

Inferring cell-cell interactions from quantitative analysis of microscopy images

Gustave Ronteix

▶ To cite this version:

Gustave Ronteix. Inferring cell-cell interactions from quantitative analysis of microscopy images. Imaging. Institut Polytechnique de Paris, 2021. English. NNT: 2021IPPAX111. tel-03517974v2

HAL Id: tel-03517974 https://hal.science/tel-03517974v2

Submitted on 14 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





Inferring cell-cell interactions from quantitative analysis of microscopy images

Thèse de doctorat de l'Institut Polytechnique de Paris préparée à l'École polytechnique

École doctorale n° 626 École Doctorale de l'Institut Polytechnique de Paris (ED IP PARIS) Spécialité de doctorat: Biologie

Thèse présentée et soutenue à Paris, le 15/12/2021, par

Gustave Ronteix

Composition du Jury :

Jean-François Joanny UMR 168, Professeur, Collège de France Président Dagmar Iber CoBi, Professeure des Universités, ETH Rapportrice Raphael Voituriez UMR 8237, Directeur de Recherches, Sorbonne Université Rapporteur Danijela Vignjevic UMR 144, Directrice de Recherches, Institut Curie Examinatrice Philippe Bousso INSERM U668, Directeur de Recherches, Institut Pasteur Examinateur Charles Baroud UMR 7646, Professeur, École polytechnique Directeur de thèse Pour Titou.

Résumé

Les systèmes biologiques sont bien plus que la somme de leurs constituants. En effet, ils sont souvent caractérisés par des comportements macroscopiques complexes issus de la communication et des interactions entre leurs parties prenantes. Par exemple, la régulation et le rejet éventuel des tumeurs par le système immunitaire sont le résultat de multiples boucles de rétroaction positives et négatives, influençant à la fois le comportement des cellules cancéreuses et immunitaires. Cela rend l'étude de ces phénomènes particulièrement complexe in-vivo, où il est impossible d'isoler les différents composants et où l'observation est limitée par les contraintes expérimentales. L'utilisation de modèles in-vitro permet de modeler certains phénomènes en facilitant leur observation au prix d'une pertinence plus limitée des observations qui en découlent. Un défi majeur est donc de parvenir à simuler des systèmes biologiques de la façon la plus fiable possible dans des dispositifs expérimentaux in-vitro.

Pour imiter les comportements complexes entre cellules immunitaires et cellules tumorales in-vitro, j'ai conçu un test microfluidique permettant de confronter des sphéroïdes de tumeurs de mélanome à de multiples cellules T et d'observer les interactions qui en résultent avec une haute résolution spatio-temporelle sur de longues périodes de temps. En utilisant des méthodes d'analyse d'images combinées à de la modélisation mathématique, je démontre qu'une boucle de rétroaction positive conduit à l'accumulation de cellules T sur la tumeur, conduisant à une fragmentation accrue des sphéroïdes. Cette étude met en lumière l'initiation de la réponse immunitaire à l'échelle d'une seule cellule : elle montre que même le tout premier contact entre une cellule T et un sphéroïde tumoral augmente la probabilité que la cellule T suivante arrive sur la tumeur. Elle montre également qu'il est possible de récapituler des comportements antagonistes complexes in-vitro, ce qui ouvre la voie à l'élaboration de protocoles plus sophistiqués, impliquant par exemple un micro-environnement tumoral plus complexe.

De nombreux processus biologiques sont le résultat d'interactions complexes entre les types de cellules, en particulier au cours du développement. Le foie fœtal est le lieu de la maturation et de l'expansion du système hématopoïétique, mais on sait peu de choses sur sa structure et son organisation. De nouveaux protocoles expérimentaux ont été récemment mis au point pour imager cet organe et j'ai développé des outils pour interpréter et quantifier ces données, permettant la construction d'un "jumeau réseau" de chaque foie fœtal. Cette méthode permet de combiner l'échelle unicellulaire et l'échelle de l'organe dans l'analyse, révélant l'accumulation de cellules myéloïdes autour des vaisseaux sanguins irriguant le foie fœtal aux derniers stades du développement de l'organe. À l'avenir, cette technique permettra d'analyser précisément les niches environnementales de manière quantitative. Ceci pourrait à son tour nous aider à comprendre les étapes du développement et la maturation types cellulaires cruciaux tels que les cellules souches hématopoïétiques. Les interactions entre les bactéries et leur environnement sont essentielles pour comprendre l'émergence de comportements collectifs complexes tels que la formation de biofilms. Un mécanisme d'intérêt est celui de la rhéotaxie, par lequel le mouvement bactérien est entraîné par les gradients des contraintes de cisaillement du fluide dans lequel les cellules se déplacent. J'ai développé un cadre pour calculer les équations semianalytiques guidant le mouvement des bactéries dans les contraintes de cisaillement. Ces équations prédisent des comportements qui sont démenties par les observations expérimentales, mais la divergence est résolue une fois que la diffusion rotationnelle est prise en compte. Les résultats expérimentaux correspondent bien à la prédiction théorique : les bactéries dans les gouttelettes se séparent de manière asymétrique lorsqu'un cisaillement est généré dans le milieu.

Abstract

Biological systems are much more than the sum of their constituents. Indeed, they are often characterized by complex macroscopic behaviors stemming from communication and interactions. For instance, the regulation and eventual rejection of tumors by the immune system is the result of multiple positive and negative regulation networks, influencing both the behavior of the cancerous and immune cells. To mimic these complex effects in-vitro, I designed a microfluidic assay to challenge melanoma tumor spheroids with multiple T cells and observe the resulting interactions with high spatiotemporal resolution over long periods of time. Using advanced image analysis combined with mathematical modeling I demonstrate that a positive feedback loop drives T cell accumulation to the tumor site, leading to enhanced spheroid fragmentation. This study sheds light on the initiation if the immune response at the single cell scale: showing that even the very first contact between T cell and tumor spheroid increases the probability of the next T cell to come to the tumor. It also shows that it is possible to recapitulate complex antagonistic behaviors in-vitro, which paves the way for the elaboration of more sophisticated protocols, involving for example a more complex tumor micro-environment.

Many biological processes are the result of complex interactions between cell types, particularly so during development. The fetal liver is the locus of the maturation and expansion of the hematopoietic system, yet little is known about its structure and organization. New experimental protocols have been recently developed to image this organ and I developed tools to interpret and quantify these data, enabling the construction of a "network twin" of each fetal liver. This method makes it possible to combine the single-cell scale and the organ scale in the analysis, revealing the accumulation of myeloid cells around the blood vessels irrigating the fetal liver at the final stages of organ development. In the future, this technique will make it possible to analyze precisely the environmental niches of cell types of interest in a quantitative manner. This in turn could help us understand the developmental steps of crucial cell types such as hematopoietic stem cells.

The interactions between bacteria and their environment is key to understanding the emergence of complex collective behaviors such a biofilm formation. One mechanism of interest is that of rheotaxis, whereby bacterial motion is driven by gradients in the shear stress of the fluid the cells are moving in. I developed a framework to calculate the semi-analytical equations guiding bacteria movement in shear stress. These equations predict behaviors that aren't observed experimentally, but the discrepancy is solved once rotational diffusion is taken into account. Experimental results are well-fitted by the theoretical prediction: bacteria in droplets segregate asymmetrically when a shear is generated in the media.

Contents

1	Introduction			
	1.1	Thesis overview	13	
		1.1.1 Context	13	
		1.1.2 Goal and rationale	13	
		1.1.3 Structure	14	
	1.2	Immuno-oncology on microfluidic chips	14	
	1.3	Network-based tissue analysis	15	
	1.4	Rheotactic drift in microfluidic droplets	16	
Ι	Im	muno-Oncology on chips	19	
2	Imr	nune-cancer interactions	21	
	2.1	A short introduction to immuno-oncology	21	
		2.1.1 The immune system plays a key role in controlling tumor growth	21	
		2.1.2 Introducing the innate immune system	21	
		2.1.3 The adaptive immune system	23	
		2.1.4 The tumor micro-environment modulates the immune response .	25	
		2.1.5 CTL migration: search strategies and motion	26	
	2.2	Microfluidic chips can replicate in-vivo immuno-oncology patterns $\ . \ .$	27	
3	Dro	pplet-based assays for 3D spheroid culture	31	
	3.1	From surface tension to droplet generation	32	
		3.1.1 The cost of an interface	32	
		3.1.2 Generating droplets	33	
		3.1.3 Guiding and anchoring microfluidic droplets	34	
	3.2	Droplet-made tumor spheroids	35	
		3.2.1 Spheroids as 3D cell culture models	35	
		3.2.2 Making spheroids in microfluidic droplets	35	
		3.2.3 Iterative droplet trapping and fusion	36	
		3.2.4 Building complex micro-environments in microfluidic droplets .	37	
		3.2.5 Droplet microfluidics in immuno-oncology studies	37	
4	Experimental protocols for immuno-oncology assays using droplet mi-			
	crof	fluidics	41	
	4.1	Introduction	41	
	4.2	Model tumor formation	42	
		4.2.1 Constructing a relevant tumour model	42	

