

THÈSE D' HABILITATION À DIRIGER LA RECHERCHE

Soutenue à AMU — Aix-Marseille Université
le 10 janvier 2025 par

Léo Guignard

Computer Science, Morphogenesis and Variability

Sciences du vivant

Biologie computationnelle

École doctorale 658 - Sciences du Vivant

IBDM

Consignes de présentation détaillées
des pages liminaires en annexe B

Composition du jury

Prénom NOM Titre et affiliation	Rapporteur·e
Prénom NOM Titre et affiliation	Rapporteur·e
Prénom NOM Titre et affiliation	Examinateur·rice
Prénom NOM Titre et affiliation	Président·e du jury
Pierre-François Lenne Directeur de Recherche, IBDM	Tuteur
Prénom NOM Titre et affiliation	Membre invité·e

Abstract and keywords

At the interface between developmental biology and computer science, I address together questions about embryonic morphogenesis, computer vision, data science and data visualisation. By developing novel and innovative computational algorithms I sought to quantify, model and better understand embryonic development at the single cell scale in whole organisms throughout development.

More specifically, I want to understand how the different forms and functions of tissues, organs and organisms are obtained as a result of cell changes at the morphological and molecular level. I will do so across three main axes/question:

1. Morphological scale: How do cells organise in coherent structures?
2. Molecular scale: What genes control that organisation?
3. Morpholecular scale: How do genes and cell organisation interplay?

Keywords: Developmental Biology, Developmental Variability, Computer Vision, Spatial Omics

Remerciements

I would like to thank my family, my friends, the IBDM, CENTURI, my group, my wife, my daughter, the examinateurs, the rapporteurs, <3

Table des matières

Abstract and keywords	2
Remerciements	3
Table des matières	4
Table des figures	5
Introduction	8
1 Previous research and scientific achievements	11
1.1 Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis	13
1.2 In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level	15
1.3 Spatial transcriptomic maps of whole mouse embryos	16
1.4 Side projects	18
1.5 Education and outreach	19
2 Future work	21
2.1 Scientific context	22
2.2 Objectives	23
2.3 Morphodynamic atlases of embryonic development	25
2.3.1 Morphological atlas of <i>P. hawaiensis</i>	26
2.3.2 Morphological atlas of <i>Drosophila melanogaster</i>	28
2.3.3 Reconstruction of gastruloid development	28
2.4 sc3D : a fast and easy dive into spatial near single-cell omics	30
2.4.1 sc3D	32
2.4.2 sc3D-viewer	33
2.4.3 Usefulness to the biological community in general	33
2.5 Morpholecular atlases	34
Conclusion	37
Bibliographie	38

Table des figures

0.1 Schematics of my research proposal	A. Recording of the development from microscopes or from spatial-single cell technics. Done by collaborators. B. Reconstruction of the acquired microscope time-series to quantify cell morphodynamics. C. Aggregation of the multiple morphodynamic reconstructions into an average atlas. D. Reconstruction of the spatial-omics datasets. E. Aggregation of the multiple spatial-omics reconstructions into an average atlas. F. Aggregation of the two atlases with different modalities into a multimodal single-cell atlas : The morpholecular atlas	10
1.1 Graphical abstract of L. GUIGNARD, FIÚZA et al. 2020	(Left) Quantitative analysis of Phallusia embryogenesis. We combined live light-sheet imaging of cell membranes (left images) with automated cell segmentation and tracking with colour-coded cell fates (centre images) to extract quantitative cell morphological properties (right images, colour-coded by cell compactness). From top to bottom : embryo at the 64-cell, mid-gastrula, and late gastrula stages. (Right) Cell signalling model. We first made simplifying assumptions concerning the distribution and diffusion of signalling pathway components (top) and then integrated cell contact geometry with gene expression profiles to predict pathway activation levels in single cells (centre) and binarized induction status (bottom).	14
1.2 Graphical abstract of McDOLE et al. 2018	(Self explanatory)	15
1.3 Graphical abstract of SAMPATH KUMA et al. 2023	Mouse embryos are sliced at different developmental stages (E8.5, E9.0, E9.5). The sections are sequenced while conserving the 2D spatial information thanks to the SlideSeqV2 method. The recordings are then reassembled to form 3D digital embryos, which in turn allows the mapping of the cell states and the visualisation of gene expression patterns.	17

2.1 Schematics of my research proposal. A. Recording of the development from microscopes or from spatial-single cell technics. Done by collaborators. B. Reconstruction of the acquired microscope time-series to quantify cell morphodynamics. C. Aggregation of the multiple morphodynamic reconstructions into an average atlas. D. Reconstruction of the spatial-omics datasets. E. Aggregation of the multiple spatial-omics reconstructions into an average atlas. F. Aggregation of the two atlases with different modalities into a multimodal single-cell atlas : The morpholecular atlas	24
2.2 Principles of time series registration. A. A set of n time series of recorded embryos. The black ticks show the acquisition timepoints. The grey arrows show a mapping between homologous timepoints. B. The time series are registered in time. Potential time dilation and compression can be seen. C. All embryos, at all shared registered timepoints, are spatially registered onto the reference embryo (here Embryo 1). D. The transformations that allow to register the embryos onto the average space are computed using the pre-computed spatial transformations. E. Once all embryos are registered onto the average space, they can be averaged	27
2.3 Recordings and reconstructions. Top left : 3D projections of a <i>P. hawaiiensis</i> embryo where the membranes and nuclei are labelled with fluorophores with two different excitation and emission wavelengths. Top row : 10h after egg laying (AEL), bottom row, 64h AEL. Top right : 3D projections of the reconstruction of 4 <i>D. melanogaster</i> embryos during germband elongation. Bottom : 5 different gastruloids where the nuclei are labelled with the mesoderm marker Brachyury and the dividing cells are labelled with the mitotic marker Ph3.	29
2.4 Single cell omics principles. A. In classic single cell omics, the cells are dissociated from each other before being recorded. This process clears away the spatial information. As a result, the processed data are often projected in 2 dimensions, in a way that groups cells together according to their molecular content. B. In spatial single cell omics (here I am referring specifically to SlideSeqV2 STICKELS et al. 2021), 2 dimensional sections are made from a 3D sample. These sections are then deposited onto an array of barcoded beads of the size of a single cell. Through gravity, each cell of the 2D slice falls onto a bead and is barcoded according to that bead. This primary barcode allows us to record the 2D position of the cell and therefore to replace its measurement in space a-posteriori.	31

Table des figures

2.5 Single-cell spatial transcriptomics of an E8.5 mouse embryo. A. Prototype of the 3D viewer for spatial single-cell datasets. Brachyury is projected on the cells. B. Example of an interaction where the user selected cells in the 2D projection on the left and the same cells are displayed on the 3D viewer.	34
---	----

Introduction

A single fertilised oocyte generates a highly organised living organism composed of thousands to billions of cells. Through divisions, shape changes, growth, rearrangement or programmed death, cells gradually form coherent groups that organise in tissues, themselves organising into organs and eventually creating a fully functional organism. This process, called embryogenesis, is highly complex and requires precise spatial and temporal coordination with precise regulation of hundreds to millions of cells, their geometry, their spatial organisation and their molecular content, to eventually produce the final functional organism. Despite its complexity, the result of this process must be reproducible : a fertilised oocyte must produce a functional adult.

One of the first steps to understanding the mechanisms driving any given process is to observe it as precisely as possible and quantify it. Developmental biologists have been observing processes for centuries to understand the underlying rules governing the creation of an organism from a single fertilised egg.

In the past ten years, technological breakthroughs have pushed the observations of biological samples to scales that were not previously accessible. More specifically, the continuous development of light-sheet fluorescence microscopy (method of the year 2014 “[Method of the Year 2014](#)” [2015](#)) has enabled the observation of whole embryos, throughout their development, at the single cell resolution, for periods lasting from hours to days. Another set of revolutionary methods are single-cell sequencing methods (methods of the year 2013 “[Method of the Year 2013](#)” [2014](#), breakthrough of the year 2018 [CRESPI 2018](#)) and more specifically spatially resolved single cell sequencing methods (method of the year 2020 “[Method of the Year 2020 : spatially resolved transcriptomics](#)” [2021](#)) that allow to precisely measure the molecular content of all the cells in a given embryo, at a given time in the development.

Altogether, these two complementary new methods have brought the realm of observation of the state of an organism to a whole new level. But, while these novel methods allow highly resolved and sometimes prolonged observations of the organisms, they generate datasets that are too large and too complex to be analysed manually, and often even to be analysed using standard computational methods. For example, light-sheet microscopes generate three-dimensional movies that weigh up to a few terabytes in memory and can contain tens of thousands of cells that must be detected and followed to quantify their dynamics. Developing computational methods to perform such

tasks automatically is currently a hot research area.

Furthermore, single-cell sequencing methods measure the concentration of tens of thousands of different molecules in tens of thousands of cells generating datasets with dimensionality far higher than what has already been treated in biology. Again, more and more groups are working on developing methods to treat these datasets.

These new methods allow biologists to quantitatively peek at developmental processes at unprecedented scales and precision. But, because of the extreme characteristics of these newly acquired datasets, new computational and mathematical methods are required to accurately extract the quantitative information buried within them. Moreover, even once the measurements are extracted, it is often nontrivial to make sense of them nor to extract rules from these observations. For that, too, novel theoretical methods have to be designed.

This is why it is now often necessary to first answer computational questions to then answer biological ones.

