Unravelling the mechanism(s) responsible for the therapeutic effect of histone deacetylase 6 (HDAC6) in amyotrophic lateral sclerosis (ALS)

A Data Management Plan created using DMPonline.be

Creator: Dufie Strubbe

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / **URL:** 11D4523N

ID: 193661

Start date: 01-11-2022

End date: 30-09-2026

Project abstract:

Amyotrophic lateral sclerosis (ALS) is a progressive, fatal neurodegenerative disease characterised by the degeneration of upper and lower motor neurons (MNs) followed by complete paralysis and death of the patient usually within 2 to 5 years after diagnosis. The devastating lack of effective treatments underlines the urgent need for further research towards the development of new therapeutic interventions. The host lab discovered that histone deacetylase 6 (HDAC6) could be a promising therapeutic target as its inhibition ameliorates ALS-related phenotypes in various in vitro and in vivo disease models. HDAC6 plays a unique role in many biological processes linked to ALS. However, little is known about the mechanisms underlying the observed beneficial effects of HDAC6 modulation. Therefore, the aim of this proposal is to elucidate the exact role of HDAC6 in ALS and to deliver mechanistic insights into the effect of HDAC6 modulation using FUS-ALS patient-derived induced pluripotent stem cell (iPSC) MNs alongside in vivo zebrafish models. More specifically, I will investigate the fundamental interactions and functions by which HDAC6 affects ALS pathogenesis in addition to unravelling its role in axonal transport, neuromuscular junction formation and stability, and ALS-associated proteotoxicity. Identifying and understanding the mechanism(s) of action underlying HDAC6 modulation will profoundly contribute to the development of a new therapeutic strategy for ALS.

Last modified: 26-04-2023

Unravelling the mechanism(s) responsible for the therapeutic effect of histone deacetylase 6 (HDAC6) in amyotrophic lateral sclerosis (ALS)
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.

	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
Plotted graphs	Graphs of data from excel files produced using GraphPad prism	☒ 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: .pzfx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Western blot images	Images created by developing western blot membranes containing protein samples.	☒ 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: .gel, .tiff, .JPEG ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Video files from behavioural assays	Video files recording the swimming behaviour of 48 hour old zebrafish embryos after touch (Touch-evoked escape response assay).	☒ 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: .avi ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Mass spectrometry raw data	Statistical analysis lists of identified protein targets in mass spectrometry experiments	☒ 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: .raw, .mzID ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Mass spectrometry analysed data	Data readable as LC chromatograms, massspectra or identified protein lists produced using the Max Quant programme	☒ 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: .xls ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

PCR gel images	Images obtained after exposing agarose gels containing DNA samples with UV -light	☒ 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: .tiff, .jpeg ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Microscopy slides	Microscopy slides used during imaging, consisting of formaldehyde-fixed tissue or cells immunostained or dyed using chemicals.	☒ 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	< 1000 glass microscopy slides
Immunohistochemistry images	Images of whole-mount zebrafish embryos or fixed cells using immunohistochemistry taken by confocal microscopes.	☒ 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: .tiff, .zen, .nd2 ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Fluorescent imaging files	Images of ventral root motor neurons in live zebrafish embryos.	☒ 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: .tiff, .JPEG, .nd2 ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
DNA sequencing files	Sequencing of plasmids, PCR products… performed by LGC genomics.	☒ 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: .fasta, .ab1 ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Analysis of confocal data	Quantification of confocal image data performed using ImageJ, in Microsoft Excel and GraphPad prism.	☒ 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: .xls, .pzfx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