	4.3 4.4	 4.2.2 Incorporating Matrigel in droplets	43 45 47 47 47 48
5	CTI 5.1 5.2 5.3	 L accumulation on spheroids in microfluidic droplets CTL migration in micro-device reproduces <i>in vivo</i> behavior	 51 51 53 53 53 54 55 57 58
6	A q drop 6.1 6.2 6.3	quantitative description of specific tumoroid killing by CTLs in pletsKilling of B16 cells by CTLs is heterogeneousStatistical observations on the spheroid fragmentation efficiencyTumor spheroid killing involves collective effects6.3.1Qualitative indications of collective killing6.3.2Estimating the collective fragmentation rate Γ_{frag} 6.3.3Probabilistic modeling of the independent killing of B16 cells on the spheroid by CTLsLong and short-range interactions combine to determine probability of CTLs to kill a tumor spheroid6.4.1Simulating the spheroid fate6.4.2Comparison with experimental results	61 63 64 65 67 68 68 68 69
7	Con	aclusions	71
Π	Ν	etwork-based tissue analysis	73
8	Intr 8.1 8.2 8.3	oduction Networks in biology Spatial networks in biology Why should a biologist work on spatial networks?	75 75 76 77
9	Bui 9.1 9.2 9.3 9.4	Iding Griottes A short introduction to graphs Motivation Program pipeline From image to network	79 79 80 81 83

		9.4.1	Extracting single-cell information from the image \ldots	83
		9.4.2	Building the network	86
		9.4.3	Comparing network construction methods	87
		9.4.4	Edge weighting by contact area	89
	9.5	Conclu	$sion \ldots \ldots$	90
10	Stru	ctural	analysis of the fetal liver as a hematopoietic expansio	n
	nich	e duri	ng development	93
	10.1	Introd	uction \ldots	93
	10.2	Ontoge	eny of the hematopoietic system	94
		10.2.1	Known ontogeny of the hematopoietic system	94
		10.2.2	HSC localization in the fetal liver	94
		10.2.3	Spatial description of the fetal liver niche	96
	10.3	Netwo	rk-based structural analysis of the fetal liver niche $\ . \ . \ . \ .$	97
		10.3.1	Imaging of the fetal liver	97
		10.3.2	Project goals	98
		10.3.3	Description of the fetal liver tissue characteristics	99
		10.3.4	Processing pipeline of the fetal liver images	103
	10.4	Using	Hidden Markov Random Fields to characterize tissues	108
		10.4.1	Why use Hidden Markov Random Fields?	108
		10.4.2	From Markov chains to Hidden Markov Random Fields	108
		10.4.3	Presenting the HMRF algorithm	108
		10.4.4	The fetal liver is irrigated by different vessel types	110
	10.5	Netwo	rk-based description of cell niches	113
		10.5.1	Investigating the immediate cell neighborhood	113
		10.5.2	Long-range spatial distribution properties	115
		10.5.3	Border analysis reveals specific spatial organization around fetal liver features	117
	10.6	Quanti	ifying the fetal liver maturation process	120
	10.0	10.6.1	Characterizing the evolution of tissue composition	120
		10.6.2	CD45+/Gr-1+ cells accumulate around fetal liver vessels during	120
		10.00	maturation	122
		10.6.3	The FL mesothelium undergoes significant reorganization be-	
	10 -		tween E12.5 and E18.5 \ldots \ldots \ldots \ldots \ldots \ldots	123
	10.7	Conclu	181011	127
TTI	F	Bacter	ial rheotaxis	129
		lactor		120
11	Sem	i-analy	ytical results on bacterial rheotaxis	131
	11.1	Analyt	cical results – deterministic case	132
		11.1.1	Definition of the problem	132
	11 0	11.1.2	Brief Observations	135
	11.2	Retriev	ving the instantaneous bacterial dynamics	136
		11.2.1	Plugging in the geometric properties	136
		11.2.2	Bacterial rotation in the shear plane θ	136
		11.2.3	Instantaneous dynamics of the pitch angle ψ	137
		11.2.4	Instantaneous dynamics of the position z	138

11.3	Long-term bacterial dynamics	139
	1.3.1 Analytical trajectories	139
	1.3.2 The effect of diffusion \ldots	141
11.4	Experimental results	144
11.5	Summary and outlook	146
11.6	Materials and Methods	148
	1.6.1 Bacteria	148
	1.6.2 Microfluidics \ldots	149
IV C	onclusion	151

12 Conclusions and Future work 1					
	12.1	Immuno-oncology in droplets	153		
	12.2	Building a network representation of tissues	154		
	12.3	A structural description of the fetal liver	155		
	12.4	Rheotactic motion of bacteria in shear flow	155		

Introduction

Chapter 1

Introduction

1.1 Thesis overview

1.1.1 Context

In his prescient article *More is Different*, P. W. Anderson counters the reductionist argument by highlighting the crucial role of emergent properties in science [1]. Reductionists argue that the properties of a large system is the consequence of the physical laws describing its components (in other words, a complex system is the sum of its parts). However, since Anderson's article, it has been shown time and time again that unpredicted behaviors can emerge from well-understood constituents. Complex macroscopic behaviors in living systems often stem from communication and interaction loops between much simpler elements: biological systems are not just more, but different from the sum of their constituents.

There, however, lies the rub: understanding and deciphering the communication patterns between cells in a tissue or bacteria in their medium requires the development of new experimental protocols, new measurement methods and new frameworks to interpret the resulting data. This is the mission of quantitative biology, who use tools borrowed from engineering, computer science and mathematics to understand how biological systems behave.

1.1.2 Goal and rationale

The initial goal for my PhD project was to adapt the microfluidic chip developed for spheroid culture in the lab to be able to challenge tumor spheroids with T cells on it. After demonstrating that the microfluidic platform was suitable for such a study, I decided to investigate the T cell accumulation and killing processes at the single-cell level. Developing the analysis tools to understand the experimental data took a large part of my PhD, and triggered my interest in writing software for the quantitative analysis of biological data.

This became useful in the second part of my PhD when I decided to create a Python toolkit for spatial analysis of biological images. By including spatial information and single cell attributes (phenotype, size etc.) in a single graph object, it allows the user to easily compute statistics with a high level of granularity and highlight previously unnoticed patterns. My final goal was to use this approach to describe the spatial structure of the fetal liver and shed light on its evolution during development.

In-between these projects, I worked on the motility of bacteria in shear flow. Previous experiments conducted in the lab have yielded intriguing, but yet unexplained, behaviors. Here, I proposed a semi-analytic description of the bacterial motion and proposed a model to explain the measurements.

The three projects described above deal with very different biological objects ranging from bacteria to T cells. However, there is a continuous thread binding them together: a deeply interdisciplinary approach and an interesting in connecting singlecell behaviors to population-level outcomes. On a more personal level, I had a lot of fun combining new hardware, protocols, measurements and analysis techniques and this led to the creation of some (hopefully) useful tools.

1.1.3 Structure

The three past years have allowed me to conduct six main tasks. First, the development immuno-oncology protocols using droplet microfluidics, which took me the greater part of my first year. Then came the development of the image analysis pipeline, eventually allowing the characterization of the T cell accumulation and killing process. Fourth, during the first part of my third year, I wrote most of the spatial analysis toolkit, which I have been using in the final months of my PhD to analyze the development of the fetal liver. Finally, the study of the bacterial motion was done during lulls between projects since the beginning of my PhD.

For the purpose of clarity, and given their independence, I have structured this thesis in three main parts. In the first one I discuss the development of the immunooncology assay and the results we get from challenging tumors with T cells in it. Then I describe **Griottes**, my tissue analysis toolkit, and the first results on the fetal liver. Finally, in the third part of my thesis, I analyze the motion of bacteria in shear flow. A short summary of the context and the results from these independent projects is given in the following sections.

1.2 Immuno-oncology on microfluidic chips

The immune system plays an essential role in controlling tumor growth and patient fate. In particular, it is known that T cell concentration in the tumor site is an important predictor of patient response to treatment: the more T cells, the higher their survival rate [2]. Conversely, other immune cell types are associated with negative prognosis such as tumor-associated macrophages (TAMs). Previous work demonstrates that T cell migration to the tumor site is associated with the presence of cytokines. These small molecules, produced by immune cells, trigger phenotypic changes in the tumor microenvironment (TME) [3] and guide the T cells to the tumor site [4]. However, it remains challenging to study the initiation of the immune response *in vivo* for experimental reasons; new experimental protocols are therefore necessary to understand how environmental factors (such as T cell concentration, TME composition etc.) affect the efficacy of the immune response.