Objectives

My research programme consists in better understanding the interplay between cell dynamics and cell gene expression patterns. I will do so by developing novel methods to extract biologically relevant quantitative information from the previously mentioned acquisition methods. Another goal of my research program is to make this extracted quantitative information accessible to visualise, browse and query, hence easily available to me and the biology community. This work will be carried in close relationship with other biologists, computer scientists and physicists while following these three axes :

- **Axis 1 : Morphological** : Deciphering cell dynamics by quantify them in developing organisms from in-toto recording and build dynamic averages of these quantifications
- **Axis 2 : Molecular** : Systematically identifying genes of interest by quantifying single cell molecular content of developing organisms from near single-cell spatial omics and build averages of these quantification
- **Axis 3 : Morpholecular** : Bridge morphological and molecular information to understand how they interplay to control embryogenesis at the single-cell scale

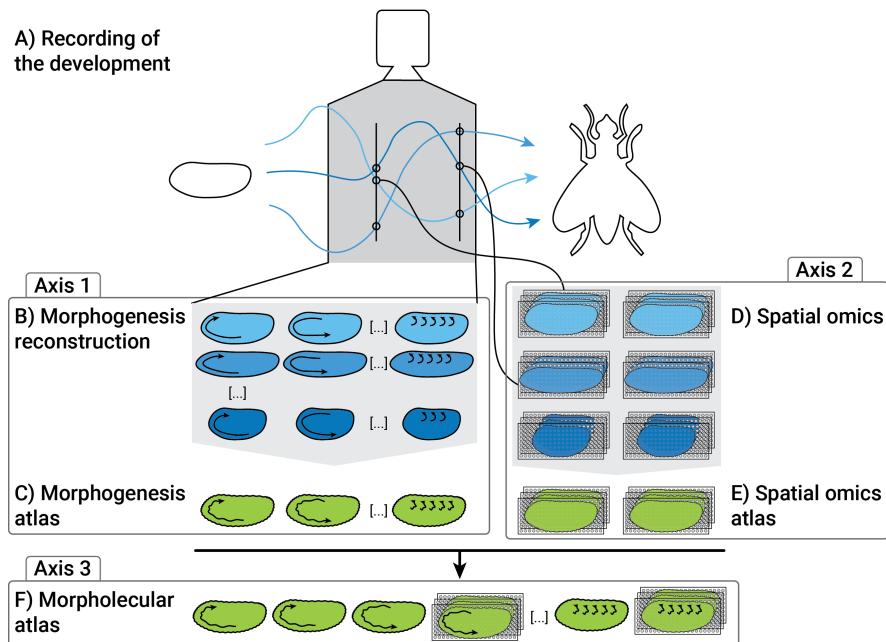


FIGURE 0.1 : Schematics of my research proposal A. Recording of the development from microscopes or from spatial-single cell technics. Done by collaborators. B. Reconstruction of the acquired microscope time-series to quantify cell morphodynamics. C. Aggregation of the multiple morphodynamic reconstructions into an average atlas. D. Reconstruction of the spatial-omics datasets. E. Aggregation of the multiple spatial-omics reconstructions into an average atlas. F. Aggregation of the two atlases with different modalities into a multimodal single-cell atlas : The morpholecular atlas

My research programme aims at bringing novel insights on both model and non-model organisms in developmental biology. While the algorithms and libraries will be primarily developed to answer my own biological questions, they will most likely be useful to the larger biology community when they will want to observe their organisms quantitatively, at the single-cell scale. Finally, the outputs of my research will be potentially useful to the applied physicist community wanting to develop data driven models of embryonic development.

1 Previous research and scientific achievements

Sommaire

1.1	Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis	13
1.2	In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level	15
1.3	Spatial transcriptomic maps of whole mouse embryos	16
1.4	Side projects	18
1.5	Education and outreach	19

From coding algorithms to decoding development, with roots in theoretical computer science, I have transitioned to a biologist.

I am and have been interested in understanding how single fertilised oocytes form highly complex, fully functional organisms. To understand the mechanisms driving embryogenesis, I have been developing computer vision and graph-based algorithms to quantitatively probe embryonic development at the single cell scale.

After my training in theoretical computer science I did a Ph.D. in biology where I sought to understand cell-cell communication in developing ascidians. This research led to the modelling of cell-cell signalling in the ascidian Phallusia mammillata. I then did a postdoc at Janelia Research Campus in the USA where I wanted to understand how cell movements shape mouse embryogenesis. To this end, I developed graph-based algorithms to reconstruct the very first dynamic single-cell scale average of the development of mouse embryos. I then wanted to link cell dynamics and gene expression still at the single cell scale. I therefore paired with experts in spatial transcriptomics. This collaboration aimed at building a 3D anatomical atlas of cell molecular content of mouse embryos at 8.5 and 9.5 days after fertilisation. This work led to the first 3D virtual *in-situ* of mouse embryos where about 27000 gene can be queried across more than 85000 cells.

I am now at the Institut de Biologie du Développement de Marseille since September 2023. There I am the head of the group : “Computer Science, Morphogenesis and Variability” where we work on understanding how developing embryos are shaped by

1 Previous research and scientific achievements –

the interplay between morphological changes and molecular dynamics in cells and tissues.

Since the start of my Ph.D. until now as the head of the group “Computer Science, Morphogenesis and Variability” at the Institut de Biologie du Développement de Marseille, I have been trying to address the following question :

How do cells organise to form a fully functional organism?

I am doing so by studying morphological dynamics as well as gene expression dynamics at the single cell scale in whole embryos. I have successfully studied and developed quantitative tools for a wide range of model organisms, including ascidians [L. GUIGNARD](#), [FIÚZA et al. 2020](#), [Drosophila MALIN-MAYOR et al. 2022](#), mouse [MCDOLE et al. 2018](#); [SAMPATH KUMA et al. 2023](#), [Parhyale WOLFF et al. 2018](#), zebrafish and Trunk-Like-Structure (organoids) [VEENVLIET et al. 2020](#). These experiences led me to be invited to write the book chapter entitled *Mathematical and bioinformatic tools for cell tracking* [HIRSCH et al. 2022](#) in the book *Cell Movement in Health and Disease*.

My work have been proven useful to the scientific community and it shows through the citation number (913 in the past 5 years¹) and the number of times my software have been downloaded (for example more than 700 thousand times² for the latest single cell omics library)

In this document I present the past research projects for which I was the main contributor and that led to publications in peer-reviewed journals :

Project 1 : Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis

Project 2 : In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level

Project 3 : Spatial transcriptomic maps of whole mouse embryos

Side projects : Brief presentation of projects that have led to publications but for which I was not the main contributor

Teaching and outreach : Brief presentation of my work towards education and outreach

¹according to [Google Scholar](#)

²according to [pypinfo](#)

1.1 Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis

While working on embryos of a marine animal from the tunicate family, *Phallusia mammillata* (*P.m.*), I investigated the dichotomy between the high developmental stereotypy and the strong genomic divergence observed within the ascidian class. As in *Caenorhabditis elegans*, ascidian cell positions and lineages are tightly conserved; in ascidians, the developmental stereotypy has also been observed down to the shape of the cells. This conservation, however, is found between ascidian species despite their rapid genomic divergence. To investigate this paradox, I collaborated with U.-M. Fiúza, who acquired several *in-toto* 3D time-series of *P.m.* embryos, to quantitatively study cell dynamics (divisions, displacements and shape changes) during embryogenesis. I developed ASTEC (for Adaptive Segmentation and Tracking of Embryonic Cells), a set of algorithms that reconstructs the shape of each cell and tracks them from a 3D time-series of images of fluorescently-labelled embryo membranes. Cell membrane segmentation and tracking provided information on the shape of every cell and its lineage throughout the recorded period of development, covering gastrulation and neurulation.

ASTEC was successfully used to build the cell segmentations and tracking of 10 wild-type and 2 perturbed embryos. Using these segmentations, we quantified *P.m.* developmental variability at the single-cell level, looking at division patterns, cell shapes and cell area of contacts. While already observed to be highly stereotyped, this study quantified for the first time developmental stereotypy. Remarkably, we found that *P.m.* development is highly constrained, down to the cell area of contact. Moreover, this quantification showed that *P.m.* developmental variability is comparable within and between embryos. This similarity, coupled with the fact that the embryos we observed were from different genetic backgrounds, showed, for the first time, that the low developmental morphological variability observed in *P.m.* development is dominated by stochastic effects rather than genetic fluctuations.

The 12 reconstructed embryos, combined with a model of cell-cell signalling we developed, revealed that communication between cells in ascidians is at least in part regulated by their surfaces of contact. This mechanism is different to the more widespread French flag model which relies on ligand diffusion and concentration decay. While this cell-cell signalling control by area of contact has been previously observed it was, to the best of our knowledge, the first time it was studied quantitatively for diffusing ligands. This mechanism is not specific to ascidian development and could very well be used in other organisms, at least for part of their development, especially when the number of cells is low, making signalling controlled by diffusion harder to tune.

1 Previous research and scientific achievements – 1.1 Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis

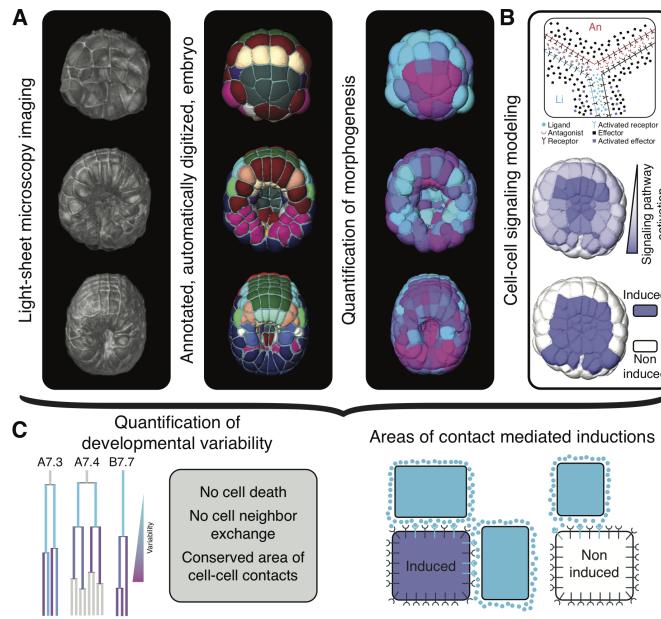


FIGURE 1.1 : Graphical abstract of L. GUIGNARD, FIÚZA et al. 2020 (Left) Quantitative analysis of Phallusia embryogenesis. We combined live light-sheet imaging of cell membranes (left images) with automated cell segmentation and tracking with colour-coded cell fates (centre images) to extract quantitative cell morphological properties (right images, colour-coded by cell compactness). From top to bottom : embryo at the 64-cell, mid-gastrula, and late gastrula stages. (Right) Cell signalling model. We first made simplifying assumptions concerning the distribution and diffusion of signalling pathway components (top) and then integrated cell contact geometry with gene expression profiles to predict pathway activation levels in single cells (centre) and binarized induction status (bottom).

We also showed *in-silico* using our model and *in-vivo* by changing the cell area of contacts in a developing embryo, that about half of the cells of the embryo saw their induction state altered by small changes in their geometry. This observation implied that the shape of these cells was under tight constraints. The maintenance of correct cell-cell communication by cell geometry rather than by protein concentration levels led us to conjecture, as a result, that the constraints on the genome were likely to be relaxed, allowing its rapid evolution and therefore the genomic divergence between ascidian species, while still conserving the cell geometry. This research led to proposing a likely mechanism that enables rapid genomic divergence while conserving cell geometry and lineage.

This project led to the publication of the article entitled Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis in the journal Science L. GUIGNARD, FIÚZA et al. 2020. This article was then relayed in

1 Previous research and scientific achievements – 1.2 In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level

the french and international press : a 1-page article in Le Monde, an article in Pour la Science and an article in The New-York Times (see my curriculum vitae for all the links). This project also led to three conference articles. One of which I was the first author L. GUIGNARD, GODIN et al. 2014 and the two others in which I was second author MICHELIN, L. GUIGNARD, FIUZA et MALANDAIN 2014; MICHELIN, L. GUIGNARD, FIUZA, LEMAIRE et al. 2015.