Analysis of Western blot images	Numerical data. Quantification of western blot images, performed using ImageJ, in Microsoft Excel and GraphPad prism.	☒ 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: .xls, .pzfx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Protein samples	Samples of denatured or non-denatured proteins extracted from tissue or cells using detergents. Stored at -20° for short term storage or -80°C for long-term storage	☒ 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	500-1000 liquid-containing tubes
cDNA samples	DNA samples produced from extracted RNA from tissue or cells using cDNA transcription kits. Stored at -20°C for short term storage or -80°C for long term storage.	☒ 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	500-1000 liquid-containing tubes
RNA samples	RNA samples extracted from tissue or cells using RNA extraction kits. Stored at -80°C	☒ 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	500-1000 liquid-containing tubes
DNA plasmids	DNA plasmids produced during the product, derived from existing DNA plasmids provided by commercial and non-commercial suppliers. Stored at -20°C	☒ 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	50 liquid-containing tubes
Risk assessments	Written risk assessments associated with standard operating procedures for experimental procedures performed within the lab.	☒ 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: .docx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

Standard operating procedures	Written protocols for experimental procedures performed in the lab	☒ 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: .docx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Text manuscript for publications	Text files associated with submitted publications.	☒ 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: .docx ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Lab books	Electronic lab notebook and paper notes	☒ 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	Paper notes associated with performed experiments
Video files for axonal transport tracking	Live axonal transport videos + images from iPSC-derived motor neurons as well as zebrafish embryos. Analysed by ImageJ and quantified by using kymographs	☒ 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: .avi, .TIFF, .raw ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Bacterial strains generated	Bacterial strains conserved as glycerol stocks frozen at 80°C	☒ 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	Cryovials containing glycerol stocks of bacterial strains frozen at -80°C
Biological tissue samples	Eppendorf tubes containing: - 30-100 anaesthetised zebrafish embryos frozen at -80°C - Cell pellets from iPSC lines or other cell lines frozen at -80°C	☒ 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	500-1000 tubes

Transgenic zebrafish lines	Information on date of creation, crossing information, way of transgenesis, and name given to lines will be stored in lab books and stored in a list on the server.	☒ 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	
Cell lines	Cell lines derived from ALS patients' fibroblasts as well as commercially available human cell lines or relevant mouse cell lines	☐ 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	< 15 different cell lines
Nucleic acid sequences	Used for primer design, antisense (morpholino) oligonucleotides, generation of transgenic zebrafish lines	☒ 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: .ab1, .fasta ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

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:

Transgenic zebrafish lines: through MTAs from EZRC or upon request from other laboratories.
DNA plasmids: through MTAs from Addgene or upon request from other laboratories.
Risk assessments and SOPs already present in the lab.

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.

- Yes, human subject data

The ethical approval to use human derived cells in the context of our research was already obtained before the start of this project.
The use of patient fibroblasts for the generation of hiPSCs was approved by the ethics committee of University Hospital Leuven (n° S50354 and S63792), while the use of mmyoblasts was approved by the ethical commission (n° NH019-2020–04-02). Upon collection of primary tissue from human donors, donors were fully informed using an informed consent that their cell material may be usedfor the production of iPSCs and that these iPSCs may be shared internationally to support research.

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

We do not exclude that the proposed work could result in research data with potential for tech transfer and valorisation. VIB has a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed.
The use of material obtained within this project will be subjected to the terms described in their respective MTAs.

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

No third-party agreement restricts dissemination or exploitation of the data or strains generated from this project. In particular, existing agreements between VIB and KU Leuven do not restrict publication of data.

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).

Data will be generated following standardised protocols. Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the electronic laboratory notebook (E-notebook) and/or in hard copy lab notebooks that refer to specific datasets.
Cryotubes of biological samples (bacterial and yeast strains) stored at -80°C will be labelled with a reference number that links to an entry in or strain database.
All datasets will be accompanied by a README.txt file containing all the associated metadata (see more details below).
The data will be generated following standardised protocols. Clear and detailed descriptions of these protocols will be stored in our lab protocol database, and published along with the results.

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

Where available accepted metadata standards will be employed. See table below:

Type of Data	Metadata standard
Confocal images cells	OME-TIFF
Confocal images zebrafish	ND2, TIFF
Immunohistochemistry images	OME-TIFF
PCR gel images	OME-TIFF, JPEG
Western blot images	OME-TIFF, JPEG
Mass spectrometry raw data	mzML
...	

