Improving animal welfare by rescuing male chicks from culling with a next-generation test for genespecific in ovo sexing

A Data Management Plan created using DMPonline.be

Creators: Simão Monteiro Belo dos Santos ⓘ https://orcid.org/0000-0002-9505-3943, Dragana Spasic

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Simão Monteiro Belo dos Santos ⓘ https://orcid.org/0000-0002-9505-3943

Project Administrator: Dragana Spasic

Grant number / URL: 1S54823N

ID: 197335

Start date: 01-11-2022

End date: 30-10-2026

Project abstract:

The global population increase next to the galloping development of countries has raised pressure for a higher quantity of safe(r) and cheap(er) food. The poultry industry remains one of the primary food sources, where high production stress resulted in new ethical issues over the years. As such, >6 billion male chicks worldwide are killed every year right after hatching because they have no use. The consumers consider this act barbaric, while the industry finds it nonlucrative. Therefore, a solution is urgently needed to improve industrial practices and please consumers. The most socially acceptable solution in this context is sex identification before hatching (so-called in ovo sexing), which is not yet present, fulfilling all the required criteria. To solve this issue, this project intends to determine and validate a female-specific gene for the embryos' sex determination in an accessible egg fluid (allantoic fluid), using standard and isothermal amplification methods (WP1), develop novel microfluidic solutions for on-chip DNA isothermal amplification, including sample manipulation (WP2), and design an autonomous microfluidic chip compatible with the eggs, for sample extraction, reagents conservation and DNA amplification (WP3). Finally, the knowledge will be combined with a chip touchless activation system to obtain a completely autonomous device that can work under incubation conditions without disturbing the embryos (WP4).

Last modified: 10-03-2023

Improving animal welfare by rescuing male chicks from culling with a next-generation test for genespecific in ovo sexing
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
Biological samples	Samples extracted from eggs and embryos	N	P	Sample	Tubes of liquid	NA	500-1000 mL
gDNA	DNA extracted from embryos samples	N	P	DNA	Tubes of liquid	NA	100-500
tDNA	Template DNA of the sequences of interest	N	P	DNA	Tubes of liquid	NA	10-100
Primers	Used primers during the research for the several stages and applications	N	P	DNA	Tubes of liquid	NA	10-100
Probes	Used probes during the research for the several stages and applications	N	P	DNA	Tubes of liquid	NA	10-100
raw data and associated statistical analysis	Numerical data. Counting of flies, recording of count data, analysis of data, saved in Microsoft excel and graph pad prism	N	D	N	.excel, .pzfx	50-200MB	
DNA sequencing files	Sequencing of PCR products	N	D	N	.txt	10-100MB	
Figures of data for publication	Figures of amalgamated data produced during the the project was created using excel, graph pad prism, and Inkscape.	N	D	I	.tif	10-100GB	
Lab books and E-lab notebook	Dated written notes associated with carrying out experimental procedures	N	P/D	T	Notebooks	2GB	5 books
qPCR data analysed	Statistical analysis lists of identified gene targets in qPCR experiments	N	D	N	.excel	50-200GB	
RPA lateral flow images	Captured images from the lateral flow assay and other steps of the RPA	N	D	I	.tif	10-100GB	
Microfluidic videos	Videos captured using Logitech camera and software for analysis of microfluidic behaviour	N	D	A	.mp4	100-500GB	
Microfluidics raw data	Data gathered from pressure and volume measurements	N	D	N	.csv	50-200MB	
Microfluidic prototype	Physical chips built for prototyping effects	N	P	Physical	Microfluidic chips		50 microfluidic chips
Microfluidic chip design	Stamp designs prepared in inkscape and/or other CAD software	N	D	A	.svg, .tif	5-10GB	
Risk assessments	Written risk assessments associated with standard operating procedures for experimental procedures performed within the lab	N	D	T	.docx, .pdf	50-200MB	
Standard operating procedures	Written protocols for experimental procedures performed in the lab	N	D	T	.docx, .pdf	50-200MB	
Text manuscript for publication	Text files associated with submitted publications	N	D	T	.docx, .pdf	1-20GB	

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

The digital data will be stored on the university's central servers (with automatic backup procedures) dedicated for long-term data storage (K: drive) for at least 5 years, conform the KU Leuven RDM policy.

The physical samples will be stored in the labs of the research group for up to 5 years after the project.

The accompanying metadata will be stored in the electronic lab notebook (eLABJournal, Bio- ITech). At the end of the project, all (meta)data on the electronic lab notebook will be exported as .pdf files and stored together with all other data onto the K: drive for archival purposes.

