Hermetia screening services
Thomas More 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 | Only for digital data | Only for digital data | Only for physical data |
				Digital data Type	Digital data format	Digital data volume (MB/GB/TB)	Physical volume
BSF -20 bank	Flies of different BSF strains, of each generation reared in IPP, stored in falcon tubes with ethanol at -20°C, for subsequent DNA analyses	reused	physical				2 m3
logbook BSF strains	info of rearing parameters linked to the flies in the -20 bank	reused	digital	Observational	.xls	<100MB	
DNA extract	extracted DNA of different BSF strains stored at -20	generate new data	physical				micro-tubes N =±200
sequencing data	WGS data (>10X) illumina 150bp PE& Pooled SEQ data	generate new data	digital	Observational	.fastq / .bam / .sam	< 1 Tb	
Publication	Publication on WGS diversity	generate new data	digital	Other	.pdf / .docx	<100MB	
BSF inbred strains	we will maintian the strains created from the inbreeding experiment	generate new data	physical				12 m3
BSF -20 benk inbred strains	we will store the flies from the inbreeding experiment in the BSF -20 bank	generate new data	physical				2 m3
logbook BSF strains	info of rearing parameters linked to the flies from the inbreeding experiment in the -20 bank	generate new data	digital	Observational	.xls	<100MB	

phenotyping data inbreeding experiment	inbreeding experiment: In generation 1, 5 and 10, larvae will be phenotyped for their performance, such as egg laying and hatching rate, pupa hatching rate, survival rate, growth rate, larval mass and feed conversion.	generate new data	digital	Observational/experimental	.xls	<100MB	
data long-term perservation experiments	while optimizing long term perservation all paramters of conditions (temperature, humidity, substrate WHC, composition..) and larval performance will be monitored. Besides survival after 1, 2, 3, 4, 5 & 6 months of storage, performance will be investigated as well. This includes feed conversion, growth rate, larval mass and survival, pupation and adult hatching, egg-laying and egg-hatching.	generate new data	digital	Experimental	.xls	<100MB	
DNA extract	extracted DNA of different generations in WP2 stored at -20	generate new data	physical				micro-tubes N =±100
sequencing data	WGS data (>10X) illumina 150bp PE	generate new data	digital	Observational	.fastq / .bam / .sam	< 1 Tb	
Publication	Publication on inbreeding, genomic characterisation, population size	generate new data	digital	Other	.pdf / .docx	<100MB	
Protocol on strain management practices	Protocol on strain management practices	generate new data	digital		.pdf / .docx	<100MB	
Protocol for long term storage of BSF larvae	Protocol for long term storage of BSF larvae	generate new data	digital		.pdf / .docx	<100MB	

Data diet-dependent performance of BSF strains	Fitness-relevant and economically desired performance traits, such as average larval (dry) mass, larval survival rate, bioconversion efficiency and waste reduction will be monitored during the feed trial on 3 different diets with different strains	generate new data	digital	Experimental	.xls	<100MB	
Data diet-dependent chemical properties of BSF strains	The chemical composition (fat & protein content) of the larvae on each diet will be analyzed for the different strains	generate new data	digital	Experimental	.xls	<100MB	
DNA extract	extracted DNA of different generations in WP2 stored at -20	generate new data	physical				micro-tubes N =±200
sequencing data	WGS data (>10X) illumina 150bp PE	generate new data	digital	Observational	.fastq / .bam / .sam	< 1 Tb	
Publication	Publication on microbiome diversity in relation to BSF genetic diversity and larval performance	generate new data	digital	Other	.pdf / .docx	<100MB	
market analysis	market analysis	generate new data	digital	Compiled/aggregated data	.pdf / .docx	<100MB	
strain passports	strain passports	generate new data	digital	Compiled/aggregated data	? Tbd	< 1 Tb	

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.

BSF -20 bank and logbook is established and maintained by TM-EC SBC

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

Passport Services, BSF Stock Offering, Strain Preservation, and Consultancy Services
within the project arrangements will be made on IP together with KUL LRD and TM legal office

Do existing third 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.

- Yes

Potentially, strains for which restrictions in use due to Nagoya Protocol

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.

- Yes

Potentially, strains for which restrictions in use due to Nagoya Protocol

2. Documentation and Metadata

Describe the documentation that will be created for the data. 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).

Data files within our research framework will be thoughtfully paired with information on the project, the experimental setup and methodology used.
Where possible a readme.txt file will be established (eg. logbook BSF strains, sequencing data) using the full meta-data model (https://www.kuleuven.be/rdm/en/rdr/full-metadata-model) , README files will be created following https://www.kuleuven.be/rdm/en/guidance/documentation-metadata/README, and https://libraries.ou.edu/content/how-make-readmetxt-file and https://data.research.cornell.edu/data-management/sharing/readme/.

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
- No

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

Where will the data be stored during research? Explain in the comment section.

- In a cloud service offered by Thomas More (Microsoft 365: SharePoint)
- Centrally on storage facilities of the research unit

TM -EC SBC: digital data: TM-sharepoint, physical data: insect pilot plant

KUL-IP&P: digital data: KUL-sharepoint for the specific research project, physical data: insect pilot plant and research labs, P-building, Campus Geel

KUL-CABG: extracted DNA will be stored in the freezers (-20°C) of the CABG Bio Sample Database, if possible in different aliquots, stored over different freezers on different electric circuits. All data generated in the project will be stored on KU Leuven internal servers (J: (Shared) , K: (Archive) and L: (Large Volume Storage) drives).