A solution to this problem is to use advanced *in vitro* devices to mimic the TME and study the tumor rejection by T cells [5,6]. This allows for prolonged monitoring and precise control of the tumor model, and existing studies have demonstrated that these devices reproduce in vivo behaviors [7,8]. However, in many situations it would be useful to generate a large number of parallel samples in order to increase the experiment statistics or to test multiple conditions in a single experiment. For this, a possible solution is the use of droplet microfluidics, which allow the analysis of multiple samples in parallel, increasing throughput. The Physical Microfluidics and Bioengineering lab at the Pasteur Institute has developed devices and protocols to conduct complex high-throughput measurements linking the single-cell behavior to population properties (dubbed "Multi-scale cytometry") [9]. In order to model the in vivo immune-tumor interactions, I adapted protocols previously developed in the lab for loading and cultivating cells in droplets to the culture of tumor spheroids. Building upon previous work, I developed a method to insert the T cells into the spheroid-containing droplet. This allowed us to image and monitor the T cells and the tumor spheroids in the droplets for prolonged periods of time. In addition to improving the experimental protocols, I developed new image analysis tools to extract information from the measurements done on the chip.

Having access to single-cell spatiotemporal information over a large number of replicas allowed me to extract minute signals from the T cell accumulation process on the spheroids. They show that the arrival rate increases linearly with the number of T cells already on the tumor: the more T cells on the spheroid, the higher the chance that an additional T cell comes and attaches. This is the hallmark of a positive feedback loop. Even more spectacular, the effect was measurable starting from the very first T cell: even for low T cell numbers on the target, at the very start of the initiation of the immune response, we can measure a positive feedback loop driving the enhanced accumulation of T cells on the tumor. Preliminary analysis suggests that the accumulation is driven by the appearance of gradients in the effective diffusion coefficient of the T cells inside the droplet, concentrating the T cells near the spheroid. This is the first characterization of the initiation of the collective immune response at the single-cell scale. Then analyzing the spheroid fragmentation caused by the T cells, I show that T cell killing of target tumor cells is driven by cooperative dynamics, in line with previously data in the literature [10, 11].

This work shows that two collective processes, T cell accumulation and then killing, lie at the core of successful tumor rejection *in vitro* and that we can recapitulate complex immune cell behaviors in microfluidic droplets.

1.3 Network-based tissue analysis

A key question in many studies is that of the spatial variation in biological signals. For example, hepatocyte cells killed by drugs seemed to be mainly distributed on the outer layers of the spheroids [12]. Similarly, previous work conducted in the lab shows that the MSCs organize spatially segregated based on heterogeneities in CD146 receptor expression [13]. The adoption of more capable imaging techniques such as confocal and light sheet microscopy, and new segmentation tools (such as CellPose [14]) have made it possible to image and segment samples with unprecedented resolution. However, frameworks capable of making sense of the large amounts of generated data have not been democratized yet.

The need for such useful frameworks has not only been felt by the lab. Indeed, the emergence of new techniques such as spatial transcriptomics [15] or CODEX imaging [16] has enabled researchers to measure unprecedented amounts of data (gene expression, marker expression etc.) with a high spatial resolution. A natural tool to store the information relative to each detected cell and to represent the connections between them is the graph. Graphs have been widely used in biology, for example to represent interactions between genes. They have also been used to analyze systems constrained by their spatial properties, for instance, the distribution of train stations on a railway network [17]. The strength of a graph (or network) representation lies in its ability to include in a single object both the individual properties of each cell and the relationship between cells themselves.

By building a graph from an imaged tissue, it is possible to generate a "network twin" that represents each cell from the tissue and that would contain all the relevant information (cell type, neighbourhood etc.) necessary for downstream analysis. I - with help from Jérémie Sobel and Valentin Bonnet - wrote a Python library to make this "network twin" from any segmented tissue image. An application for this program called **Griottes** - came after a discussion with Marcia Peixoto and Francisca Soares da Silva from Ana Cumano's Lymphocytes and Immunity lab. They study the fetal liver to understand the development of the hematopoietic system. The cells in such tissue will ultimately compose the blood system: red and white blood cells, plaquettes etc. In the adult, the hematopoietic stem cells (HSCs) are located in the bone marrow, but during development, they expand in the fetal liver (FL). In the long run, characterizing the fetal liver and understanding the factors driving HSC expansion could pave the way for expanding HSCs in the lab, which would be of immense interest for the research towards treating numerous diseases affecting the blood system. Yet, today very little is known about the fetal liver and the specific niche in which HSCs expand.

I constructed the network-representation of the FL from images of tissue sections. This gave us access to the precise environment of up to 90 000 cells per image. The sheer amount of data means that not only can we characterize the tissue composition with FACS-like statistics, but also we can connect the macroscopic tissue structure with the immediate niche of each individual cell. Focusing on markers staining progenitor and then myeloid cells, we show that these sub-populations have unique spatial distributions in the FL and we show how they vary during the maturation of the fetus. This is the first work mapping vast portions of the fetal liver with single-cell precision in 3D. The combination of new sample preparation protocols, imaging pipelines and network-based analysis methods enabled us to describe the tissues in new detail and uncover the organization of the fetal liver at critical development stages.

1.4 Rheotactic drift in microfluidic droplets

The motion of bacteria in aqueous media has long been of interest to physicists. At the scale of a bacterium, inertia terms can be neglected in the Navier-Stokes equation, leading to counter-intuitive behaviors. For instance, a very small swimmer (say, 0.1mm big) beating his feet would move back and forth without any net motion forward: it is the inertia of the fluid displaced by his feet that enable him to propel himself forward. These phenomena was elegantly described by Purcell in his classic study *Life at Low* Reynolds Number [18].

If you place a small elongated object in a shear flow it will undergo periodic rotations in the shear planes. The exact motion of the object is called Jeffery orbit [19], after the mathematician who first described the equation governing this type of movement. An object undergoing Jeffery orbits bobs in and out of the shear plane periodically, but the net motion orthogonally to the shear plane is null. Yet, it has been recently shown that bacteria can have a net motion in the plane orthogonal to the shear plane, a phenomenon that has been named rheotaxis [20]. Numerical simulations of the long-term bacteria motion confirm these experimental observations and reveal that the displacement is driven by shear flow. However the analytical derivation of the rheotactic force hasn't been found yet.

In order to quantify the behavior of bacteria in shear flow, I derived a semi-analytical formulation of the rheotactic force as a function of bacteria parameters. This allowed me to compute long-term, deterministic bacteria trajectories. Intriguingly, these equations predict that the rheotactic force tilts all the bacteria orthogonally to the shear plane, which differs significantly from experimental observations. I show however that the discrepancy can be resolved by considering the effect of diffusion and the Jeffery forces that tend to align the bacteria in the shear plane. Finally, I propose a relationship between the shear that the bacteria is subject to and the rheotactic force, and I confirm this relationship experimentally by looking at the bacterial segregation in the droplets as a function of the oil flow in the microfluidic chip.

This study proposes the first analytical derivation of the rheotactic force exerted on a bacteria in a shear flow and it links the theoretical results with experimental measurements conducted in microfluidic droplets.

Part I Immuno-Oncology on chips

Chapter 2

Immune-cancer interactions

The current part of my PhD dissertation describes the immuno-oncology assay on microfluidic chips that I developed. This introduction will be constituted in a short description of the immune system, then it will describe the nascent field of microfluidics and how we used this technology to grow tumours in droplets and challenge them with T cells.

2.1 A short introduction to immuno-oncology

2.1.1 The immune system plays a key role in controlling tumor growth

The role of immune system in regulating cancer was revealed already in the late 1800s by William Coley when he observed tumor remission in patients with bacterial infections [6]. In the beginning of the 20th century, Paul Ehrlich showed that the immune system could detect and destroy tumors in patients [21]. Half a century later, in 1957, MacFarlane Burnet developed the theory of immunosurveillance which states that immune cells constantly patrol the organism to destroy developing tumors [22]. Today, cancer is routinely treated with approaches designed to harness the power of the immune system such as immune checkpoint inhibitors or CAR-T therapies. These treatments have enabled us to make significant steps forward in treating cancer, even bagging the Nobel prize in Physiology or Medicine for Tasuku Honjo and James Allison for their discovery of the programmed death molecule-1 (PD-1) and the cytotoxic T-lymphocyte antigen-4 (CTLA-4) on T cells.

Before delving into the biology of tumor surveillance, we will briefly introduce the immune system and its main protagonists in the following paragraphs.

2.1.2 Introducing the innate immune system

A wide range of cell types belong to the immune system, each one of them occupying a specific function to protect the organism. These cells can be classified into two groups: the innate or the adaptive immunity (Fig. 2.1). These two categories can be distinguished by their reaction time, with innate immune cells quickly reacting to new threats in an unspecific manner, whereas the defense conducted by adaptive immune



FIGURE 2.1: Description of the main cell types composing the immune system and their role. Taken from McComb et al. [23]

cells is delayed in time, though specific and can be preserved in the immunological memory.