1.2 In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level

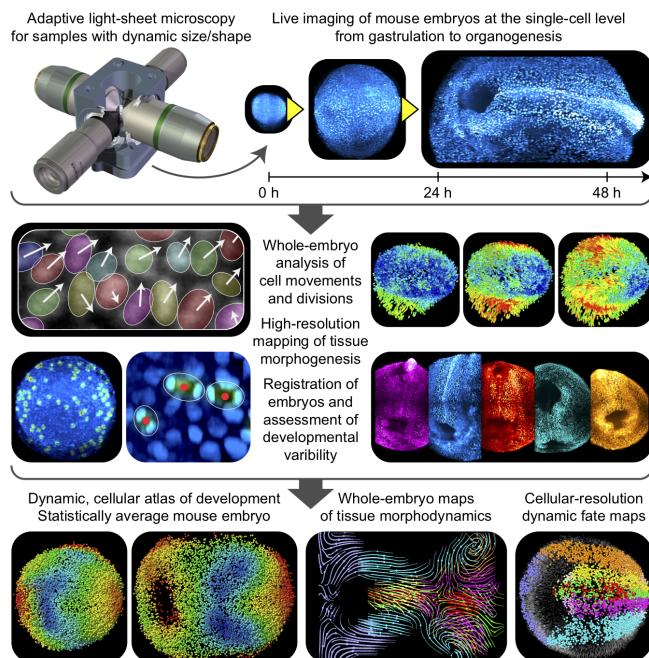


FIGURE 1.2 : Graphical abstract of McDole et al. 2018 (Self explanatory)

While working with mouse embryos, I sought to quantify, at the single-cell level, the cell dynamics that helped disentangle the multiple tissues and organs formed during the two days of development following mouse embryo implantation. Knowledge about tissue patterning at these stages of development had previously mainly come from static snapshots, leaving cell dynamics unexplored. I worked in collaboration with K. McDole, who developed the imaging protocol and acquired the *in-toto* 3D time-series of fluorescently-labelled nuclei of the mouse post-implantation embryo. Using these 3D time-series, I developed a new pipeline of algorithms to track cell

movements throughout the entirety of the 48 hours-period imaged. This pipeline allowed us to determine the spatial origin and destination of every cell at any point in time. Combining this output together with a manual labelling of tissues at the final time point of the time-series, we were able to back propagate the origin of each tissue type and to observe their disentanglement. This showed an unexpectedly already present spatial organisation of cells in tissues at early stages of development.

In order to offer a more complete quantification of cell dynamics, I developed TARDIS, an algorithm that non-linearly co-registers, in space and time, multiple developing embryos and builds a synthetic average embryo embedded in an average space. This average embryo was the first quantitative atlas of mouse post-implantation development, at the single-cell scale. It gives access not only to average cell behaviours but also to the variability observed across the embryos used to build the atlas. Moreover, we integrated the manual tissue labelling from our set of embryos to build a statistical fate map allowing us to quantify for any cell its probability to acquire a specific future fate, given its spatio-temporal position.

Ultimately, we built tools to register single segmented time points onto our atlas. These tools provide the ability to observe and quantify the most likely dynamics of targeted cells. Our atlas can therefore be used as a proxy to observe dynamics of cells that could not be identified *in vivo*. The statistical morphodynamics atlas that we created has been described as “the most detailed view ever of a developing mouse” and is the starting point of further quantitative analyses of the mouse post-implantation development.

This project led to an article entitled In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level that was published in the journal Cell McDole et al. 2018. This article was extremely well received by the scientific community, so much so that Science and Nature methods each dedicated a press article about the project. Moreover, the general press also relayed the project with articles published in The New-York Times and in Wired (see the curriculum vitae for the link to the articles).

1.3 Spatial transcriptomic maps of whole mouse embryos

During my previous projects I had studied embryogenesis through the lens of cell physical dynamics. While cell physical dynamics do carry a tremendous amount of information, they are missing a crucial one : the molecular content of the cells. To circumvent this issue, I started to work in collaboration with the groups of Fei Chen at the Broad institute in Cambridge, MA, USA and the group of Alexander Meissner at the Max Planck Institute for Molecular Genetics in Berlin. The main aim of our project was

1 Previous research and scientific achievements – 1.3 Spatial transcriptomic maps of whole mouse embryos

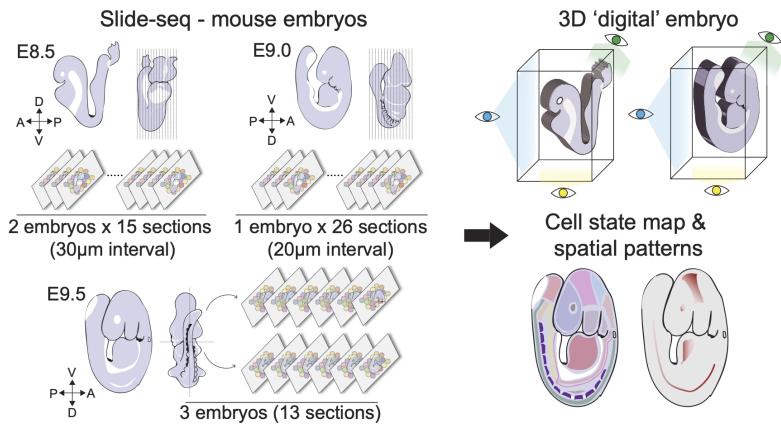


FIGURE 1.3 : Graphical abstract of SAMPATH KUMA et al. 2023. Mouse embryos are sliced at different developmental stages (E8.5, E9.0, E9.5). The sections are sequenced while conserving the 2D spatial information thanks to the SlideSeqV2 method. The recordings are then reassembled to form 3D digital embryos, which in turn allows the mapping of the cell states and the visualisation of gene expression patterns.

to provide to the community a tool that allowed looking at virtually all genes expressed in a whole 3 dimensional embryo at a near single-cell resolution while keeping the spatial context of these cells. With the follow-up idea that one could then combine the morphological and molecular recordings into one comprehensive and multimodal atlas.

The group of F. Chen developed SlideSeqv2 and the group of A. Meissner developed the protocol to use SlideSeqv2 with mouse embryos. SlideSeqv2 is a method that allows the measurement of the molecular content of single cells while keeping their 2D spatial information.

While these methods output 2D point clouds, they could not directly be exploited to look for spatially differentially expressed genes. In that context I developed a set of graph-based methods to reconstruct, display and query the datasets created using SlideSeqv2. The challenges to fully exploit these dataset are multiple. First it is necessary to register together the different 2D point clouds to reconstruct a coherent 3D structure. Then, it is necessary to develop methods to query the datasets. One main question that we were asking was about automatically and systematically finding all the genes that have specific expression patterns. In other words, we were interested in finding genes that were expressed in a localised manner within a given tissue. To automatically and systematically detect genes that were locally expressed, I developed a method that compared the average density of the spatial graph of the cells of the embryo. The two densities that I compared were the initial density of a given tissue against the density of the sub-graph of the expressing cells in that tissue. If the gene

expression is randomly distributed, the two densities are linearly correlated, if the gene expression is localised, this linear correlation is lost. By computing the deviation from the expected linear correlation I was able to find known locally expressed genes but also to find novel genes that might be of interest for the biologists.

Finally, to be able to investigate these datasets, it is important to visualise them, in 3D, dynamically. So, I hijacked an existing program (napari) dedicated to image visualisation to make it display 3D point clouds that incorporate about 27000 measurements per point. Doing so it is possible to visualise the 85000 cells and their 27000 measurements, live, dynamically, and to project onto it any kind of precomputed information.

This project led to the publication of an article entitled [Spatial transcriptomic maps of whole mouse embryos](#) that was published in *Nature Genetics* SAMPATH KUMA et al. 2023. Following up this project I was interviewed by *The Biologist* as an expert in spatial omics. Moreover, the core software that I developed for this project has been installed about 700 thousand times according to [pypinfo](#).

1.4 Side projects

In parallel to developing the above mentioned projects, I punctually worked in collaboration with biologists to help answer specific questions.

Parahyale. I worked with A. Pavlopoulos from the Institute of Molecular Biology and Biotechnology, Greece, on Parahyale limb development. In this context, I wanted to understand if it was possible to predict the final position of a cell in a *P. hawaiensis* limb from its division pattern. By developing an algorithm that compares different cell lineage trees, I showed that indeed cells closer to each other in space have more similar division patterns than cells that live further away. That led to a publication entitled [Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb](#) in the journal *eLife* WOLFF et al. 2018.

Drosophila. I also worked in collaboration with C. Malin-Mayor from the group of J. Funke at JRC to help develop a deep-learning algorithm to detect and track nuclei in developing embryos using a minimal number of sparse annotations. While deep-learning has displayed great performances for cell detection and tracking, one of its bottlenecks today is that these algorithms need extensive and comprehensive annotations to be trained. In this project we developed an algorithm which does not need extensive annotation which greatly reduces the load and the difficulty of manual annotation for the training. This project led to the publication of the article entitled [Automated reconstruction of whole-embryo cell lineages by learning from sparse annotations](#) in the journal *Nature Biotechnology* MALIN-MAYOR et al. 2022.

Zebrafish. I worked with Y. Wan, again from the group of P.J. Keller, to develop a pipeline to stabilise 3D+t movies of Zebrafish spinal cord calcium activity. As the embryo was imaged in non-immobilised conditions, muscular contractions induced a non-linear deformation of the spinal cord therefore preventing the computation of neural activity. The pipeline I developed allowed the registration of consecutive time points in a non-linear fashion allowing the computation of neural activity and ultimately confirming that the spinal cord was behaving similarly in immobilised and non-immobilised conditions. This project led to an article entitled Single-Cell Reconstruction of Emerging Population Activity in an Entire Developing Circuit in the journal *Cell* WAN et al. 2019.

Organoids. I worked in collaboration with J. Veenvliet and A. Bolondi from the groups of B.G. Herrmann and A. Meissner at the Max Planck Institute for Molecular Genetics in Berlin, to develop tools to segment and quantify the shape of neural tubes and guts in Trunk-Like-Structures resulting from aggregates derived from mouse embryonic stem cells in an extracellular matrix compound (organoids). In this study, I quantified the morphology of the neural tube and the gut validating their similarity to the orthologous tissues grown in wild-type conditions. This project led to an article entitled Mouse embryonic stem cells self-organise into trunk-like structures with neural tube and somites published in the journal *Science* VEENVLIET et al. 2020.

1.5 Education and outreach

I firmly believe that part of the duty of a researcher is to teach science to younger generations and to promote science to non-scientists. This is why I have been teaching since the start of my Ph.D.. I have been teaching computer science mainly to biologists but also to computer scientists, mathematicians and physicists.

During my Ph.D. I started teaching the basics of computer science independently of its use to biologists to university students. First to students from Nime university but also, more punctually, to students from École Normale Supérieur de Paris during one of their summer schools. During these classes I was teaching initiation to programming, graph theory and data analysis.