How will the data be backed up?

TM-EC SBC:Data will be backed up using cloud service offered by Thomas More (Microsoft 365: SharePoint/OneDrive/Teams). These cloud services take automatic backups.

KUL-CABG: Standard back-up provided by KU Leuven ICTS. DNA samples: if possible, all samples will be aliquoted (2 or more aliquots) and divided over different freezers, all with temperature alarm signals.

KUL-IP&P: Data will be backed up using cloud service offered by KU Leuven ICTS (Microsoft 365: SharePoint/OneDrive/Teams). These cloud services take automatic backups.

Is there currently sufficient storage and 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

TM-EC SBC: Cloud service offered by Thomas More has 5 TB storage and backup capacity through the corporate Office 365 licence. This is sufficient for the data created during the project

KUL-CABG: In the project budget, a part of the consumables is allocated to computing and data storage. The estimated volume we expect to generate in this project is within lines of the available space on the KU Leuven storage servers. If insufficient, additional server space will be acquired (funds from research group).

KUL-IP&P: Cloud service offered by KU Leuven has 1 TB storage and backup capacity per team through the corporate Office 365 licence. This is sufficient for the data created during the project

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

TM EC SBC: Access to the data stored in the cloud service offered by Thomas More (corporate Office365 storage) is arranged via the Azure Active Directory and Multi Factor Authentication

KUL-CABG: Access to the digital data is restricted to the team members of the research group CABG. DNA samples are stored in a locked (freezer) room with restricted access only to authorized members of the research group CABG.

KUL-IP&P: Electronic data will be stored using a combination of Sharepoint and Teams-sites. Both of these methods are recommended as suitable for confidential data by the KU Leuven storage guide (https://icts.kuleuven.be/storagewijzer/en). Teams allows granting of individual access to users
both inside and outside of KU Leuven, thus facilitating easy collaboration between the research consortium while maintaining security of data access.

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

TM-EC SBC: No extra costs are to be expected since the cloud service offered by Thomas More will be used.

KUL-CABG:Server storage: for LVS € 95.14 / TB / year. Expected total (1TB): 100€/year. This was foreseen in the project budget.

KUL-IP&P: No extra costs are to be expected since the cloud service offered by KU Leuven will be used.

4. Data preservation after the end of the research project

Which data will be retained for at least ten 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 generated (raw and processed) data will be retained for 10 years after the end of the project for the purposes of reproducibility, verification, and potential reuse.

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

TM EC SBC: physical data: -20 °C IPP, digital data: The data will be stored on SharePoint (with automatic back-up procedures) for at least 10 years, conform the Thomas More RDM policy.
KUL-CABG: Digital data will be preserved for 10 years according to KU Leuven RDM policy. KU Leuven K and LVS drives. DNA samples will be stored at -20°C.
KUL-IP&P:physical data: -20 °C Research Lab IP&P, P-building, Geel Campus. Digital data will be preserved for 10 years according to KU Leuven RDM policy.

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

TM EC SBC: No extra costs are to be expected since the cloud service offered by Thomas More will be used.
KUL-CABG: Costs are expected to be in range with 50-150 € year. In the project budget, a part of the consumables is allocated to data storage. If insufficient for the long-term storage, additional costs will be covered by the CABG research group.
KUL-IP&P: Sharepoint and Teams sites are free within KU Leuven and can also be used by other members of the research consortium. Space and costs associated with storage of physical samples are envisioned within the normal operations of the labs involved.

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.

- Yes, in a restricted access repository (after approval, institutional access only, ...)

If you answered 'Yes, in a restricted access repository' or 'Yes, upon request by mail', please specify who will be able to access the data and under what conditions. Explain the procedures to control access.

Potential application for patents limits the availability of the data until the patent is granted.
If no patents are filed, the published data (raw and processed) will be available to other researched on request and under certain conditions. The project collaborators (TM-Centre of Expertise sustainable biomass and chemistry, KUL-IP&P and KUL-CABG) will evaluate each 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 third party, legal restrictions)? Please explain per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

The project holds the potential of generating new processes, information with industrial relevance. Hence, the project collaborators may submit patent applications making the data private and not fit for public access (even on request).

If the data (or part of the data) will be made available for reuse after/during the project, explain where the data will be made available. If already known, provide a repository per dataset or data type.

Data will be available on request

When will the data be made available?

- Upon publication of research results

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

Data from the project that can be shared will be made available under a Creative Commons Attribution-NonCommercial license (CC-BY-NC), a creative commons license that bans commercial use

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.

- No

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

There are no expected costs for data sharing. Data sharing using Zenodo / OSF is free of charge.

6. Responsibilities

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

Project contributors will manage data documentation and metadata during the research project. Project leader bears the end responsibility.

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

Project contributors will manage data storage and backup during the research project. Project leader bears the end responsibility.

Who will manage data preservation and sharing?

Project contributors will manage data preservation and sharing during the research project. Project leader bears the end responsibility.

Who will update and implement this DMP?

The project leader bears the end responsibility of updating and implementing this DMP. The DMP will be updated during and at the end of the project.

Which additional resources are needed for the execution of the Data Management Plan?

No additional resource needed.

Did you read the Research Data Management Vision text of Thomas More?

- Yes

7. Internal project registration

Please provide the internal Thomas More project number.

C3 24 088

Did you read the Research Data Management Vision text of Thomas More?

GDPR

Have you registered personal data processing activities for this project?

- No