The innate cell type population is mainly composed of granulocytes (for more information on the maturation process of granulocytes we refer the reader to the section 10.6.2), mast cells, neutrophils, macrophages and dendritic cells. The rapid response can be important in initiating inflammation and wound healing and is crucial in the early responders to parasites, extracellular bacteria, and tumors. For instance it is the arrival of granulocytes during infection that induces the symptoms of wound inflammation and dilation of the surrounding blood vessels that we all have experienced [23]. Neutrophils and macrophages are adept phagocytic cells which are capable of ingesting and destroying invading microbes. Unlike neutrophils, which are bloodresident and short-lived, macrophages take up residence in all tissues of the body and are relatively long-lived. Macrophages and neutrophiles produce cytokines that draw in and activate other immune cells to the site of infection. Dendritic cells (DCs) are also capable of internalizing and destroying invading microbes through phagocytosis, however they are better known for their important role of activating the cells of the adaptive immune system. Both DCs and macrophages are often called antigen presenting cells (APCs) due to their efficient ability to internalize pathogens and present pathogen peptides on their cell surface. They play an essential role in warning the immune system of a possible infection.

2.1.3 The adaptive immune system

The adaptive immune system is composed of T and B lymphocytes and is characterized by its high specificity: it is primed towards pathogens that the organism has already encountered. B cells carry out antibody responses, secreting proteins that bind to specific sites on target antigenes. T cells on the other hand carry out cell-mediated responses. They are generally classified into two groups expressing either cell surface CD4 or CD8 receptors. These receptors play an important role in the formation of the immune synapse between the T cell receptor (TCR) and MHC of lymphocytes and target cells (see Fig. 2.2). These cells are also known as cytotoxic T lymphocytes (CTL) because upon recognition of the MHC complex by the TCR they secrete cytotoxic granules and perforin into the immune synapse that penetrate the target cell and induces apoptosis [23]. CD4 T cells are commonly referred to as helper T cells because they play an important role in contributing to the cytokine response that drives either cell mediated immunity by macrophages and CD8 T cells or humoral immunity mediated by B cells (Fig. 2.2).

We will focus on CD8 T cells in the following paragraphs since these cells are the main focus of the study presented in this part of my thesis.



FIGURE 2.2: Behavior of CD8 and CD4 cells upon TCR activation. Taken from McComb et al. [23]

In contrast to the broad specificity of innate immune cells towards pathogens, adaptive immune cells are antigen specific. Indeed, out of the billion or so T and B cells in the human body only a handful (10 to a 100) are specific for any given pathogen protein [23]. These cells recirculate in the lymph nodes or the spleen until they meet an APC presenting their specific antigene [24]. When a T cell encounters an APC and its T cell receptor has a high affinity for the MHC-peptide complex, the cell-cell interaction is reinforced with adhesion receptor/ligand interactions. The Major Histocompatibility Complex I (MHC-I, or MHC in the current context) is a protein complex on the cell membrane that presents antigens to immune cells and help trigger an immune reaction. The point of interaction between the T cell and the antigen presenting cell is the immunological synapse (IS). These co-receptors modulate the immune response and are expressed upon reception of activation signals such as inflammatory cytokines such as IFN- γ . The activation of the T cells therefore is conditioned upon the recognition of the right peptide presented on an MHC complex and in the right co-stimulatory environment. The activation of the IS drives a cascade of biochemical signaling leading to the rapid clonal proliferation of the specific CTL population [24]. From a handful of cells expressing the right receptor, the immune system can rapidly produce millions of clones to fight back an infection.

The CTLs then move into the periphery where they scan for their cognate antigene. If the activated CD8 T cell binds tightly to a target cell, it releases cytolytic granules into the cell synapse that penetrate the target cell causing its death as shown in figure 2.4. The immunological synapse can form between multiple target cells simultaneously



FIGURE 2.3: Behavior of CD8 and CD4 cells upon TCR activation. Taken from McComb et al. [25]

It is important to point out that the activation of the IS is conditioned by the expressing of co-receptors at the IS site (see Fig. 2.3a). These can modulate the immune response as for example the case of immune checkpoint inhibitors PD-1 and CTLA-4 that upon expression from the target cell inhibit the activation of the TCR [26], preventing target cell death and promoting immune escape. Upon successful activation of the TCR, a calcium influx occurs within the CTL [27, 28], leading to the exocytosis of the cytolytic granules and performs at the IS site (Fig. 2.3b). The performs form pores in the target cell membrane, causing cell lysis and death. Simultaneously, the granzymes encapsulated in the granules penetrate the cells and trigger target cell apoptosis via multiple parallel mechanisms. This leads to selective killing as shown in the intravital imaging in figure 2.4: the T cells (green) migrate and kill target tumor cells causing them to die by apoptosis (blue).



FIGURE 2.4: Target tumor cells are killed by T cells in the bone marrow. Taken from Cazaux et al. [29].

2.1.4 The tumor micro-environment modulates the immune response

The efficacy of the immune reaction against tumors is modulated by the interplay of multiple different factors, among which the physical structure of the matrix, oxygen levels, the chemical cues as well as the cellular composition of the tumor microenvironment (TME). The additive effect of these multiple signals is complex and still is actively studied. In the following lines I will describe some of these effects to illustrate the kind of influence the TME can have on the immune response.

The presence of dense ECM fibers around tumor islets can restrict the T cell access to tumor cells, despite immune-immune cell signaling attracting the CTLs to the tumor site [30]. Quantification of the displacement of CTLs shows that the CTLs preferentially migrate along the surfaces parallel to the fiber axis, and the tumor-stroma interface. This limits the entry of immune cells into the tumor and indicates that a mechanism of immunosuppression limiting the efficacy of CTLs in the tumor is the inhbition of their locomotion in the stroma [30]. Indeed, *in-vitro* studies have demonstrated that CD8 concentration significantly affects tumor clearing efficacy [2], and so limiting the local CTL concentration limits the immune response.

A second factor that can influence the effect of the immune response is the TME composition. Indeed, certain cell types such as tumor associated macrophages (TAMs) limit the efficacy of immune responses [31]. For instance TAMs engage inhibition receptors on T cells and secrete multiple cytokines down-regulating the activity of CTLs. TAMs are known to impede CTL motility to the tumor cell site [32]. This is done by engaging the CTLs in long-lasting interactions limiting the motility of the CTLs towards the tumor and creating "immune-excluded" regions. Conversely, Peranzoni et al. [33] show that depleting the TAMs from the TME restores CTL migration and improves the efficacy of anti-PD-1 immunotherapy. But the action of the immune system itself also contributes in reshaping the TME, by for example restricting the genetic diversity of the tumor cells: as tumor-specific CTLs infiltrate and kill the tumor cells in the TME, tumor sub-populations escaping the immune response expanded leading to a narrowing of the genetic diversity in the TME [29, 34]. This changes the genetic composition of the tissue an can potentially impact the efficacy of later treatments.

The TAMs are one among the many different cell types secreting chemokines that influence the CTL behavior. These cytokines can promote immune cell efficacy, promote CTL migration to the tumor or conversely inhibit CTL activity in the tumor environment. For example, interferon γ (IFN- γ) is secreted by CTLs (as well as other immune cell types) and has pleiotropic effects during anti-tumor immune responses; it triggers the production of CXCL9, CXCL10 and CXCL11 attracting more immune cells to the tumor [4], it also upregulates the production of MHC-I by the target tumor cells facilitating TCR activation and subsequent killing [3]. The multiple cytokines secreted by tumor cells, effector immune cells and bystander cells contribute in the complex dynamics determining the efficacy - or not - of the immune response [35].

2.1.5 CTL migration: search strategies and motion

As shown by Peranzoni et al. [33], immune cell trafficking to the tumor site plays a key role in successful tumor rejection. CTL migration in the tumor micro-environment has been abundantly studied in both in-vivo and in-vitro systems and CTLs have been reported to undergo several movement types depending on their environment [36].

Theoretically, it would be advantageous for CTLs to optimize the exploration of the tumor environment with superdiffusive walks in low-antigen concentration areas and diffusive or subdiffusive behavior where the antigens are highly present [36]. Indeed, superdiffusive walks increase the chances that CTLs "miss" their target cells in a given region, but are highly effective for exploring larger volumes: this is the "explorationexploitation trade-off" [36]. Superdiffusive Levy-walk style behavior of CTLs has been reported in-vivo [37], although other motility models have been used to explain observed CTL motility patterns [38]. In vivo, studies demonstrate that cells move faster and meander less in antigen-depleted regions compared to regions rich in cognate antigens [3]. This heterogeneity in motility patterns leads to the concentration of CTLs in regions rich in antigens, as can be experimentally observed in figure 2.5. To demonstrate the accumulation of CTLs in cognate-antigen expressing regions, Thibaut et al. created tumors in mice models constituted of a mixture of OVA-expressing and OVA-nonexpressing B lymphoma cells labeled with CFP and YFP (Fig. 2.5a). The injection of GFP+ CD8+ OT-I T cells recognizing the OVA peptide led to their preferential accumulation in the OVA+ tumor regions (Fig. 2.5b).

