and sharing?	***Jolien Vreys will be the main responsible and Prof. Patrick Van Dijck will be co-responsible for the data preservation and sharing.***
Who will update and implement this DMP?	***Jolien Vreys and Prof. Patrick Van Dijck bear the overall responsibility for updating & implementing this DMP.***