While specific data types might require particular metadata, as a general rule the metadata will be based on a generalized metadata schema such as Dublin Core or DataCite. The following information will be stored:

- Title: free text
- Creator: Last name, first name, organization
- Date and time reference
- Subject: Choice of keywords and classifications
- Description: Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) (including version number) used to produce and to read the data, the purpose of the experiment, etc.
- Format: Details of the file format,
- Resource Type: data set, image, audio, etc.
- Identifier: DOI (when applicable)
- Access rights: closed access, embargoed access, restricted access, open access.

The final dataset will be accompanied by this information under the form of a README.txt document. 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.

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

Where will the data be stored?

All other electronic files (text documents, images, sequences): will be stored on KU Leuven servers, with hourly on-site backup and mirroring or stored on a cloud-based service offered by KU Leuven (OneDrive). In addition data will be regularly backed -up and stored on an external harddrive.

- Tissue samples: Tissues will be stored locally in the laboratory.

- Omics data: omics data generated during the project will either be stored on KU Leuven servers or on The Flemish Supercomputer Centre (VSC), initially in the staging area and later in the archive area.

- Nucleic acid and protein sequences: All nucleic acid and protein sequences generated during the project will be stored on KU Leuven servers. Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes), NCBI Gene Expression Omnibus (microarray data / RNA-seq data / CHIPseq data), the Protein Database (for protein sequences), the EBI European Genome-phenome Archive (EGA) for personally identifiable (epi)genome and transcriptome sequences.

- Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer). All published vectors and the associated sequences will be sent to the non-profit plasmid repository Addgene, which will take care of vector storage and shipping upon request.

- Bacterial and yeast strains will be stored in a -80°C freezer in the lab of Ludo Van Den Bosch at KU Leuven. Costs are covered by general lab expenses.

- Cell lines: Newly created human cell lines will be stored locally in the laboratory in liquid nitrogen storage and will be deposited in the UZ Leuven-KU Leuven Biobank. Other human cell lines will be stored locally in liquid nitrogen cryostorage of the laboratory when actively used for experiments.

- Genetically modified organisms: Zebrafish will be maintained in the zebrafish facility of the Ludo Van Den Bosch lab which applies Standard Operation Procedures concerning housing, feeding, health monitoring to assure consistent care in accordance with European and national regulations and guidelines. All animals will be registered in the Leuven Animal Information System (LAIS) database, along with corresponding genotyping information, ethical approval documents and animal provider receipts. Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate: -80°C for nucleic acids, protein samples, living organisms (zebrafish) will be stored with appropriate backup copies of the stock in the Van Den Bosch lab at KU Leuven. Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate

How will the data be backed up?

KU Leuven drives are backed-up according to the following scheme:
- data stored on the “L-drive” is backed up daily using snapshot technology, where all incremental changes in respect of the previous version are kept online; the last 14 backups are kept.
- data stored on the “J-drive” is backed up hourly, daily (every day at midnight) and weekly (at midnight between Saturday and Sunday); in each case the last 6 backups are kept.
- data stored on the digital vault is backed up using snapshot technology, where all incremental changes in respect of the previous version are kept online. As standard, 10% of the requested storage is reserved for backups using the following backup regime: an hourly backup (at 8 a.m., 12 p.m., 4 p.m. and 8 p.m.), the last 6 of which are kept; a daily backup (every day) at midnight, the last 6 of which are kept; and a weekly backup (every week) at midnight between Saturday and Sunday, the last 2 of which are kept.
- All omics data stored on the Flemish Supercomputer Centre (VSC) will be transferred on a weekly basis to the archive area which is backed up.
Incremental backups are done daily from one 20 TB QNAP NAS to a second 20 TB QNAP NAS.