In vivo, T-cell motility is function of the physical matrix properties [39, 40]. The motility of CTLs is also known to be influenced by the production of cytokine gradients, among which the CXCR3-dependent ligands CXCL9/CXCL10/CXCL11 [37, 41] and the CCR5-dependent ligands CCL3/CCL4 [4]. They actively bias the motility of CTLs towards their targets, leading to enhanced accumulation of effector cells at the target site. This bias has been reported to have two different origins: preferential movement towards the target site [4] or CTL slow-down in the vicinity of the target [41]. CTL motility changes have also been reported in 2D in-vitro settings after presentation with the cognate antigen with CTL velocities depending on cognate antigen concentrations [42].

However, the efficacy of chemotactic cell signaling at low effector cell concentrations remains unknown. Indeed, experimental constraints require the use of high CTL concentrations in-vivo and in-vitro [4,37,41,42]. This probably leads to high cytokine concentrations in the tumor micro-environment that aren't representative of the initiation of the immune response In vivo. Furthermore, conventional methods for studying

2.2. MICROFLUIDIC CHIPS CAN REPLICATE IN-VIVO IMMUNO-ONCOLOGY PATTERNS



FIGURE 2.5: **a** OVA+ (blue) and OVA- (yellow) $E\mu$ -myc cells are implanted in mice to form tumor models. Subsequently, OVA-specific CD8+ cells are injected and the tumors are imaged using 2 photon microscopy. **b** The CD8 cells mainly accumulate in OVA+ tumor regions and induce tumor cell killing. Taken from Thibaut et al. [3]

CTL motility require a clear-cut signal and are highly sensitive to noise [43]. This means the effect on CTL motility of the cytokine production by a few (1 or 2 cells) contacts between tumor and target cells at the onset of the immune response has yet to be measured.

2.2 Microfluidic chips can replicate in-vivo immunooncology patterns

One of the (many) problems slowing down the development of new anticancer therapies is the lack of preclinical models that can be used to identify key molecular, cellular and biophysical features of human cancer progression. Conventional 2D cell cultures are widely used to assess the effects of drugs on tumor viability, however the results are often disappointing. This is because most in vitro cancer models fail to faithfully recapitulate the local tissue and organ microenvironment in which tumours form and which substantially contribute to the complex pathophysiology of the disease [44]. In response to this, more complex in vitro cancer models have been developed, including transwell cell cultures, spheroids and organoids grown within flexible extracellular matrix gels, which better mimic normal and cancerous tissue development than cells maintained on conventional 2D substrates [45]. This allows investigators to use models of intermediate complexity, replicating complex physiological tumor behaviors (metastasis, hypoxia etc.) in easily manipulable and controllable protocols (Fig. 2.6).



FIGURE 2.6: The schematic summarizes the different *in vitro* and *in vivo* models available for studying biological processes and testing drugs in order of increasing physiological relevance. Taken from Boussomier-Calleja et al. [6].

The issues described in the response to anticancer therapies also holds true in the response to immune challenges. Many parameters are known to influence patient response to immunotherapies. Among those, the number, phenotype and distribution of CTLs can have a strong predictive value in several types of cancer [46]. Understanding the interactions between immune and tumor cells is therefore key to improve diagnostics and treatments of cancer patients.

Multiple approaches are available to evaluate T cell cytotoxicity against tumors. In-vitro assays in cell suspension have been used to measure CTL killing capacity and, when performed at the single-cell level, provide information on the extent of functional heterogeneity within a T cell population [47]. These *in-vitro* assays however lack the complexity of the 3D tumor microenvironment, which strongly impacts T cell behavior and function. At the other end of the spectrum, intravital imaging offers direct insights into the dynamics, signaling and killing behavior of single T cells within a developing tumor [29,48–50]. Limitations of these approaches however include the fact that they provide a view of the interactions in a limited spatial and temporal window. Indeed continuous observation periods are generally limited to a few hours, precluding a full understanding of T cell histories in the TME.

An interesting emerging platform comes from advanced *in-vitro* models that recapitulate some aspects of the TME while providing access to the system dynamics [52, 53]. These include organoids [54, 55], where cells are allowed to organize in three dimensions



FIGURE 2.7: Upper panel shows examples of real trajectories performed by WT spleen cells, while lower panel depicts the same for KO spleen cells. WT, but not KO, splenocytes migrate from top-right through bottom-left following a chemokine gradient. Taken from Agliari et al. [51]

(3D), or organ-on-a-chip devices [6], where the microfluidic device represents the organ geometry and the microfluidics enable temporal control of the flows and physical conditions. Recent work has also dealt with combining the advantages of both approaches to produce organoids-on-a-chip [56]. Microfluidic chips have already been used to prove the complex relationship between immune cells, targets and soluble factors, for example demonstrating the role of chemokines in guiding spleen cell motion (Fig. 2.7) or the environmental variables influencing spheroid killing by T cells (Fig. 2.8) [57]. In Aref et al. [8], they use ex vivo cultures to model the efficacy of immune checkpoint blockade. Other studies have mimicked the complex interplay between tumor, support and immune cells on chip [58].



FIGURE 2.8: **a** Hepatoma tumor spheroid models are grown in a microfluidic chip. **b** The CD8 cells accumulate in tumor regions. Taken from Pavesi et al. [57].

However, a lingering issue with these devices is throughput. Mimicking in-vivo behavior is an important first step, but for many questions it is required to be able to produce a large number of parallel experiments. For example, to investigate the heterogeneity of the immune cell population it is necessary to sc'reen thousands of cells. Similarly, if one investigates inherently stochastic processes, one has to access a large number of samples in order to infer the probability distributions at stake. Many microfluidic chip devices proposed in the literature do not permit to increase the throughput significantly, which calls for new methods to be developed. One solution to this issue is to use droplet microfluidics. In the next section we will describe how microfluidic tools can be used to generate immune-tumor models.

Chapter 3

Droplet-based assays for 3D spheroid culture

Sections of this chapter are excerpts from:

Sart S., Ronteix G., Jain S., Amselem G. and Baroud C. (2021). *Cell culture in microfluidic droplets*, Under review.

Using droplets to encapsulate small volumes of fluids and their contents offers many attractive features. Indeed the physical and chemical differences between the two phases, such as water and oil, can be leveraged to manipulate the fluids. For example the chemical contrast between the two phases allows some species to be dissolved in the droplets with no losses to the surrounding oil [59], while the interfacial tension between the two phases can be lead to forces that allow the droplets to be manipulated [60]. As a result the field of droplet microfluidics has flourished through a large body of work that developed a toolbox to create, transport, and further manipulate the drops and their contents. The resulting technologies were then used as a basis for applications to chemistry and biology.

Droplets have been used to encapsulate biological and chemical assays long before microfluidics. Some early works date back to André Lwoff who used a pneumatic micro-manipulator to encapsulate single bacterial cells in femto-liter drops [61]. This encapsulation allowed Lwoff to culture and observe individual bacterial cells under a microscope in order to study lysogeny [61]. The method was later used to study single cell infections in monkey by the poliomyelitis virus [62], as well as for measuring the production of antibodies from individual cells [63]. Later still, Rotman used a nozzle to spray droplets of diameter around 10 μ m (volumes in the pl range), in order to observe the product of reactions from single enzyme molecules [64].

But since the beginning the of 21st century, microfluidics has matured rapidly finding applications in chemistry and biology. Here I will describe some of the physical principles of microfluidics, the technology from the lab and how it can be used to generate 3D spheroids at scale.

3.1 From surface tension to droplet generation

3.1.1 The cost of an interface

Individual molecules forming a liquid interact with their neighbors constantly. Indeed, it the very existence of these interactions which promote the creation of a liquid phase and without them the molecules would disperse into the air. But the existence of favorable interactions between molecules in the liquid implies that any interface of surface S comes at an energy cost. A change in the surface if the liquid-liquid or the liquid-gas interface represents an energy cost given by equation 3.1.

$$\Delta E = \gamma \Delta S \tag{3.1}$$

 γ is called surface tension and the higher it is, the higher the energy penalty for the creation of supplementary interfaces. It is typically worth 72 mN/m for the water-air interface we see in our everyday lives. The water-oil surface tension on the other hand is worth approximately 50 mN/m [65]. This is the force the cook is fighting against when making mayonnaise: the multiple oil-in-water droplets create an excess of water-oil interfaces, increasing the energy of the system. The struggle to minimize interfacial energy also influences the shape that droplets of liquid will take: the droplet, of a fixed volume V, will seek to minimize its surface S, leading to the formation of a sphere. Captain Haddock discovers this to his dismay: in space, free from the constraints of gravity, whiskey forms spherical drops instead of staying put in a glass (Fig. 3.1).



FIGURE 3.1: In space, only capillary forces shape Captain Haddock's whiskey, leading to perfectly spherical drops.