After my Ph.D. I continued teaching as a teaching assistant during the Advanced School of Quantitative Biology – Mechanics and Mechanisms of Morphogenesis of the Kavli Institute for Theoretical Physics, Santa-Barbara, CA, USA (2016). There I taught biologists how to extract and analyse data from images acquired by fluorescence microscopy.

But most of my teaching efforts arose from my years as a group leader at the Turing Centre for living system. There, I fully contributed to developing teaching units for the

1 Previous research and scientific achievements – 1.5 Education and outreach

interdisciplinary master Computer Science, Mathematics and Biology (CMB). First by co-leading the Professional Path teaching unit for Master 1 CMB students. Then by teaching “Implementing Turing Patterns using design patterns in Python” to Master 2 CMB students. Finally, I have been organising a yearly course : “Introduction to biological Data training course” on data analysis in Python for Ph.D. students and postdoctoral researchers of the CENTURI programme. In this yearly crash course I teach initiation to Python to pre/postdocs who would like to develop their own methods to analyse the data they have generated.

In parallel to teaching computer science to computer scientists, biologists and more, I have been engaged in outreach actions. I gave interviews to broad audience journals such as Le Monde, Pour la Science and Howard Hugh Medical Institute outreach as an expert in developmental biology and image analysis for the articles mentioned above L. GUIGNARD, FIÚZA et al. 2020; McDOLE et al. 2018. I also have been interviewed by The Biologist for my work in SAMPATH KUMA et al. 2023 as an expert in spatial omics.

I have also been part of the making of two outreach videos. First the video “Cries or whispers between cells” that explain my Ph.D. work for a wide audience as a cartoon. Second for a video to promote COVID vaccination : “Letter to Jade, the Covid vaccination explained to teenagers”. For the first video I was consulted as an expert in developmental biology and computer science and for the second video I was consulted as an expert in data analysis.

Finally, for the last 3 years I have been promoting computer science applied to biology by co-organising a Hackathon to bring computer scientists and engineers to work on biological problems. These hackathons have been proven popular with overwhelming positive feedback from the participants and the project leaders.

2 Future work

Sommaire

2.1	Scientific context	22
2.2	Objectives	23
2.3	Morphodynamic atlases of embryonic development	25
2.3.1	Morphological atlas of <i>P. hawaiensis</i>	26
2.3.2	Morphological atlas of <i>Drosophila melanogaster</i>	28
2.3.3	Reconstruction of gastruloid development	28
2.4	sc3D : a fast and easy dive into spatial near single-cell omics	30
2.4.1	sc3D	32
2.4.2	sc3D-viewer	33
2.4.3	Usefulness to the biological community in general	33
2.5	Morpholecular atlases	34

At the interface between developmental biology and computer science, I address together questions about embryonic morphogenesis, computer vision, data science and data visualisation. By developing novel and innovative computational algorithms I sought to quantify, model and better understand embryonic development at the single cell scale in whole organisms throughout development.

More specifically, I want to understand how the different forms and functions of tissues, organs and organisms are obtained as a result of cell changes at the morphological and molecular level. I will do so across three main axes/question :

1. Morphological scale : How do cells organise in coherent structures?
2. Molecular scale : What genes control that organisation?
3. Morpholecular scale : How do genes and cell organisation interplay?

Major expected outcomes of my future work :

My future work will produce the following major outcomes :

1. Developmental Biology : Quantitative and statistical understanding of cell dynamics during embryogenesis across multiple model and non-model organisms

2. Computer science : Improved algorithms to detect and track cells from fluorescence microscopy and algorithms to reconstruct and mine near single-cell spatial transcriptomics
3. Data science : Ways to display, browse and query datasets in a seamless manner
4. Dataset : Statistical atlases of the development of multiple model and non-model organisms

2.1 Scientific context

A single fertilised oocyte generates a highly organised living organism composed of thousands to billions of cells. Through divisions, shape changes, growth, rearrangement or programmed death, cells gradually form coherent groups that organise in tissues, themselves organising into organs and eventually creating a fully functional organism. This process, called embryogenesis, is highly complex and requires precise spatial and temporal coordination with precise regulation of hundreds to millions of cells, their geometry, their spatial organisation and their molecular content, to eventually produce the final functional organism. Despite its complexity, the result of this process must be reproducible : a fertilised oocyte must produce a functional adult.

One of the first steps to understanding the mechanisms driving any given process is to observe it as precisely as possible and quantify it. Developmental biologists have been observing processes for centuries to understand the underlying rules governing the creation of an organism from a single fertilised egg.

In the past ten years, technological breakthroughs have pushed the observations of biological samples to scales that were not previously accessible. More specifically, the continuous development of light-sheet fluorescence microscopy (method of the year 2014 “[Method of the Year 2014](#)” [2015](#)) has enabled the observation of whole embryos, throughout their development, at the single cell resolution, for periods lasting from hours to days. Another set of revolutionary methods are single-cell sequencing methods (methods of the year 2013 “[Method of the Year 2013](#)” [2014](#), breakthrough of the year 2018 [CRESPI 2018](#)) and more specifically spatially resolved single cell sequencing methods (method of the year 2020 “[Method of the Year 2020 : spatially resolved transcriptomics](#)” [2021](#)) that allow to precisely measure the molecular content of all the cells in a given embryo, at a given time in the development.

Altogether, these two complementary new methods have brought the realm of observation of the state of an organism to a whole new level. But, while these novel methods allow highly resolved and sometimes prolonged observations of the organisms, they generate datasets that are too large and too complex to be analysed manually, and

often even to be analysed using standard computational methods. For example, light-sheet microscopes generate three-dimensional movies that weigh up to a few terabytes in memory and can contain tens of thousands of cells that must be detected and followed to quantify their dynamics. Developing computational methods to perform such tasks automatically is currently a hot research area.

Furthermore, single-cell sequencing methods measure the concentration of tens of thousands of different molecules in tens of thousands of cells generating datasets with dimensionality far higher than what has already been treated in biology. Again, more and more groups are working on developing methods to treat these datasets.

These new methods allow biologists to quantitatively peek at developmental processes at unprecedented scales and precision. But, because of the extreme characteristics of these newly acquired datasets, new computational and mathematical methods are required to accurately extract the quantitative information buried within them. Moreover, even once the measurements are extracted, it is often nontrivial to make sense of them nor to extract rules from these observations. For that, too, novel theoretical methods have to be designed.

This is why it is now often necessary to first answer computational questions to then answer biological ones.

2.2 Objectives

My research programme consists in better understanding the interplay between cell dynamics and cell gene expression patterns. I will do so by developing novel methods to extract biologically relevant quantitative information from the previously mentioned acquisition methods. Another goal of my research program is to make this extracted quantitative information accessible to visualise, browse and query, hence easily available to me and the biology community. This work will be carried in close relationship with other biologists, computer scientists and physicists while following these three axes :

- **Axis 1 : Morphological** : Deciphering cell dynamics by quantify them in developing organisms from in-toto recording and build dynamic averages of these quantifications
- **Axis 2 : Molecular** : Systematically identifying genes of interest by quantifying single cell molecular content of developing organisms from near single-cell spatial omics and build averages of these quantification
- **Axis 3 : Morpholecular** : Bridge morphological and molecular information to understand how they interplay to control embryogenesis at the single-cell scale

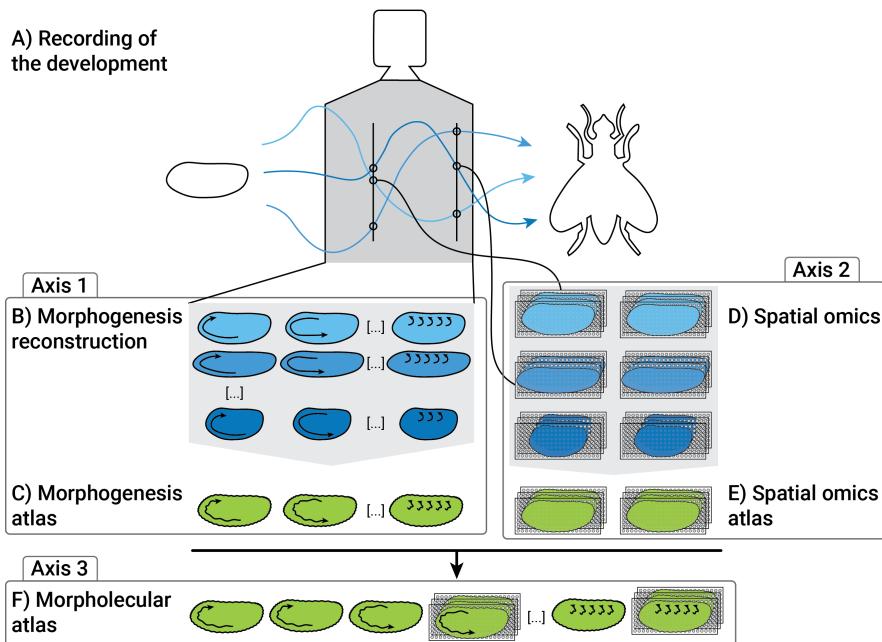


FIGURE 2.1 : Schematics of my research proposal A. Recording of the development from microscopes or from spatial-single cell technics. Done by collaborators. B. Reconstruction of the acquired microscope time-series to quantify cell morphodynamics. C. Aggregation of the multiple morphodynamic reconstructions into an average atlas. D. Reconstruction of the spatial-omics datasets. E. Aggregation of the multiple spatial-omics reconstructions into an average atlas. F. Aggregation of the two atlases with different modalities into a multimodal single-cell atlas : The morpholecular atlas

My research programme aims at bringing novel insights on both model and non-model organisms in developmental biology. While the algorithms and libraries will be primarily developed to answer my own biological questions, they will most likely be useful to the larger biology community when they will want to observe their organisms quantitatively, at the single-cell scale. Finally, the outputs of my research will be potentially useful to the applied physicist community wanting to develop data driven models of embryonic development.

Host institution : I have recently moved my own group to the Institut de Biologie du Développement de Marseille (IBDM, UMR 7288) in Marseille. Therefore, the IBDM would be my preferred host institution. The IBDM focuses on developmental biology which fits the biological questions I want to address. Moreover, the IBDM already has interdisciplinary groups : a bio-computational group (B. Habermann, bioinformatics) and a biophysics group (P.-F. Lenne group, physical approaches to morphogenesis). The IBDM greatly benefits from my current group as we are doing image analysis,

2 Future work – 2.3 Morphodynamic atlases of embryonic development

spatial transcriptomics and data analysis as well as I am greatly benefiting from all the potential collaborations with groups generating dataset that would need quantitative analysis. Moreover, my research about embryonic development across 6 different species fits well within specific groups such as the ones of P.-F. Lenne or A. Le Bivic, two groups with which I already have ongoing collaborations. Other groups, such as the one of T. Lecuit, B. Prud'homme and V. Bertrand would be fruitful terrains for new collaborations.