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

There is sufficient storage and back-up capacity on all KU Leuven servers:
- the “L-drive” is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp eseries storage systems, and a CTDB samba cluster in the front-end.
- the “J-drive” is based on a cluster of NetApp FAS8040 controlers with an Ontap 9.1P9 operating system.

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

Both the “L-drive” and “J-drive” servers are accessible only by laboratory members, and are mirrored in the second ICTS datacenter for business continuity and disaster recovery so that a copy of the data can be recovered within an hour.

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

The total estimated cost of data storage during the project is based on the following costs:
- The costs of digital data storage are as follows: €173,78/TB/Year for the “L-drive” and €519/TB/Year for the “J-drive”.
- Total maintaining of zebrafish costs, including multiple lines, the fish food, housing system, personnel’s costs... is about €10.000/year. The PI and the lab manager if applicable are responsible for the preservation of zebrafish lines as well as administrative and experimental data. All published lines will be preserved for the remainder of the PI’s research career. All unpublished lines will be preserved for a minimum of 5 years after the end of the project.
- Electricity costs for the -80° freezers present in the labs are included in general lab costs.
- Data storage and backup costs are included in general lab costs.

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 minimum preservation term of 10 years after the end of the project will be applied to all datasets. All datasets 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.

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

As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication. For all other datasets, long term storage will be ensured as follows:
- Digital datasets: files will be stored on the “L-drive”.
- Novel transgenic zerbrafish lines will be stored locally in the laboratory.
- Biological samples (protein, RNA, cDNA etc.) will be stored locally in the laboratory.
- Tissue samples: Tissues will be stored locally in the laboratory.
- Omics data: datasets will be stored on the “L-drive”.
- Vectors: As a general rule at least two independently obtained clones will be preserved for each vector in the form of purified DNA (in -20°C freezer).
- Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.
Following publication, the results associated with each study will also be deposited in the Dryad repository, where they will be preserved indefinitely.

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

-The costs of digital data storage are as follows: 173,78€/TB/Year for the “L-drive” and 519€/TB/Year for the “J-drive”.
- Maintaining a zebrafish line alive costs about 500 euro per year.
These costs will be met partially by the benchfee provided by the FWO and partially by existing lab grants.
Electricity costs for the -80° freezers present in the labs are included in general lab costs.
Data storage and backup costs are included in general lab costs.

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
- Yes, in a restricted access repository (after approval, institutional access only, …)
- Other, please specify:

Also, upon request by e-mail.

As a general rule, datasets will be made openly accessible via existing platforms that support FAIR data sharing (www.fairsharing.org). Sharing policies for specific research outputs are detailed below:

- Omics datasets will be deposited in open access repositories such as the PRIDE Archive for proteomics data.
- Vectors: Upon publication, all vectors supporting a manuscript will be made publicly available via the non-profit plasmid repository Addgene, along with the corresponding DNA sequences. Addgene in turns performs quality control on the DNA, curates the plasmids online with all relevant information (maps, sequences), and for a minimal cost (typically $65) ships the vectors upon simple request and signature of a material transfer agreement. The MTA will be prepared before depositing the vectors with the help of our organization's Tech Transfer office. For transfer between nonprofit or academic institutions, Addgene typically uses the Uniform Biological Material Transfer Agreement (https://www.addgene.org/terms/1047/). All non-published vectors and the associated documentation will be shared by the PI upon request and after signature of a material transfer agreement, at no cost except the cost of shipment.
- Cell lines: All human pluripotent cell lines supporting publications will be registered in hPSCreg, the European human embryonic stem cell registry supported by the European Commission (https://hpscreg.eu/). Information about the deposited lines (including donor information, derivation method, availability and characterization) will also be made accessible. Registration of cell lines in hPSCreg will provide visibility, confirm ethical procurement and facilitate comparison with other hPSC lines. The PI will remain the distributor of the pluripotent cell lines.
- Genetically modified organisms: All genetically modified organisms used in publications will be made available to researchers upon request at the time of publication.
- Other digital datasets that support publications (including image, video or audio files, spectroscopy data and simulation data) will be made publicly available via an open research data platform such as Mendeley Data or Zenodo.
- Antibodies, synthetic and recombinant compounds: samples will be stored as appropriate in the laboratory. Within availability, they will be shared with interested researchers upon request.
- Research documentation: All protocols used to generate published data will be described in the corresponding manuscript(s), and the related documentation will be included as supplementary information. These data and all other documents (daily logs, raw data) are accessible to the PI and the research staff, and will be made available upon request.
- Manuscripts: All scientific publications will be shared openly. Manuscripts submitted for publication will be deposited in a pre-print server such as bioRxiv, arXiv, Nature Precedings or ASAPbio). At the time of publication, research results will be summarized on the PI's website (add website address) and postprint pdf versions of publications will be made available there if allowed by copyright agreements, possibly after an embargo as determined by the publisher. Before the end of the embargo or in cases where sharing the post-print is not allowed due to copyright agreements, a pre-print version of the manuscript will be made available. Publications will also be automatically listed in our institutional repository, Lirias 2.0, based on the authors name and ORCID ID.
- Nucleic acid and protein sequences: Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes).
- Data that do not support publication will be either deposited in an open access repository or made available upon request by email.