The cost of the ample volume storage will be € 128,39 per TB and year. We anticipate needing 1 TB for five years to keep the essential data available. This will amount to € 614,95 and will be covered by the research group.

Prof Jeroen Lammertyn (PI of the Biosensors group and promotor of this FWO-SB project), will enforce the preservation of the data.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

No deviation from the principle will take place.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Not applicable.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

Not applicable.

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

Created using DMPonline.be. Last modified 10 March 2023

4 of 9

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Created using DMPonline.be. Last modified 10 March 2023

5 of 9

Improving animal welfare by rescuing male chicks from culling with a next-generation test for genespecific in ovo sexing
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 / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
Biological samples	Samples extracted from eggs and embryos	N	P	Sample	Tubes of liquid	NA	500-1000 mL
gDNA	DNA extracted from embryos samples	N	P	DNA	Tubes of liquid	NA	100-500
tDNA	Template DNA of the sequences of interest	N	P	DNA	Tubes of liquid	NA	10-100
Primers	Used primers during the research for the several stages and applications	N	P	DNA	Tubes of liquid	NA	10-100
Probes	Used probes during the research for the several stages and applications	N	P	DNA	Tubes of liquid	NA	10-100
raw data and associated statistical analysis	Numerical data. Counting of flies, recording of count data, analysis of data, saved in Microsoft excel and graph pad prism	N	D	N	.excel, .pzfx	50-200MB	
DNA sequencing files	Sequencing of PCR products	N	D	N	.txt	10-100MB	
Figures of data for publication	Figures of amalgamated data produced during the the project was created using excel, graph pad prism, and Inkscape.	N	D	I	.tif	10-100GB	
Lab books and E-lab notebook	Dated written notes associated with carrying out experimental procedures	N	P/D	T	Notebooks	2GB	5 books
qPCR data analysed	Statistical analysis lists of identified gene targets in qPCR experiments	N	D	N	.excel	50-200GB	
RPA lateral flow images	Captured images from the lateral flow assay and other steps of the RPA	N	D	I	.tif	10-100GB	
Microfluidic videos	Videos captured using Logitech camera and software for analysis of microfluidic behaviour	N	D	A	.mp4	100-500GB	
Microfluidics raw data	Data gathered from pressure and volume measurements	N	D	N	.csv	50-200MB	
Microfluidic prototype	Physical chips built for prototyping effects	N	P	Physical	Microfluidic chips		50 microfluidic chips
Microfluidic chip design	Stamp designs prepared in inkscape and/or other CAD software	N	D	A	.svg, .tif	5-10GB	
Risk assessments	Written risk assessments associated with standard operating procedures for experimental procedures performed within the lab	N	D	T	.docx, .pdf	50-200MB	
Standard operating procedures	Written protocols for experimental procedures performed in the lab	N	D	T	.docx, .pdf	50-200MB	
Text manuscript for publication	Text files associated with submitted publications	N	D	T	.docx, .pdf	1-20GB	

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:

NA

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

This study holds great potential for tech transfer and valorisation. The submission of patents will be evaluated in collaboration with KU Leuven Research & Development (LRD) and the IOF manager of the Biosensors group. In these discussions, we will determine which data should be protected and in which way (e.g. patent, no disclosure). After safeguarding intellectual property, the inventions can still be published in open-access journals or presented at international conferences.

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

There are no 3rd party agreements in place regarding this project.

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 Biosensors group uses the electronic lab notebook in which a number of predetermined topics have to be described for each experiment (objective, protocol, results, and conclusion). The electronic lab notebook facilitates searching for particular metadata through a search engine. By mimicking the folder structure of the electronic lab notebook in the server-based folder with the experimental data, linking of the metadata to the actual data will be facilitated.
As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org). When depositing data in a local or public repository, the final dataset will be accompanied by this information in a README.txt document, following the Dublin Core Metadata standard if no other meta-standard is available yet. This file will be located in the top-level directory of the dataset and will also list the contents of the other files and outline the file-naming convention used. This will allow the data to be understood by other members of the laboratory and add contextual value to the dataset for future reuse.
For each peer reviewed article, a separate folder will be made on the server, containing the latest Word version and all raw and processed data used in the article. In addition, a separate file will be made in the electronic lab notebook for each article, containing clickable links to all metadata files of data that were used in that particular article, to facilitate tracing back of protocols, results and conclusions.

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.

- No

To guarantee reusable aspect of data, sufficient documentation and methods information will be provided, whereas CC-BY license will be attached to data through data repositories.

