

---

## Plan Overview

*A Data Management Plan created using DMPOnline.be*

**Title:** A whole-organism single-cell approach to uncover gene regulatory mechanisms in *Drosophila melanogaster*: mapping the developmental enhancer space

**Creator:** Hannah Dickmanken

**Principal Investigator:** n.n.

**Data Manager:** Hannah Dickmanken

**Project Administrator:** Hannah Dickmanken

**Affiliation:** KU Leuven (KUL)

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

**Template:** FWO DMP (Flemish Standard DMP)

**Principal Investigator:** n.n. n.n.

**Data Manager:** Hannah Dickmanken

### Project abstract:

*Single-cell transcriptomics and single-cell epigenomics provide a powerful means of cataloguing known cell types, discovering novel ones, and deciphering their underlying regulatory principles. I propose to generate a single-cell multi-omics atlas, combining scRNA-seq and scATAC-seq, across the fruit fly *Drosophila melanogaster* from embryo to larvae, to pupae, to adult. (1) So far, I covered the whole larval development in chromatin accessibility and continue with the pupal development, both in closely set time windows. I annotate the scATAC-seq atlas using bridge integration with the previously published Fly Cell Atlas and developmental tissues. (2) My multi-ome atlas will be used to infer gene regulatory networks at the whole-organism level focusing on cell type identity and function in development. Along inferred trajectories, I will study transcription factor (TF) and enhancer pleiotropy; I will determine the optimal combinations of TFs per cell type; and I will study network motifs (e.g., feedback and feedforward loops) and gene repression. (3) Importantly, I will combine the computational predictions with genetic experiments in an iterative fashion, leading to new biological insight and better computational models. For the in vivo validations, I will take advantage of the advanced fly genetics toolbox. This will include the generation of transgenic flies for synthetic enhancer design and the use of single-cell high-throughput TF-binding assays adapted for *Drosophila*.*

**ID:** 211158

**Start date:** 01-11-2024

**End date:** 31-10-2028

**Last modified:** 21-11-2024

**A whole-organism single-cell approach to uncover gene regulatory mechanisms in *Drosophila melanogaster*: mapping the developmental enhancer space**

**DPIA**

---

DPIA

- Not applicable

**A whole-organism single-cell approach to uncover gene regulatory mechanisms in *Drosophila melanogaster*: mapping the developmental enhancer space**

**GDPR**

---

**GDPR**

- Not applicable

# A whole-organism single-cell approach to uncover gene regulatory mechanisms in *Drosophila melanogaster*: mapping the developmental enhancer space

## Application DMP

---

### Questionnaire

*A major aim of this project is to generate single-cell RNA-seq and single-cell ATAC-seq data on *Drosophila melanogaster*. We will generate, process and store:*

*(a) Experimental data including digital images (microscopy pictures, gel scans, graphs, illustrations, figures) and omics data (genomics, transcriptomics, epigenomics data);*

*(b) Simulation data;*

*(c) Compiled data including research documentation (generated by the staff or collected from online resources, including laboratory notes, protocols), manuscripts, algorithms, scripts and software;*

*(d) Canonical data: nucleic acid sequences.*

1. *Responsible person: Stein Aerts (PI).*

*During the research omics data will be stored on the servers of KU Leuven or of the Flemisch Supercomputer Centre. Algorithms, scripts, and software will be stored in private online git repositories of the PI.*

*After the research, datasets will be made openly accessible, whenever possible via existing platforms. Nucleic acid sequences supporting the manuscript will be made publicly available via repositories (e.g. NCBI Gene Expression Omnibus). Published vectors and associated sequences will be sent to the non-profit plasmid repository Addgene. For all other datasets, long-term storage will be ensured on the aforementioned servers and local computers.*

NA

NA

NA

# A whole-organism single-cell approach to uncover gene regulatory mechanisms in *Drosophila melanogaster*: mapping the developmental enhancer space