In our day-to-day lives, we seldom encounter perfectly spherical droplets. This is because capillary forces, although significant at smaller scales are quickly dominated by gravity when the studied objects become bigger. In order to distinguish between these different scales, physicists have introduced a characteristic number, the capillary length κ .

$$\kappa = \sqrt{\frac{\rho g}{\gamma}} \tag{3.2}$$

Looking at κ , we see that it is the result of the balance between the surface tension (represented by γ) and gravity (represented by the product between the liquid volumic mass ρ and gravity g). For problems where the typical dimensions are much smaller

than κ , capillary forces rule and gravity can be neglected. On the other hand if κ is smaller than the typical dimension of your problem, then gravity will dominate. For water, $\kappa \approx 2.5$ mm, so the captain Haddock is right to be surprised to see his drink forming a perfectly round sphere since gravity would rule the physics of any typical whiskey glass on Earth.

On the other hand, physicists conducting experiments in microfluidics, where the typical dimensions are of the order of 100μ m, can safely neglect the influence of gravity on their experiments and study the formation of droplets whilst only considering surface tension effects. Once the physics guiding the manipulation of liquids are micrometric scales has been established it is possible to engineer devices to generate and guide microfluidic droplets as will be described in the following section.

3.1.2 Generating droplets

Many techniques have been developed to produce monodisperse microfluidic droplets [9,66], but one of the most commonly used is the flow focusing junction (Fig 3.2a). In such a junction the aqueous phase is injected at the center of the junction, it is then pinched by two symmetric oil flows causing the aqueous flow to break and generate droplets at regular intervals. The size of the droplets depend on the junction geometry, the respective capillary forces and the viscosity. Modifying the composition of the droplet, for example by adding polymers such as non-gelified Matrigel to the aqueous phase, also considerably changes the physics of the droplet formation, requiring the experimental protocols to be adapted.



FIGURE 3.2: **a** Droplets can be generated at a microfluidic junction or by using step emulsification (**b**) [60]. **c** Combining step emulsification and microfluidic junction allows for a more robust droplet generation. **d** Adding a supplementary oil bypass prevents droplets from coming in contact after being generated. Figures are extracted from reference [67].

Droplets can also be produced by step emulsification (Fig 3.2b). Here the droplet breaking mechanism comes not by pinching from an oil phase but from a sharp increase of the channel height that changes the curvature balance in the aqueous thread [68]. In order to increase the reproductibility of the droplet production (even in the case of a high viscosity dispersed phase) a step is added just after the junction in the chip
design used in the Physical Microfluidics and Bioengineering lab (Fig 3.2c) [9]. Another addition to the design is to add bypasses on both sides of the junction allowing oil to flow and ensuring that the droplets are well separated immediately after their formation (Fig 3.2d).

A major issue once the droplets have been formed is to prevent them from merging inside the microfluidic device. This is energetically favorable as it would decrease the total energy of the system. One way to stabilize them and prevent them from merging is to add surfactants to oil-water interfaces. These are usually amphiphilic molecules that adsorb at the interface, lowering the surface tension and therefore the energy gained by merging with another droplet. The surfactants also generate Marangoni stresses [69] making fusion between droplets more difficult. The combination of these kinetic and thermodynamic effects stabilize the droplets and prevents them from merging. In the current experimental protocol, the surfactants are added to the oil and adsorb to the interface in an emulsification channel [70].

3.1.3 Guiding and anchoring microfluidic droplets

Once the droplets have been generated, it is necessary to develop the tools to manipulate them within the microfluidic device in order to conduct experiments on the droplet contents. Passive manipulation techniques, requiring no exterior energy are well-suited for this task. The lab has developed several techniques over the years to trap [71] and guide [72] droplets in microfluidic chips. The etches or the wells in the chip walls allow the droplets to reduce their surface by a surface difference ΔS at constant volume. That is, their surface energy is lowered by $\Delta E = \gamma \Delta S$. If we call d the well dimension we expect the trapping force to be of the order of $F_{\gamma} \approx \frac{\Delta E}{d}$ [71], anchoring the droplet in the well despite an oil flow pushing it away (Fig. 3.3a). As shown in figure 3.3b, microfluidic rails make it possible to guide droplets in the chip chamber using the same physical principles.



FIGURE 3.3: **a** A microfluidic droplet is kept in place despite the oil flow in the chamber thanks to an anchor (left), the anchoring force is driven by the relative decrease in the interface sie (right). **b** Droplets can be guided thanks to microfluidic rails. Illustration from [67]

Subsequently, the well dimensions have been optimized to improve the trapping efficiency of the microfluidic droplets in the chamber [67], allowing for the creation of droplet arrays in microfluidic devices [9]. Combining the different microfluidic techniques allows to conduct complex, multi-step experiments involving droplet generation, stabilization and storage in a single device as shown in figure 3.4.

We have now described the essential building blocks necessary for conducting experiments on microfluidic devices. We have shown how droplets can be generated,



3.2. DROPLET-MADE TUMOR SPHEROIDS



FIGURE 3.4: Schematic of the microfluidic chip developed in the lab. It allows for the simultaneous generation and storage of droplets in a single device [67].

guided and stored in microfluidic chips, permitting the generation of independent biological samples and their downstream analysis. In the next section, we will describe how microfluidic droplets enable the culture of tumoroids on a chip and how we can use the microfluidic chip described above to generate immuno-oncology models.

3.2 Droplet-made tumor spheroids

3.2.1 Spheroids as 3D cell culture models

In the past decade, consistent efforts have been made to increase the relevance and complexity of *in vitro* cell cultures, so as to recapitulate *in vivo* behaviours [73]. One method is to promote the self-aggregation of cells into coherent spheres (or spheroids). These objects can be mono- or multi-cellular depending on the number of cell-types composing it. Not all cell types can form spheroids [74] and the formation of spheroids can be enhanced by the addition of extra-cellular matrix to the culture medium. Spheroids composed of a larger number of cell types and recapitulating more complex behaviours are sometimes named organoids to underline their increased physiological relevance.

Spheroids have been extensively used in pharmacology for drug discovery and testing, and a large number of methods have been proposed to fabricate them [73]. In the following sections, we will expand on the protocol developed in the Physical Microfluidics and Bioengineering group to make spheroids in microfluidic chips.

3.2.2 Making spheroids in microfluidic droplets

Once one can reproducibly fabricate droplets one can dilute cells in the aqueous solution to generate cultivate them in the droplets. This technique has been adopted on a large variety of cell types such as the MCF-7 breast cancer cell line [75], hepatocytes [9], mesenchymal stem cells [76] and many more. Once the droplets have been generated they are trapped in microfluidic anchors [9], sieves [75], traps [77] or moved and cultured in conventional culture vessels such as 96 well plates [78] for further study.

In figure 3.5 we present an example protocol for generating spheroids on the device presented in figure 3.4. Differences between cell lines and cell types mean that the cell



FIGURE 3.5: Protocol for culturing spheroids in microfluidic droplets on chips. Modified from Sart et al. [9]

culture conditions may be modified to suit the experiment. For instance, variations in cell metabolism mean that culture times may have to be reduced from 5 days for B16 melanoma cells to just 24h for neural stem cells. Similarly, certain cell types need an extracellular matrix (ECM) to form spheroids. Therefore special care needs to be given to the droplet micro-environment.

3.2.3 Iterative droplet trapping and fusion



FIGURE 3.6: **a** Smaller microfluidic droplets introduced into a microfluidic chamber containing double-anchors (left) leads to the trapping of the droplets in the secondary anchors (right). Scalebar is 400 μ m. Taken from [12] **b** Schematic of the CTLs (green) in the secondary droplet before the fusion of the two droplets and the introduction of the CTLs in the gelified, spheroid-containing droplet [79].

In sub-section 3.2.2, we described how microfluidic technology could be used to generate an array of spheroids encapsulated in gel-droplets. Many protocols require the successive controlled addition of new cell types. In the lab, we have developed a system for trapping smaller droplets in secondary anchors, adjacent to the main anchor present in each well [12]. These anchors are too small to efficiently trap the main droplets (i.e. the force generated by the oil flow on the droplets is much larger than the trapping force from the smaller traps), however they efficiently capture smaller droplets as shown in figure 4.5a. Previous work shows that the addition of surfactants in the oil causes the droplets to fuse, liberating the content of the smaller droplet into the big one [12]. To fuse the primary with the secondary droplets, 20%(v/v) of 1H,1H,2H,2H-perfluoro-1-octanol (PFO) was dissolved in NovecTM-7500 Engineered Fluid (3M) and is perfused in the microfluidic chip. This causes the fusion of adjascent droplets. PFO is toxic for many cell types, and so after the fusion of droplets, a fresh solution of FC40 and surfactant is flushed-in to remove the PFO from the microfluidic chamber.

3.2.4 Building complex micro-environments in microfluidic droplets

The cellular microenvironment plays an important role in governing functional processes in different organ specific tissues [80] and its alteration can give rise to several pathological conditions such as inflammation or fibrosis [81]. In the case of cancerous tissues, the tumor microenvironment comprises both physical and biochemical factors that can impact the tumor fate [82–85]. Specifically the tumor microenvironment (TME) is known to play a crucial role in dictating the tumor growth, tumor angiogenesis, cancer metastasis and drug resistance [86–89]. Since the objective of droplet-based tumor assays is to challenge the tumor spheroids in relevant conditions it is necessary to recapitulate some of the complexity of the TME in the droplets.