The IBDM is also part of the Turing Centre for living systems (CENTURI), a centre for interdisciplinary research within which I am currently a group leader.

2.3 Morphodynamic atlases of embryonic development

To understand morphogenesis and embryogenesis, biologists observe the normal development of their organisms of choice WALLINGFORD 2019. They also perform targeted disruptions on their organisms of interest and compare the disrupted development to the original one. Both methods of investigation require looking at a set of samples, as precisely as possible. Moreover, to compare two sets of behaviours (for example wild-type and perturbed) it is important to have a statistical view of the set of samples that were recorded to precisely measure significant differences. In this axis I propose to develop computational methods to quantitatively look at and measure organisms in 3 dimensions, through time and at the single-cell scale. I propose to also develop methods to aggregate multiple samples onto an average representative, a statistical atlas. These methods extract the quantitative information from recordings of embryonic development made with fluorescence microscopes where cell membranes and/or cell nuclei are labelled HUISKEN et al. 2004; KELLER et al. 2008; KRZIC et al. 2012; MEDEIROS et al. 2015.

To extract quantitative measurements from such movies and enable the quantification of organism development at the single cell scale it is necessary to develop methods to detect the cells, reconstruct their shape and track them throughout the recordings, preferably in a fully automatic fashion. Once reconstructed, these movies can then be combined to build an average representative of the wild-type (normal) development. To do so, the cell trajectories from different recordings of the organism, at the same developmental time points, must be combined to implement a more general statistical view of the average development. These statistical views can then be used for a precise, quantitative description of the variability of wild-type behaviour WOLFF et al. 2018; LI et al. 2019; L. GUIGNARD, FIÚZA et al. 2020. The statistical views can also be used to perform quantitative and statistically relevant comparisons between wild-type behaviour and perturbed one.

Therefore, this axis aims at developing computational frameworks to reconstruct, align and aggregate recordings of cell morphodynamics into average statistical representatives (Fig. 2.2). The building of these atlases not only relies on the existence of recordings of the development but also on the method used will change with the properties of the imaged sample. This axis is split into three sub-axes, each tackling the reconstruction and aggregation of different organisms that exhibit different degrees of variability and come from recordings made from either fluorescence light-sheet microscopy or bi-photon microscopy. First, I will work with already acquired *in-toto* recordings of *Parhyale hawaiensis*. Second, I will work with 50, already acquired, *in-toto* recordings of *Drosophila melanogaster* that last 7 hours from the beginning of the gastrulation. And third, I will work with already acquired recordings of cell aggregates, namely gastruloids, that mimic one of the most important developmental processes, gastrulation, where a group of undifferentiated cells gives rise to the three germ layers. Each of these axes is supported by an independent biological question but a common computational one : **How to build a quantitative and statistical view of a developing biological system?**

2.3.1 Morphological atlas of *P. hawaiensis*

One of the most pervasive, yet poorly understood, hallmarks of most metazoans is bilateral symmetry GENIKHOVICH et al. 2017. The external left and right sides of bilaterians develop separately but somehow manage to produce symmetric matching halves, presumably via deterministic developmental programmes, homeostatic buffering mechanisms or combinations thereof. These mechanisms are poorly understood WOLPERT et al. 2010; GRIMES 2019. In collaboration with Anastasios Pavlopoulos, a developmental biologist at the Institute of Molecular Biology and Biotechnology of the Foundation for Research and Technology Hellas in Heraklion, Greece, I will probe these mechanisms in the embryonic epithelial monolayer of the crustacean *Parhyale hawaiensis* that acquires a stereotypic, geometrically ordered and bilaterally symmetric tissue architecture composed of less than a thousand cells, while maintaining the capacity to restore left/right symmetry after unilateral cell ablations. Through live imaging with multi-view light-sheet fluorescence microscopy WOLFF et al. 2018 (Fig. 2.3) and image analysis of intact and ablated embryos, **I will combine novel computer vision and data analysis algorithms to decompose the shaping of each side into the individual cell contributions.** By doing so, I will quantify their levels of stereotypy and variability across sides, embryos and conditions during the normal emergence of bilateral symmetry and during its restoration after cell ablation.

In this project, I will advance the development of computational algorithms for aligning and integrating segmented image data (consisting of cell proliferation and cell death, cell shape changes and rearrangements) with cellular accuracy across sides, biological replicates and conditions. I will also develop the mathematical framework for the unbiased quantification of left/right (a)symmetries across spatiotemporal scales based on the cellular observables collected. To quantify the left/right (a)symmetries I

2 Future work – 2.3 Morphodynamic atlases of embryonic development

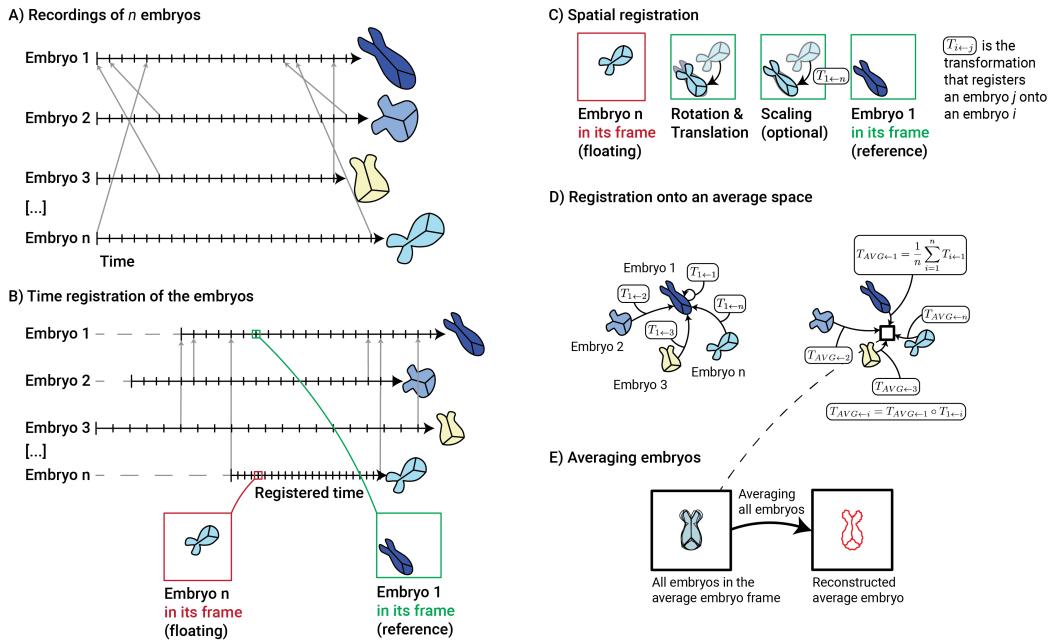


FIGURE 2.2 : Principles of time series registration. A. A set of n time series of recorded embryos. The black ticks show the acquisition timepoints. The grey arrows show a mapping between homologous timepoints. B. The time series are registered in time. Potential time dilation and compression can be seen. C. All embryos, at all shared registered timepoints, are spatially registered onto the reference embryo (here Embryo 1). D. The transformations that allow to register the embryos onto the average space are computed using the pre-computed spatial transformations. E. Once all embryos are registered onto the average space, they can be averaged

will extend the algorithm that I used in my previous work on ascidians [L. GUIGNARD, FIÚZA et al. 2020](#). This algorithm [ZHANG 1996](#), whose principles are similar to that of DNA sequence distance/alignment algorithms such as BLAST [ALTSCHUL et al. 1990](#), compares lineage trees by defining the difference between two lineage trees T_1 and T_2 (let it be $d(T_1, T_2)$) as the minimum number of weighted operations that rewrites the binary tree T_1 into T_2 : $d(T_1, T_2) = \operatorname{argmin}_{d \in D} \sum_{o \in d} w_o \cdot o$ where D is the set of all possible rewritings, d is a given rewriting, o is an atomic rewriting operation of d and w_o is the associated cost of the operation o . These measurements in wild-type and perturbed conditions will allow the detection of statistically relevant deviation from the wild-type behaviour in the perturbed embryos, pointing to the timing of interest during bilateral formation of *P. hawaiensis* embryos. These times will be used for single cell molecular content analysis (Axis 2).

2.3.2 Morphological atlas of *Drosophila melanogaster*

Drosophila melanogaster is one of the main model organisms for developmental biology, yet, to the best of my knowledge, there is no available extensive reconstruction of the cell dynamics of this fly. A reconstruction of the average cell dynamics of the embryogenesis of *D. melanogaster* would be a rich and precious asset to the community. To remedy this missing information, I will develop novel computational methods to build a 3D+time atlas of *D. melanogaster* development. To do so, I will use 50 already existing 3D wild-type time-series provided by W. Lemon from the group of P.J. Keller at Janelia Research Campus, USA. These movies cover the embryonic development of the fruit fly *D. melanogaster* for 7 hours on average, starting at the beginning of gastrulation with a frame rate of one 3D acquisition every 30 seconds. The movies were acquired using multiview fluorescence light-sheet microscopy ROYER et al. 2016. Taking advantage of state-of-the-art cell detection and tracking algorithms that I have and helped developed (TGMM AMAT et al. 2014 and SVF McDOLE et al. 2018), I will first track cell dynamics across the entire time-series (Fig. 2.3). I will then redefine the TARDIS framework I developed for building the mouse post-implantation atlas *ibid.* to scale to a high number of recorded embryos. Currently, two key limitations prevent TARDIS from being scalable to constructing an atlas that integrates a cohort of embryos : the average embryo is biassed towards a reference embryo that must be arbitrarily chosen, and manual landmarks are required. The manual process of landmark annotation in the current scheme prohibits scaling the method to a larger number of embryos. I will remove the manual landmarks by automatically detecting clusters of cells that have specific, unique properties such as local density or velocity. These clusters of cells will be used in place of the landmarks. For example, the germ-band cells during its elongation and retraction can inform about the dorso-ventral and antero-posterior orientation of the embryo and its developmental rate. Based on these automatic landmarks, I will then use combinatorial optimisation algorithms to find the set of transformations that registers every embryo onto the space of the average embryo. This set of transformations will be optimised to minimise the sum of the distances between the average and all the original embryos. This set of transformations will create a genuinely unbiased average space for the average synthetic embryo, removing the potential bias towards a reference.

2.3.3 Reconstruction of gastruloid development

As stated, morphogenesis is orchestrated by complex mechanisms involving biochemical signalling and mechanical interactions. One remarkable event during the development of a given organism is the breaking of the initial symmetry. This breaking is necessary to define the anterior and posterior parts of the embryo and later on, to define all the essential tissues. Even though microscopy has dramatically improved within the past few years, recording symmetry breaking events at the single cell scale remains difficult in-vivo.