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

Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. These repositories clearly describe their conditions of use (typically under a Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication or an ODC Public Domain Dedication and License, with a material transfer agreement when applicable). Interested parties will thereby be allowed to access data directly, and they will give credit to the authors for the data used by citing the corresponding DOI. For data shared directly by the PI, a material transfer agreement (and a non-disclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted.

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

We aim at communicating our results in top journals that require full disclosure of all included data. Biological material will be shared upon simple request following publication, unless we identify valuable IP, in which case we will first protect commercial exploitation, either through patenting or via an MTA that restricts the material from commercial use.

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

In an Open Access respository, in a restricted access repository, upon request by mail.

When will the data be made available?

As a general rule all research outputs will be made openly accessible at the latest at the time of publication. No embargo will be foreseen unless imposed e.g. by pending publications, potential IP requirements – note that patent application filing will be planned so that publications need not be delayed - or ongoing projects requiring confidential data. In those cases, datasets will be made publicly available as soon as the embargo date is reached.

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

Datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. These repositories clearly describe their conditions of use (typically under a Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication or an ODC Public Domain Dedication and License, with a material transfer agreement when applicable). Interested parties will thereby be allowed to access data directly, and they will give credit to the authors for the data used by citing the corresponding DOI. For data shared directly by the PI, a material transfer agreement (and a non-disclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted.

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

Manuscripts: Metadata information will be submitted alongside the final version of the manuscript, including the names, titles, email addresses, ORCIDs and affiliations of all authors. Upon publication, this metadata information will also be submitted to bibliographic databases such as Medline. All manuscripts will be assigned a unique Digital Object Identifer (DOI) by the publisher. Manuscripts will be given a descriptive title, and will be accompanied by keywords provided by the authors in order to maximize their findability.

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

It is the intention to minimize data management costs by implementing standard procedures e.g. for metadata collection and file storage and organization from the start of the project, and by using free-to-use data repositories and dissemination facilities whenever possible. Data management costs will be covered by the laboratory budget.

6. Responsibilities

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

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the laboratory notebook that refer to specific datasets, and additionally compiling applicable metadata along with the data in the manner described above.

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

The research and technical staff will ensure data storage and back up, with support from René Custers and Alexander Botzki for the electronic laboratory notebook and from Raf De Coster for the KU Leuven drives.

Who will manage data preservation and sharing?

The PI is responsible for data preservation and sharing, with support from the research and technical staff involved in the project, from René Custers and Alexander Botzki for the electronic laboratory notebook and from Raf De Coster for the KU Leuven drives.

Who will update and implement this DMP?

The PI is ultimately responsible for all data management during and after data collection, including implementing and updating the DMP.