Several physical signals like the 3D geometry or the architecture of the surrounding matrix, its stiffness, rigidity, shear or compressive stresses present in the TME can alter cancer cell proliferation, Epithelial to Mesenchymal Transition (EMT) and cancer cell invasion [84]. For example, the growth of cancer cells and the cellular composition within a tumor are altered in 3D microenvironments, often with the apparition of the multilayered structure of cancer cells within the tumor. This type of complex organization is absent in conventional 2D monolayers [78,90]. Several studies have aimed to reconstruct a relevant 3D architecture within microfluidic droplets, for instance Yu et al. tested the compatibility of different hydrogels for allowing spheroids of MCF-7 cells to form in droplets [91]. They co-encapsulated small groups of cells with either Matrigel or collagen, and compared the ability of cells to form spheroids to the case of cells encapsulated in passive alginate hydrogel. The study showed that spheroids could not form within the alginate, while they formed compact 3D structures in the other hydrogels (Fig. 3.7). Separately the interaction between matrix rigidity and tumor growth was studied by Lee et al. who found that matrix rigidity also alters tumor growth, whose rate was found to be directly proportional to the surrounding rigidity [92,93].

With the recent technical advancement in the field of droplet microfluidics it is now possible to mimic the TME and approach a more physiologically relevant state by adding gels to the droplets, although this requires some technical improvements as described in section 4.2.2.

3.2.5 Droplet microfluidics in immuno-oncology studies

Microfluidic droplet technology has opened up the possibility to study the immune cell dynamics with a unique level of precision and granularity. Several teams have used these droplets as model microenvironments in order to probe immune cell activity.



This protocol has been adapted to several similar experiments, with droplet volumes typically in the picoliter-nanoliter range [96, 97] and containing a few cells each. The concentration and relative composition of the droplets is controlled by fine-tuning the composition of the initial aqueous solution. The drops are then stored in a droplet array composed of traps in a similar manner to what is described in section 3.2.2 [97–99] and imaged at regular intervals. Specific biological cues (such as cell death, cytokine secretion etc.) can be tracked via fluorescence microscopy by adding specific markers into the aqueous phase forming the droplets. This experimental protocol has been used to challenge single tumor cells with antigen-specific immune cells in microfluidic droplets. For instance, to investigate the interplay between Interferon- γ (IFN- γ) secretion and anti-tumoral cytotoxic capabilities by Natural Killer (NK) cells, Antona et al. developed a droplet-based protocol where they simultaneously measured in-drop IFN- γ concentration using antibody functionalized polystyrene beads and the killing of tumor cells by fluorescence microscopy [100]. This procedure revealed that killing events were correlated with increased detection of IFN- γ in the droplets. They also showed that the addition of IFN- γ to the initial droplet reduced the cytotoxic capacities of NK cells.

In summary droplet microfluidics can be useful for probing the immune system. It has been adapted for a wide range of immune-cell types, from Jurkat cells [101], NK cells

al. [91].

[102,103], dendritic cells [104] and B-cells [105]. The complexity of the immune system resides in part in the multiple interacting spatio-temporal scales. Cell signaling can be modulated by the secretion of chemokines and span several hundreds of micrometers [41] or by cell-cell contacts [11]. Conversely, immunological processes can be of the order of a couple of minutes [29] to several hours [106]. Droplet microfluidics is a relevant platform to construct improved immunological models since the small droplet volumes increase the chances for observable interactions to happen. Furthermore, the high throughput of the technology allows for improved quantitative characterization of the immunological models. Taken together, these elements suggest that droplet-based immunological models have the potential to improve the study the immune system in a wide range of contexts. Droplet-based functional assays have been developed for screening individual T cells, but they have yet to be developed to investigate the complex interplay between tumor and immune cells in a quantitative manner. This is where the work presented in this PhD is of particular interest for the community.

CHAPTER 3. DROPLET-BASED ASSAYS FOR 3D SPHEROID CULTURE

Chapter 4

Experimental protocols for immuno-oncology assays using droplet microfluidics

The condensed version of chapters 4, 5 and 6 is available in:

Ronteix G.*, Jain S.*, Cazaux M., Angely C., Khazen R., Bousso P. and Baroud C. (2021). A Multiscale Immuno-Oncology on-Chip System (MIOCS) establishes that collective T cell behaviors govern tumor regression, Under review. Available on biorXiv, doi:10.1101/2021.03.23.435334.

(*): joint first authors.

4.1 Introduction

The Physical Microfluidics and Bioengineering group has developed protocols for cell culture in microfluidic droplets, experimenting with MSCs [76] and hepatocytes [9]. During my PhD, I updated these methods to culture melanoma spheroids in the droplets and challenge them with T cells.



FIGURE 4.1: A short schematic describing the experimental pipeline of the current project. It combines microfluidic protocol development, cell culture, image processing and analysis, as well as mathematical modeling to interpret the generated data.

CHAPTER 4. EXPERIMENTAL PROTOCOLS FOR IMMUNO-ONCOLOGY ASSAYS USING DROPLET MICROFLUIDICS

The experimental pipeline is summarized in figure 4.1. The interactions between CTLs and tumour spheroids were tracked with microscopy imaging, recording the droplets and their content every two minutes for up to 24h. In order to extract the information from these data, we developed an image analysis program that from an image stack automatically segments the spheroid, extracts the T cell positions and records the position of the T cells with regard to the model tumour. Finally, downstream analysis of these processed data enable us to measure the accumulation of the T cells on the spheroid and quantify the interactions between the CTLs in the droplet.

Combining experimental and theoretical work is necessary in order to access the information on T cell behavior. The high spatio-temporal resolution and the parallelization of the recently developed "multiscale cytometry" technique [9] requires new methods to extract the maximum amount of information from the generated data. In the following sections I will detail the different steps of the experimental procedure allowing us to study the immune-tumour interactions in the microfluidic droplets.

4.2 Model tumor formation

4.2.1 Constructing a relevant tumour model

The experiment begins by producing aqueous droplets (Volume = 50 nl) containing Matrigel and a suspension of B16 cells at a concentration of $1.5 \cdot 10^6$ cell/ml. Once the droplets are anchored, the device is placed in an incubator at 37°C overnight, which allows a single B16 spheroid to form in each droplet (Fig. 4.2b). After 24h, this led to the formation of 234 single spheroids in independent droplets separated by an oil phase [9,13].



FIGURE 4.2: **a** Microfluidic chip on a standard glass slide. **b** Expanded view of the trapping region of the chip (dashed box) showing an array of 234 trapped droplets. Each droplet contains a single B16 spheroid in Matrigel, as shown in the inset.

The tumor cells used were B16-WT melanoma and B16-Ovalbumin (B16-OVA) peptide expressing cell-lines. The use of a model antigen like OVA expressed by tumor cells serves to mimic a tumor neo-antigen, since most natural tumor neoantigen in tumor cell lines are not characterized. The B16-OVA model is a classic preclinical melanoma model that has been used in many other studies (>300 papers reported on

pubmed) to analyze T cell responses and responses to immunotherapies in multiple settings including anti-PD1, STING agonist, dendritic cell vaccines, oncolytic viruses etc. The availability of transgenic mice expressing a T cell receptor recognizing this epitope (OT-I TCR T cells) offers the opportunity to study the impact and fate of tumor-specific T cells on model tumors. In particular, to interpret imaging experiments, it is often essential to know the specificity of the T cells analyzed, therefore the use of T cells expressing a transgenic TCR is a very common approach. The B16-OVA tumor model combined with OT-I T cells has also been used for intravital imaging to dissect intratumoral T cell behavior and killing by several groups [10, 107].

B16-OVA were generated by retrovirally transducing parental B16-WT cells, and B16 and B16-OVA cells only differ by the expression of the model antigen. The expression of an additional antigen (OVA) could in theory increase the immunogeneceity of the tumor cells when injected in vivo. However, a study inoculating a mixture of B16 and B16-OVA reported that 'two cell populations proliferated similarly, maintained the original injection ratios, and formed thoroughly admixed tumors [107].

4.2.2 Incorporating Matrigel in droplets

The choice of the ECM is constrained by the experimental requirements. We needed a collagen-rich environment to allow the free movement of the CTLs in the droplets and to account for interactions between the tumor spheroid and the extracellular matrix and this led to the choice of Matrigel for our experiments.

However, loading the chips with Matrigel was an experimental challenge. The gel permanently polymerizes at temperatures above a dozen of degrees (the exact temperature depends on the concentration). This means that the chip, as well as the oil and the tubing need to be cooled down in order to prevent gel polymerization before the droplets are trapped in the microfluidic anchors on the device. If the gel polymerizes in the tubing then the droplets can't be formed at the junction. Similarly, polymerization in the chip before the droplets are trapped prevents them from being properly anchored in the wells, keeping the gel droplets liquid is therefor crucial. The trick is to cool the oil and the chip to a couple of degrees above zero in the fridge before the experiment, if the loading is conducted rapidly enough then the temperature in the device stays below polymerization temperatures and the droplet trapping protocol functions properly.