2 Future work – 2.3 Morphodynamic atlases of embryonic development

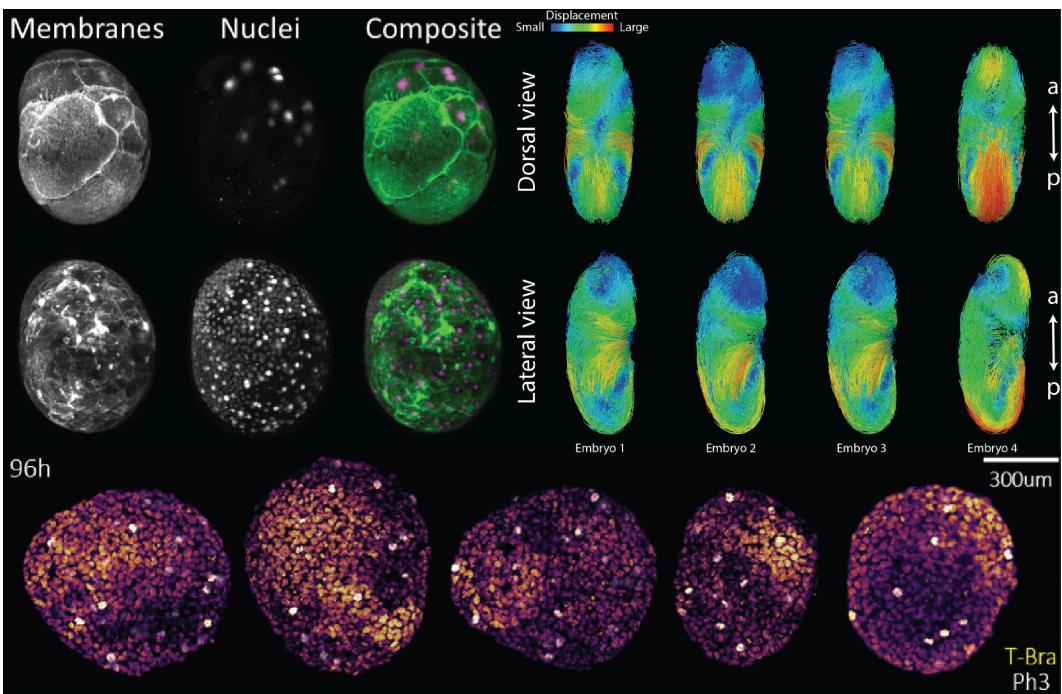


FIGURE 2.3 : Recordings and reconstructions. Top left : 3D projections of a *P. hawaiiensis* embryo where the membranes and nuclei are labelled with fluorophores with two different excitation and emission wavelengths. Top row : 10h after egg laying (AEL), bottom row, 64h AEL. Top right : 3D projections of the reconstruction of 4 *D. melanogaster* embryos during germband elongation. Bottom : 5 different gastruloids where the nuclei are labelled with the mesoderm marker Brachyury and the dividing cells are labelled with the mitotic marker Ph3.

For this reason, recent in-vitro systems mimicking mammalian embryogenesis have been developed. It has been shown that when cultured under biomimetic conditions, embryonic stem cells aggregate and organise into structures that recapitulate the early developmental stages of mouse embryogenesis BRINK et al. 2014. These systems of stem cell aggregates are called gastruloids and constitute convenient models for understanding how stem cells can self-organise to develop tissues. They present another advantage : they allow high-throughput experiments, which is essential to study the variability from one sample to another and therefore the mechanisms ensuring the robustness of development. In collaboration with S. Tlili in the group of P.-F. Lenne, at the IBDM, I am co-mentoring A. Gros, a Ph.D. student, to develop algorithms to reconstruct, track cells and quantify the observable developmental variability in gastruloids. More specifically, in this project, the biological question addressed focuses on **the role of cell divisions during the morphogenesis of gastruloids, specifically during the formation of the antero-posterior axis**. The 3D geometry and complex dynamics of such multicellular systems require developing or adapting specific tools

for imaging, segmentation and analysis. To study the mechano-genetics coupling during symmetry breaking, I am currently and will continue to co-lead the development of a pipeline to quantify in 3D the heterogeneities in gene expression and mechanical properties in gastruloids. This pipeline includes the production of experimental data, here immunostainings on fixed samples to image gene expression patterns with a two-photon microscope. From the 3D images, the pipeline is an adaptation of a deep neural network WEIGERT et al. 2020 to segment nuclei shapes, and extract a range of nuclear properties (volumes, deformations, divisions..), averaged over variable spatial scales and compared with genetic observables to look for spatial correlations with genetic observables, for instance, the mesoderm marker Brachyury. Ultimately, I would like to probe the spatio-temporal scales of variability that could be observed during gastruloid development to enable the construction of a morphogenetic atlas of their development. It would create an invaluable source of information for the rapidly growing field of organoids.

2.4 sc3D : a fast and easy dive into spatial near single-cell omics

The first axis is focused on reconstructing cell morphodynamics and building statistical representatives of embryogenesis at the single cell scale. While understanding cell movements and shape changes is important to understand embryogenesis, one crucial piece of information is missing from these measurements : cell molecular content. Indeed, precise spatiotemporal orchestration of gene expression is required for proper embryonic development. Recently, the development of high throughput single-cell omics (scO) technologies has generated comprehensive transcriptomic definitions of cell states within the embryo, shedding light on many mechanisms. But, the design of the first scO methods is so that spatial information is lost (Fig. 2.4). In other words, while it is possible to measure the molecular content of every single cell within a given sample, it is not possible to retain their positional information. Even more recently, efforts have been made to alleviate this issue. Novel protocols, namely near single-cell spatial omics (scSO), have been newly developed to keep the spatial localization of the measured cells (Fig. 2.4). As scO methods did when they came out, scSO methods bring new exciting computational challenges. Amongst these are two unavoidable ones :

- **How to incorporate spatial dependency information to the already complex high dimensional analysis ?**
- **How to visualise and browse efficiently such large datasets ?**

In this axis, I propose to develop two computational libraries, namely sc3D and sc3D-viewer, to efficiently handle and analyse (sc3D) and dynamically display and browse (sc3D-viewer) scSO datasets. These two libraries split the axis in two sub-axes :

- **Sub-axis 2.1 : sc3D.** The sc3D library will extend the already existing AnnData WOLF et al. 2018 library, allowing handling of classic single-cell transcriptomic datasets. sc3D will include the organisation of the spatial (position and neighbourhood relationship) information in efficient data structures. sc3D will also contain scSO specific preprocessing and analysis methods.
- **Sub-axis 2.2 : sc3D-viewer.** sc3D-viewer will be based on the sc3D data structure and will allow the display of scSO datasets in 3D. sc3D-viewer will allow the projection, on the fly, of expression and/or co-expression of measured genes or any post-computed metrics.

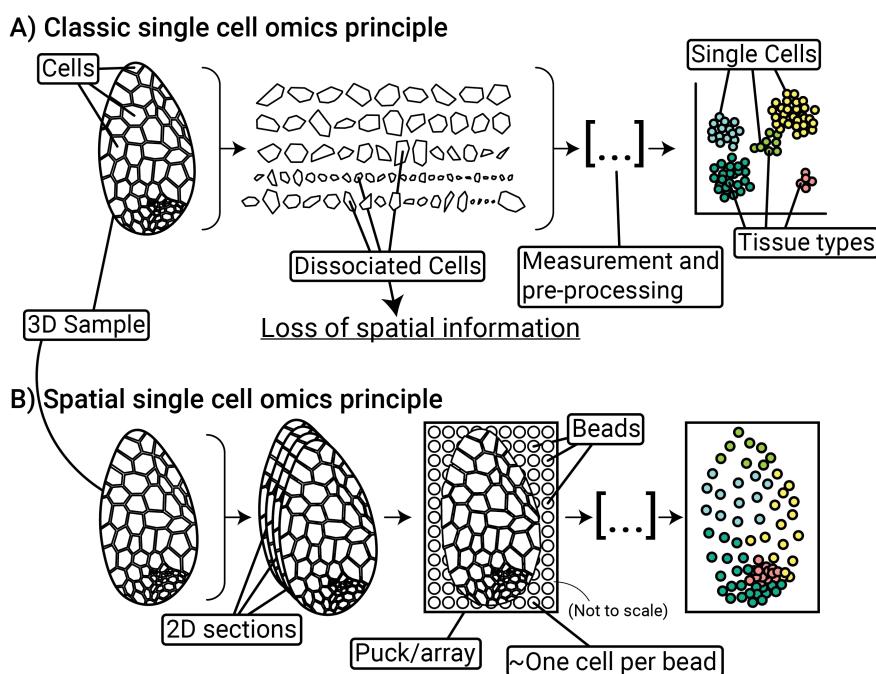


FIGURE 2.4 : Single cell omics principles. A. In classic single cell omics, the cells are dissociated from each other before being recorded. This process clears away the spatial information. As a result, the processed data are often projected in 2 dimensions, in a way that groups cells together according to their molecular content. B. In spatial single cell omics (here I am referring specifically to SlideSeqV2 STICKELS et al. 2021), 2 dimensional sections are made from a 3D sample. These sections are then deposited onto an array of barcoded beads of the size of a single cell. Through gravity, each cell of the 2D slice falls onto a bead and is barcoded according to that bead. This primary barcode allows us to record the 2D position of the cell and therefore to replace its measurement in space a-posteriori.

2.4.1 sc3D

sc3D will address one of the primary challenges brought by scSO : how to store relevant spatial information efficiently. As a first step and the way it is usually done, spatial information is stored as if it were a classic measurement : three values representing the x, y, z coordinates of the position of the measured beads. While memory efficient, it does not directly encompass the neighbouring information between beads and therefore prevents the use of many computational methods designed for spatial datasets. To resolve this problem, I propose to extend the state-of-the-art data structure for Python (AnnData WOLF et al. 2018) to add a network representation of the beads where the nodes of the networks are the beads and the edges of the network are the neighbourhood relationships between beads. The neighbourhood relationship will be computed as the Gabriel graph of the spatialised beads similarly to McDOLE et al. 2018 where it was proven to be a representative neighbourhood reconstruction while keeping a low computational time complexity.

Another challenge that has arisen for scSO methods is the reconstruction, in 3D, of the datasets. Indeed, scSO datasets are usually acquired as a set of 2 dimensional stacks. However, the spatial coherence is often not kept from slice to slice. Therefore algorithms such as Tangram BIANCALANI et al. 2021, STUtility BERGENSTRÄHLE et al. 2020 or more recently PASTE ZEIRA et al. 2022 have been developed. Here I propose to incorporate the previously developed algorithms to sc3D so that users can choose amongst all the different methods. I will also develop a new algorithm based on pre-computed cluster alignments. Cell clusters can be computed using classic methods implemented in the reference libraries for scO : Seurat HAO et al. 2021 or scanpy WOLF et al. 2018 for example. When correctly parameterized, these clusters represent the different tissues or sub-tissues. I will take advantage of the fact that most tissues are continuous in space and use that property to align the different slices. By going from the single cell/bead scale to the tissue scale, our algorithm will be significantly faster than the already existing algorithms.