For more details, please see section 2.1.

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

Where will the data be stored?

- OneDrive (KU Leuven)
- Shared network drive (J-drive)
- Personal network drive (I-drive)
- Other (specify below)

The time-stamped digital data will be stored in a dedicated folder for this FWO-SB project on the shared drive (J:) of KU Leuven. The corresponding time-stamped digital metadata will be stored on the server of the electronic lab notebook (eLABJournal, Bio-ITech). The dedicated folder will be open for all research group members and is secured and backed up by the ICTS service of KU Leuven. Remote access to all data will be possible through KU Leuven OneDrive, which will also be used for the active use of the data during the project. Physical data samples will be stored in the freezers of the labs of the Biosensors group and will be documented in a digital inventory in the electronic lab notebook. At the end of the project, all information gathered on the electronic lab notebook will be exported as .pdf files and stored on the shared drive for long-term storage (K:), as to ensure all metadata is accessible even if the electronic lab notebook is not, and thus to minimize the risk of potential data losses.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

Data will be stored on OneDrive, which is synced approximately every 10 minutes. Backup to external hard drives will be performed every week. KU Leuven servers have hourly on-site backup and mirroring.

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

KU Leuven provides sufficient storage and back-up capacity during and after the project. A dedicated folder is made for the project to store data files.

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

The network drive for the project shared folder and the large volume storage folder are secured by the ICTS service of KU Leuven with a mirror copy. Confidential data can and will be protected with a password (available only for PI Jeroen Lammertyn). Visitors, MSc thesis students and internship students in the groups as well as other unauthorized persons will not have access to the data on the shared folder. Data storage in the cloud will be avoided, unless for temporary use only, e.g., to transfer large files between the researchers involved in the project.

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

Type 1 server back-end storage with mirror backup for the project shared folder will cost 57 Euro per Tb per year. Costs will be covered by the project consumables budget.

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...).

The data to be retained for the expected 5 years after the project's end are: dissemination data (source files of publications and presentations) and the most relevant measurement data.

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

The research data will be stored on an external hard drive after the end of the project. Dissemination data, namely files corresponding to papers and presentations, will be stored on the PCs of PI (J Lammertyn), and backed-up daily on the departmental server for long term storage.

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

The volume corresponding to dissemination data is expected to be relatively low (<10 GB), and therefore can be seamlessly embedded in the PIs' allocation on the departmental server. The costs (1000 EUR/year) will be covered by other on-going projects at that point in time.
For research data, at current archiving costs of 10 Euro/(TB*year), we estimate a cost of 2000 Euro/year. These costs will be covered by funding acquired by the project PIs in the context of other research projects.

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 an Open Access repository

Only researchers participating in the project will be able to access the data for the duration of the project. As soon as the article associated with the data is ready for publication, the data will be made open through the institutional repositories mentioned in 2.1. The data will be deposited in the institutional repositories: (KU Leuven: Research Data Repository (RDR) Research Data Repository (RDR) - RDR - Research Data Repository (kuleuven.be). Data will be assigned with DOIs to create trustworthy and persistent links for online content.

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

NA

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.

- Yes, Intellectual Property Rights

Before making data and other research output from the project (e.g. journal articles, book chapters and conference proceedings) openly available, they will be aligned with the project IP strategy to avoid premature disclosure, which can compromise the patent filing application(s).

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

KU Leuven RDR (Research Data Repository).

When will the data be made available?

As soon as the research results have been published, the data can be made available to other researchers.

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 license (cc-by 4.0), so that users have to give credit to the original data creators.

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

Data will be assigned with DOIs to create trustworthy and persistent links for online content.
ORCID will be included in all publications and other dissemination material for all the researchers involved in the project.

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

A restricted access repository can be implemented on a free tool, such as Dropbox, up to a certain volume. If this volume does not suffice, time-limited storage will be considered, thus limited to the time needed to download the data. The costs associated with data storage will be covered by the budget foreseen in the project agreement.

6. Responsibilities

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

Simão Santos (FWO-SB PhD student)

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

Simão Santos (FWO-SB PhD student)

Who will manage data preservation and sharing?

During the project: Simão Santos (FWO-SB PhD student). After the project: Jeroen Lammertyn (PI of the Biosensors group and promotor of this FWO-SB project).

Who will update and implement this DMP?

During the project: Simão Santos (FWO-SB PhD student). After the project: Jeroen Lammertyn (PI of the Biosensors group and promotor of this FWO-SB project).