## FWO DMP (Flemish Standard DMP)

### 1. Research Data Summary

				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
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Generate new data</li> <li>• Reuse existing data</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Digital</li> <li>• Physical</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Observational</li> <li>• Experimental</li> <li>• Compiled/aggregated data</li> <li>• Simulation data</li> <li>• Software</li> <li>• Other</li> <li>• NA</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• .por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ...</li> <li>• NA</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• &lt;100MB</li> <li>• &lt;1GB</li> <li>• &lt;100GB</li> <li>• &lt;1TB</li> <li>• &lt;5TB</li> <li>• &lt;10TB</li> <li>• &lt;50TB</li> <li>• &gt;50TB</li> <li>• NA</li> </ul>	
developmental fly data	newly generated scATAC and scRNA data of the developing fly	Generate new data	digital	experimental	NA, fragment files	<10TB	
wing disc	scmultiome data of the L3 wing disc	reused	digital	experimental	NA, fragment files	<1GB	
eye disc	sc ATAC data	reused	digital	experimental	NA, fragment files	<1GB	
brain development	sc ATAC data	reused	digital	experimental	NA, fragment files	<1TB	

Reuse of developmental *Drosophila* data:  
Wing disc: <https://elifesciences.org/articles/81173>  
eye disc: <https://www.embopress.org/doi/full/10.15252/msb.20209438>  
Fly brain: <https://www.nature.com/articles/s41586-021-04262-z>

- No
- NA, *Drosophila melanogaster* is not classified as an animal.

- No

- No

- No

- No

## 2. Documentation and Metadata

I keep an overview of my data management plan in my personal folder of the lab located at the VSC. The data will be stored in a separate data folder at the end of my PhD. Additionally, I keep a personal electronic lab book for my computational work and use the Elab electronic lab book for my wet lab experiments. My notebooks to analyse my data are also included in my projects folders for clear overview of which analysis is conducted in which data set. I also keep an excel sheet (KUL drive, shared with technicians) with an overview of my wet lab experiments and the stored data resulting out of it.

- Yes

Metadata of the wet lab experiments is stored in the Elab journal used by VIB. I also keep an excel sheet (KUL drive, shared with technicians) with an overview of my wet lab experiments and the stored data resulting out of it.

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

on the VSC partition of the lab (Tier-1)

The data will be backed up in ManGo (Tier-1 Data)

- Yes

Yes, there is enough storage space on Tier-1 for this project.

the accounts to access ManGo are password protected. The general VSC security measurements are taken.

NA

## 4. Data preservation after the end of the research project

*After the research, datasets will be made openly accessible, whenever possible via existing platforms. Nucleic acid sequences supporting the manuscript will be made publicly available via repositories (e.g. NCBI Gene Expression Omnibus). Published vectors and associated sequences will be sent to the non-profit plasmid repository Addgene. For all other datasets, long-term storage will*

*be ensured on the aforementioned servers and local computers.*

*After the research, datasets will be made openly accessible, whenever possible via existing platforms. Nucleic acid sequences supporting the manuscript will be made publicly available via repositories (e.g. NCBI Gene Expression Omnibus). Published vectors and associated sequences will be sent to the non-profit plasmid repository Addgene. For all other datasets, long-term storage will be ensured on the aforementioned servers and local computers.*

NA

## 5. Data sharing and reuse

- Yes, in an Open Access repository

*Nucleic acid sequences supporting the manuscript will be made publicly available via repositories (e.g. NCBI Gene Expression Omnibus).*

NA

- No

*Nucleic acid sequences supporting the manuscript will be made publicly available via repositories (e.g. NCBI Gene Expression Omnibus).*

*After the research, datasets will be made openly accessible, whenever possible via existing platforms.*

*After the research, datasets will be made openly accessible, whenever possible via existing platforms. All users of those platforms (e.g., NCBI) can access the data freely.*

- Yes

NA

## 6. Responsibilities

me personally

me personally

Responsible person: Stein Aerts (PI).

me personally