Because of the spatial nature of the datasets, some already existing scO methods need to be adapted and many new methods can be developed. I will translate the already existing scO methods to incorporate the spatial component. One example of such methods is detecting spatially differentially expressed genes ((S)DEG). I will develop novel approaches to integrate spatial information into the detection of DEG in space. Our SDEG detection method will be based on creating subnetworks of expressing cells and on the property that quasi-linearly links the density of a network and a random subnetwork of that network. On top of our new method to detect SDEG, I will also incorporate state of the art methods such as Splotch ÄIJÖ et al. 2019 or C-SIDE CABLE et al. 2022.

Finally, I will develop a new set of methods to build statistical atlases of scSO from a set of samples. Thanks to the spatial nature of the dataset, I will be able to take

advantage of my expertise in point cloud registration to develop methods to register scSO datasets against each other and therefore allow their aggregation into an average representation, building the first single cell omics 3D statistical atlases. This average single-cell spatial atlas of the species or tissue recorded will enable the statistical comparison between wild-type/healthy samples and perturbed/sick samples allowing us to statistically measure the gene expression deviation of an abnormal sample to a normal one.

2.4.2 sc3D-viewer

Extracting biologically relevant information from such complex and high dimensional datasets is challenging. Exploring the datasets is the first step to orienting questions and helping experimentalists design experiments. Such exploration has to be easy and seamless.

To that aim, I will develop sc3D-viewer : a 3D viewer for scSO based on napari, a software initially dedicated to 3D visualisation of microscopy images. sc3D-viewer will be both a way of exploring scSO datasets but also a graphical interface for sc3D to allow users to run the tools and methods proposed in sc3D without having to write code.

Our viewer will aim not only to display the dataset in 3D but also to be intuitively interactive. Among others, the viewer will allow the user to choose tissues or clusters to display or hide. It will also allow the user to choose to project on the cells any measured gene expression or metrics such as the pseudo-time for example. The user will also be able to have a dynamic interaction between the 3D view of their datasets and the more common 2D projections such as UMAPs and tSNEs representations. sc3D-viewer will allow the user to select cells from any 2D projection and display them on the fly in the 3D viewer. I will also enable the user to display multiple genes simultaneously to explore gene co-expression. Finally, I will develop methods to compute meshes from any set of points to visualise the different tissues or classes better.

As I have already developed a prototype : [napari-sc3D-viewer](#), I know that it will be possible to indeed have the rapid, seamless and in depth user experience that is required to explore such dense and large datasets.

2.4.3 Usefulness to the biological community in general

I have already started to develop these tools in collaboration with the groups of F. Chen at the Broad Institute, Cambridge, MA, USA and A. Meissner at the Max Planck Institute for Molecular Genetics, Berlin, Germany. The group of A. Meissner provided the samples, the group of F. Chen developed the method to measure the cell molecular content while keeping the spatial information and I developed the computational

tools to reconstruct, analyse and visualise the generated datasets (see the document about the previous works for more details) SAMPATH KUMA et al. 2023.

For this project, I built the first 3D molecular atlas of a mouse embryo at 8.5 and 9.0 days after fertilisation (Fig. 2.5). The atlas is a powerful tool for biologists who can perform in-silico 3D virtual in-situs for 27000 genes and further look in depth at tissues and genes of interest to quantify expression. Furthermore, biologists can systematically look for which genes are locally expressed in their tissues of interest. Currently developed for mouse data, expanding these methods to other organisms will be a valuable asset to an even larger community.

Finally, the tools developed for the mouse datasets from A. Meissner's group will be extended to the datasets that will later be acquired by the group of A. Pavlopoulos with whom I have an ongoing collaboration.

2.5 Morpholecular atlases

Axes 1 and 2 will produce extremely rich single-cell scale datasets of two different modalities : morphological and molecular. By themselves, these datasets have led to significant biological discoveries, however, they are not sufficient yet to reflect the complex synergy between morphology and molecular. In order to understand, at the single-cell scale, the interplays between the molecular content of a cell, group of cells or tissues and their morphodynamics, the next logical step is to integrate these two modalities together, morphological and molecular, into a common atlas. The goal of this third axis is to **create the very first single cell atlas of cell morphodynamics together with their molecular content : a morpholecular atlas**.

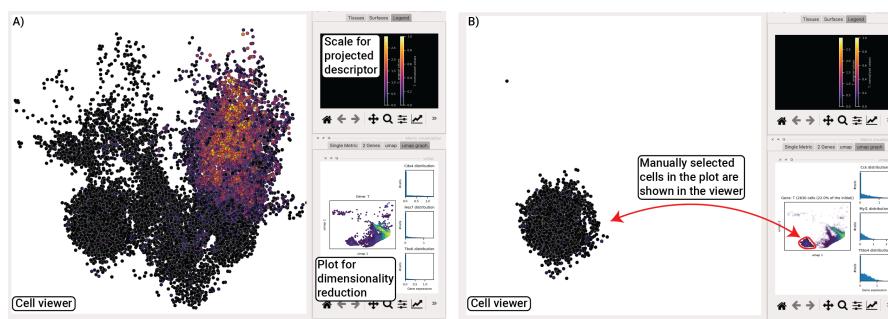


FIGURE 2.5 : Single-cell spatial transcriptomics of an E8.5 mouse embryo. A. Prototype of the 3D viewer for spatial single-cell datasets. Brachyury is projected on the cells. B. Example of an interaction where the user selected cells in the 2D projection on the left and the same cells are displayed on the 3D viewer.

The morpholecular atlas, together with the adequate visualisation tools will provide biologists with an integrated view of the development of their favourite organisms. If they are interested in the specific cell movement of a given tissue, they will be able systematically look for genes that are specifically expressed at the spatio-temporal position of their interest. If they are interested in a specific gene, they will be able to systematically look at which morphological events are correlated with the specific expression of these genes.

Building such morpholecular atlases comes with challenges, the main ones are likely the following :

1. The two modalities will probably come from different samples, therefore, **how to map two samples, from different modalities, onto each other?**
2. The recording of the single cell molecular modality is lethal to the sample, so it is only possible to record fixed snapshots, therefore, **how to propagate the static molecular information in time?**
3. The construction of a morpholecular atlas will generate one of a kind datasets, **how to validate the quality of the atlas?**
4. Such morpholecular atlases are extremely rich and span many dimensions, therefore, **how to easily and seamlessly browse and query such complex datasets?**

To answer these questions I will develop my first prototype of the morpholecular atlas on the epidermis of the *Parhyale hawaiensis* embryo working in collaboration with A. Pavlopoulos who will provide the samples. *P. hawaiensis* is the ideal model for developing such a prototype since its morphological complexity is in between the complex one of the mouse and the simpler ones of highly stereotyped species such as *C. elegans*. Moreover, while it is still an emergent model system, many tools have been developed for genetic manipulation and imaging of the embryos.

Such tools already exist and will allow the recordings of *P. hawaiensis* to be done for both modalities in a single sample. The embryos will be imaged using a light-sheet fluorescence microscope, then, at the end of the recording of the cell morphodynamics, the sample will be frozen and the spatial single cell RNA sequencing will be made with SlideSeqV2 STICKELS et al. 2021. In doing so, I will be able to overlay both datasets from the same embryo at the same developmental time point.

This approach of combining imaging and SlideSeqV2 [ibid.](#) will be performed up to different timepoints in order to acquire the co-recordings of the spatial single cell transcriptomics and the cell dynamics at different developmental times. These multimodal recordings will form a ground truth against which I will quantify the quality of the reconstruction of my methods. Using these recordings will also allow the validation of the methods to propagate the molecular information in space.

2 Future work – 2.5 Morpholecular atlases

Using these morpholecular atlases, we will address the question of the bilateral symmetry of *P. hawaiensis*. It is known that *P. hawaiensis* can recover their bilateral symmetry after specific drastic cell ablations, however, the mechanisms driving this recovery are not yet known. By comparing the wild-type morpholecular atlases to the measurements of ablated embryos, we will find what genes are responsible for the recovery of the bilateral symmetry.

Finally, I will extend the viewer developed for the Axis 2 in order to integrate spatio-temporal information and images. To do so, I will take advantage of the fact that the software basis of the sc3D-viewer is napari, a software that was initially designed for 3D+time image visualisation.

Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Bibliographie

- [1] T. ÄIJÖ, S. MANIATIS, S. VICKOVIC et al., “Splotch : Robust estimation of aligned spatial temporal gene expression data”, *bioRxiv*, 2019. DOI : [10.1101/757096](https://doi.org/10.1101/757096). eprint : <https://www.biorxiv.org/content/early/2019/09/05/757096.full.pdf>. adresse : <https://www.biorxiv.org/content/early/2019/09/05/757096> (cf. p. 32).
- [2] S. F. ALTSCHUL, W. GISH, W. MILLER et al., “Basic local alignment search tool”, *Journal of Molecular Biology*, t. 215, p. 403-410, 3 1990. DOI : [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2). adresse : [http://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) (cf. p. 27).
- [3] F. AMAT, W. LEMON, D. P. MOSSING et al., “Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data”, en, *Nat Methods*, t. 11, n° 9, p. 951-958, sept. 2014, ISSN : 1548-7105. DOI : [10.1038/nmeth.3036](https://doi.org/10.1038/nmeth.3036). (visité le 07/11/2019) (cf. p. 28).
- [4] J. BERGENSTRÄHLE, L. LARSSON et J. LUNDEBERG, “Seamless integration of image and molecular analysis for spatial transcriptomics workflows”, *BMC Genomics*, t. 21, p. 78, 1 2020. DOI : [10.1186/s12864-020-06832-3](https://doi.org/10.1186/s12864-020-06832-3). adresse : [http://doi.org/10.1186/s12864-020-06832-3](https://doi.org/10.1186/s12864-020-06832-3) (cf. p. 32).
- [5] T. BIANCALANI, G. SCALIA, L. BUFFONI et al., “Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram”, *Nat Methods*, t. 18, p. 1352-1362, 11 2021. DOI : [10.1038/s41592-021-01264-7](https://doi.org/10.1038/s41592-021-01264-7). adresse : [http://doi.org/10.1038/s41592-021-01264-7](https://doi.org/10.1038/s41592-021-01264-7) (cf. p. 32).
- [6] S. C. van den BRINK, P. BAILLIE-JOHNSON, T. BALAYO et al., “Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells”, *Development*, t. 141, p. 4231-4242, 22 2014. DOI : [10.1242/dev.113001](https://doi.org/10.1242/dev.113001). adresse : [http://doi.org/10.1242/dev.113001](https://doi.org/10.1242/dev.113001) (cf. p. 29).
- [7] D. M. CABLE, E. MURRAY, V. SHANMUGAM et al., “Cell type-specific inference of differential expression in spatial transcriptomics”, *Nat Methods*, t. 19, p. 567-5751076-1087, 9 2022. DOI : [10.1038/s41592-022-01575-3](https://doi.org/10.1038/s41592-022-01575-3). adresse : [http://doi.org/10.1038/s41592-022-01575-3](https://doi.org/10.1038/s41592-022-01575-3) (cf. p. 32).
- [8] S. CRESPI, “Video : 2018’s Breakthrough of the Year and runners-up”, *Science*, 2018. DOI : [10.1126/science.aaw4480](https://doi.org/10.1126/science.aaw4480). adresse : [http://doi.org/10.1126/science.aaw4480](https://doi.org/10.1126/science.aaw4480) (cf. p. 8, 22).

- [9] G. GENIKHOVICH et U. TECHNAU, “On the evolution of bilaterality”, Development, t. 144, p. 3392-3404, 19 2017. DOI : [10.1242/dev.141507](https://doi.org/10.1242/dev.141507). adresse : <http://doi.org/10.1242/dev.141507> (cf. p. 26).
- [10] D. T. GRIMES, “Making and breaking symmetry in development, growth and disease”, Development, t. 146, dev175976, 16 2019. DOI : [10.1242/dev.170985](https://doi.org/10.1242/dev.170985). adresse : <http://doi.org/10.1242/dev.170985> (cf. p. 26).
- [11] Y. HAO, S. HAO, E. ANDERSEN-NISSEN et al., “Integrated analysis of multimodal single-cell data”, Cell, t. 184, 3573-3587.e29, 13 2021. DOI : [10.1016/j.cell.2021.04.048](https://doi.org/10.1016/j.cell.2021.04.048). adresse : <http://doi.org/10.1016/j.cell.2021.04.048> (cf. p. 32).
- [12] P. HIRSCH, L. EPSTEIN et L. GUIGNARD, “Chapter 20 - Mathematical and bioinformatic tools for cell tracking”, in Cell Movement in Health and Disease, M. SCHNOOR, L.-M. YIN et S. X. SUN, éd., Academic Press, 2022, p. 341-361, ISBN : 978-0-323-90195-6. DOI : <https://doi.org/10.1016/B978-0-323-90195-6.00013-9>. adresse : <https://www.sciencedirect.com/science/article/pii/B9780323901956000139> (cf. p. 12).
- [13] J. HUISKEN, J. SWOGER, F. DEL BENE et al., “Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy”, Science, t. 305, p. 1007-1009, 5686 2004. DOI : [10.1126/science.1100035](https://doi.org/10.1126/science.1100035). adresse : <http://doi.org/10.1126/science.1100035> (cf. p. 25).
- [14] P. J. KELLER, A. D. SCHMIDT, J. WITTBRODT et al., “Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy”, Science, t. 322, p. 1065-1069, 5904 2008. DOI : [10.1126/science.1162493](https://doi.org/10.1126/science.1162493). adresse : <http://doi.org/10.1126/science.1162493> (cf. p. 25).
- [15] U. KRZIC, S. GUNTHER, T. E. SAUNDERS et al., “Multiview light-sheet microscope for rapid *in toto* imaging”, Nat Methods, t. 9, p. 730-733, 7 2012. DOI : [10.1038/nmeth.2064](https://doi.org/10.1038/nmeth.2064). adresse : <http://doi.org/10.1038/nmeth.2064> (cf. p. 25).
- [16] L. GUIGNARD, U.-M. FIÚZA, B. LEGGIO et al., “Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis.”, eng, Science (New York, N.Y.), t. 369, 6500 juill. 2020 (cf. p. 12, 14, 20, 25, 27).
- [17] L. GUIGNARD, C. GODIN, U.-M. FIUZA et al., “Spatio-temporal registration of embryo images”, in 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), IEEE, 2014, p. 778-781 (cf. p. 15).
- [18] X. LI, Z. ZHAO, W. XU et al., “Systems Properties and Spatiotemporal Regulation of Cell Position Variability during Embryogenesis”, Cell Reports, t. 26, 313-321.e7, 2 2019. DOI : [10.1016/j.celrep.2018.12.052](https://doi.org/10.1016/j.celrep.2018.12.052). adresse : <http://doi.org/10.1016/j.celrep.2018.12.052> (cf. p. 25).
- [19] C. MALIN-MAYOR, P. HIRSCH, L. GUIGNARD et al., “Automated reconstruction of whole-embryo cell lineages by learning from sparse annotations”, Nat Biotechnol, 2022. DOI : [10.1038/s41587-022-01427-7](https://doi.org/10.1038/s41587-022-01427-7). adresse : <https://doi.org/10.1038/s41587-022-01427-7> (cf. p. 12, 18).

- [20] K. McDOLE, L. GUIGNARD, F. AMAT et al., “In toto imaging and reconstruction of post-implantation mouse development at the single-cell level”, Cell, t. 175, n° 3, p. 859-876, 2018 (cf. p. 12, 15, 16, 20, 28, 32).
- [21] G. d. MEDEIROS, N. NORLIN, S. GUNTHER et al., “Confocal multiview light-sheet microscopy”, Nat Commun, t. 6, p. 603, 1 2015. DOI : [10.1038/ncomms9881](https://doi.org/10.1038/ncomms9881). adresse : <http://doi.org/10.1038/ncomms9881> (cf. p. 25).
- [22] “Method of the Year 2013”, Nat Methods, t. 11, p. 1-1, 1 2014. DOI : [10.1038/nmeth.2801](https://doi.org/10.1038/nmeth.2801). adresse : <http://doi.org/10.1038/nmeth.2801> (cf. p. 8, 22).
- [23] “Method of the Year 2014”, Nat Methods, t. 12, p. 1-1, 1 2015. DOI : [10.1038/nmeth.3251](https://doi.org/10.1038/nmeth.3251). adresse : <http://doi.org/10.1038/nmeth.3251> (cf. p. 8, 22).
- [24] “Method of the Year 2020 : spatially resolved transcriptomics”, Nat Methods, t. 18, p. 1-1, 1 2021. DOI : [10.1038/s41592-020-01042-x](https://doi.org/10.1038/s41592-020-01042-x). adresse : <http://doi.org/10.1038/s41592-020-01042-x> (cf. p. 8, 22).
- [25] G. MICHELIN, L. GUIGNARD, U.-M. FIUZA, P. LEMAIRE et al., “Cell pairings for ascidian embryo registration”, in 2015 IEEE 12th International Symposium on Biomedical Imaging (IEEE, 2015, p. 298-301 (cf. p. 15)).
- [26] G. MICHELIN, L. GUIGNARD, U.-M. FIUZA et G. MALANDAIN, “Embryo cell membranes reconstruction by tensor voting”, in 2014 IEEE 11th International Symposium on Biomedicine (IEEE, 2014, p. 1259-1262 (cf. p. 15)).
- [27] L. A. ROYER, W. C. LEMON, R. K. CHHETRI et al., “Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms”, Nat Biotechnol, t. 34, p. 1267-1278, 12 2016. DOI : [10.1038/nbt.3708](https://doi.org/10.1038/nbt.3708). adresse : <http://doi.org/10.1038/nbt.3708> (cf. p. 28).
- [28] A. SAMPATH KUMA, L. TIAN, A. BOLONDI et al., “Spatial transcriptomic maps of whole mouse embryos”, Nat Genetics, n° 55, p. 1176-1185, 2023. DOI : [10.1038/s41588-023-01435-6](https://doi.org/10.1038/s41588-023-01435-6) (cf. p. 12, 17, 18, 20, 34).
- [29] R. R. STICKELS, E. MURRAY, P. KUMAR et al., “Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2.”, eng, Nat Biotechnol, t. 39, p. 313-319, 3 mars 2021 (cf. p. 31, 35).
- [30] J. V. VEENVLIET, A. BOLONDI, H. KRETZMER et al., “Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites”, Science, t. 370, n° 6522, eaba4937, 2020. DOI : [10.1126/science.aba4937](https://doi.org/10.1126/science.aba4937). eprint : <https://www.science.org/doi/pdf/10.1126/science.aba4937>. adresse : <https://www.science.org/doi/abs/10.1126/science.aba4937> (cf. p. 12, 19).
- [31] J. B. WALLINGFORD, “The 200-year effort to see the embryo”, Science, t. 365, p. 758-759, 6455 2019. DOI : [10.1126/science.aaw7565](https://doi.org/10.1126/science.aaw7565). adresse : <http://doi.org/10.1126/science.aaw7565> (cf. p. 25).

Bibliographie

- [32] Y. WAN, Z. WEI, L. L. LOOGER et al., “Single-Cell Reconstruction of Emerging Population Activity in an Entire Developing Circuit”, English, *Cell*, t. 179, n° 2, 355-372.e23, oct. 2019, ISSN : 0092-8674, 1097-4172. DOI : [10.1016/j.cell.2019.08.039](https://doi.org/10.1016/j.cell.2019.08.039). (visité le 13/01/2020) (cf. p. 19).
- [33] M. WEIGERT, U. SCHMIDT, R. HAASE et al., “Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy”, in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2020, p. 3655-3662. DOI : [10.1109/WACV45572.2020.9093435](https://doi.org/10.1109/WACV45572.2020.9093435) (cf. p. 30).
- [34] F. A. WOLF, P. ANGERER et F. J. THEIS, “SCANPY : large-scale single-cell gene expression data analysis”, *Genome Biol.*, t. 19, p. 1145, 1 2018. DOI : [10.1186/s13059-017-1382-0](https://doi.org/10.1186/s13059-017-1382-0). adresse : <http://doi.org/10.1186/s13059-017-1382-0> (cf. p. 31, 32).
- [35] C. WOLFF, J.-Y. TINEVEZ, T. PIETZSCH et al., “Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb”, *eLife*, t. 7, e34410, 2018 (cf. p. 12, 18, 25, 26).
- [36] L. WOLPERT et A. D. LANDER, “Arms and the Man : The Problem of Symmetric Growth”, *PLoS Biol.*, t. 8, e1000477, 9 2010. DOI : [10.1371/journal.pbio.1000477](https://doi.org/10.1371/journal.pbio.1000477). adresse : <http://doi.org/10.1371/journal.pbio.1000477> (cf. p. 26).
- [37] R. ZEIRA, M. LAND, A. STRZALKOWSKI et al., “Alignment and integration of spatial transcriptomics data”, *Nat Methods*, t. 19, p. 567-575, 5 2022. DOI : [10.1038/s41592-022-01459-6](https://doi.org/10.1038/s41592-022-01459-6). adresse : <http://doi.org/10.1038/s41592-022-01459-6> (cf. p. 32).
- [38] K. ZHANG, “A constrained edit distance between unordered labeled trees”, *Algorithmica*, t. 15, p. 205-222, 3 1996. DOI : [10.1007/BF01975866](https://doi.org/10.1007/BF01975866). adresse : <http://doi.org/10.1007/BF01975866> (cf. p. 27).