Once the droplets are trapped, they are put in the incubator at 37 °C for 24h, allowing the gel to polymerize and the spheroids to form inside the droplet. The addition of ECM profoundly modifies the geometry of the B16 melanoma spheroids formed in the droplet. Indeed, without any Matrigel, the spheroids aggregate loosely (Fig. 4.3a - left), whereas the presence of Matrigel at high enough concentrations leads to the formation of tightly-packed, smooth spheroids (Fig. 4.3a - right). The need of a biocompatible extracellular matrix to form spheroids depends on the cell type. For instance, MCF-7 breast cancer cells or hepatocytes don't require the addition of matrigel or collagen and form spheroids in aqueous media [12, 78]. Interestingly, the spheroid phenotype in some experiments was heterogeneous within the same chip, with some spheroids starting to spread into the gel as shown in figure 4.3b. We did not however have the time to investigate further the drivers of this phenotypic heterogeneity in the chips. Matrigel concentrations higher than 2 mg/mL lead to the formation of multiple spheroids in the droplets: the increased viscosity of the media slows down sedimentation of cells at the

CHAPTER 4. EXPERIMENTAL PROTOCOLS FOR IMMUNO-ONCOLOGY ASSAYS USING DROPLET MICROFLUIDICS

bottom of the droplets and the heightened Matrigel concentration also increased the polymerisation speed. Thus, instead of falling to the bottom of the droplets, the cells remain spread out throughout the droplet and form multiple dissociated spheroids. Down the line, this made the experiment interpretation difficult and so lower Matrigel concentrations of 2 mg/mL were chosen. This concentration is in the same range as other studies of CTL behavior *in-vitro* [108].



FIGURE 4.3: **a** The addition of Matrigel in the drops modifies the dynamics and end-state of spheroid formation. **b** Matrigel-containing droplets can lead to "smooth" or "branching" spheroids within the same chip. Scalebar is 60 μ m.

At these cell and Matrigel concentrations we obtain spheroid radii in most cases ranging between 35 to 45 μ m (Fig. 4.4a). To assess the viability of the B16-OVA spheroids, the droplets were made according to the protocol described above with the addition of Propidium Iodide (PI) at a concentration of 3μ M. The chip was then imaged at 24 and 48 hours after seeding. Only spheroids positioned in the center of the droplets were imaged, in order to avoid artifacts due to the microscopy. Using a custom-made Imagej macro, the area of red fluorescent signal by PI was measured and the percentage fluorescent area was calculated when compared to the complete spheroid area. The live-dead staining shows that less than 3% of the cells were dead after 48h in the chip (Fig. 4.4b), which suggests that the cell culture protocol is well tolerated by the B16 cells.

A key question remains that of nutrient availability in the droplets. It is partially answered with the cell viability data described immediately above. The very limited death rate (ca 3-5% of all cells, see figure 1) hints that we can ignore the effects of nutrient depletion during the course of our experiment. For any long-term experiments this issue has been dealt with in previous publications from the lab. For example we can replace the oil separating the droplets with media [9] or we can add and fuse supplementary media containing droplets with the spheroid containing ones [12].



FIGURE 4.4: **a** Distribution of spheroid radii within a single chip (n=215). **b** Viability measurements using live-dead staining after 24h and 48h (n = 54).

4.3 Modeling the immune challenge on an integrated microfluidic chip



FIGURE 4.5: **a** Smaller microfluidic droplets introduced into a microfluidic chamber containing double-anchors (left) leads to the trapping of the droplets in the secondary anchors (right). Scalebar is 400 μ m. This figure is duplicated from Tomasi et al. [12]. **b** Schematic of the CTLs (green) in the secondary droplet before the fusion of the two droplets and the introduction of the CTLs in the gelified, spheroid-containing droplet.

After creating the tumour models as described in sections 4.2.1 and 4.2.2, we need to challenge the tumours by adding the T cells in the droplets in a controlled manner. This is done by generating CTL-containing secondary droplets and fusing them with the spheroid-containing droplet following the method described in section 3.2.3 (steps 1-3 in figure 4.6).

The fusion of the droplets leads to the release of the CTLs into the gel. The number of CTLs detected per droplet is highly heterogeneous and depends on the initial concentration of CTLs in the solution. The dispersion of the number of CTLs per droplet in figure 4.7 is also explained by un-controled experimental variables: the



FIGURE 4.6: Schematic showing a primary droplet with a tumor spheroid, followed by the addition and fusion of a secondary droplet containing GFP labeled CTLs, eventually leading to tumor cell killing and spheroid fragmentation. Scale bar is 200μ m. Schematic representation of the complete experimental protocol.

CTLs may aggregate or settle in the tubing and in the syringes for example (Fig. 4.6).



FIGURE 4.7: Distribution of the number of recorded CTLs per droplet.

Once the CTLs are in the main droplet, the interactions between the CTLs and spheroids are observed by time-lapse imaging, typically over 24 hours. The whole process from spheroid preparation to CTL addition and imaging is done on a single chip (Fig. 4.6) and each experiment yields up to 234 individual replicates, of which we typically obtain 50 time lapse movies, due to the small image acquisition time-intervals (2min/frame). The limiting factor for the capture of additional replicates is the imaging rate of the microscope. For long-term studies where low-frequency imaging is sufficient it is possible to image the whole chip.

4.4 Spheroid fate is measured via a tailor-made image analysis pipeline

The experimental protocol described above generates high volumes of data. A single chip will typically be imaged every two minutes for 20 hours at 50 different positions (i.e. approx. 30000 times). This requires the creation of tailored image analysis pipelines to extract information relevant for analysis. The routines were written in Python programming language and use several open-source libraries [109–111]. The pipeline to analyze the images is described below.

4.4.1 Segmenting the spheroid

The first step conducted for each image is the segmentation of the spheroid using only the information contained in the greyscale channel. The spheroid is detected by calculating the local gradient using laplace filtering and thresholding. From a greyscale image (Fig. 4.9 left) the mask outlining the high-gradient areas is constructed (Fig. 4.9 middle). Then the different detected regions are filled and filtered according to their eccentricity. Round shapes are kept and sorted by size, the biggest detected area being the categorized as the spheroid (Fig. 4.9 right).



FIGURE 4.8: Example of the segmentation procedure. The droplets are imaged in BF (left), from the BF image gradient-based analysis allows the detection of the spheroid contours (middle). The contours are used to build masks of potential spheroids, and additional filtering singles out the mask for the spheroid (right).

This procedure is repeated at every time step for more robustness. In images where the spheroid isn't detected or where the position and radius of the spheroid differ too much from the average values, we replace the spheroid position and radius with the average feature values. This increases the robustness of the final results by eliminating spurious segmentations.

The major issue leading to no spheroid segmentation were cases when the spheroid remained stuck near the droplet border. This occurred in approximately 10-15% of the wells.

4.4.2 Detecting the CTLs

When the spheroids have been segmented, the algorithm then analyzes the FITC (510 nm) fluorescence channel to detect the GFP+ CTLs in the droplet. To detect the

CHAPTER 4. EXPERIMENTAL PROTOCOLS FOR IMMUNO-ONCOLOGY ASSAYS USING DROPLET MICROFLUIDICS

individual CTLs, we use a state-of-the-art Python library used for particle-tracking called Trackpy [112]. It allows us to specifically pin-point the location of each CTL in the focal plane at each time-step. The cells are represented as the red and blue dots in figure 4.9. We then link the positions of the different cells at different times to build the tracks for each detected CTL.



FIGURE 4.9: Detected CTLs are represented on the greyscale (left) and FITC fluorescence channels (right). Cells detected as being on the spheroid are in red, those in the gel are in blue and the green dot on the left represents the detected spheroid center.

This information is then crossed with the spheroid positions extracted above to determine the position of the CTLs relative to the spheroid in the droplet. Cells on (resp. off) the spheroid are represented in red (resp. blue) in figure 4.9.

4.4.3 Post-process verification

We are highly dependent on the fidelity of the spheroid segmentation and the CTL tracking for the downstream analysis of our experiments. Any issue could lead us to the wrong conclusions and it is therefor necessary to be highly confident not only in the experiment but also in the initial steps of the experiment analysis. To control the results, for each timestep an image with the raw image, the mask covering the detected spheroid and the relative positions of the CTLs is generated for manual verification a-posteriori of the algorithm efficacy. Faulty wells (with for example mis-detected spheroids) are then thrown away before data analysis and ignored in the experiment analysis.

Key take-aways

Methods Combining of multiple techniques (microfluidics, *in-vitro* cell culture and image analysis), I have developed an experimental pipeline enabling me to get access to single-cell data over multiple replicates and over extended time-lapses of up to 24h.

Results This work shows that it is possible to study immune-tumour interactions in microfluidic droplets over 1-2 days.

Next steps In the following chapter, we use this uniquely resolved spatiotemporal data to study CTL accumulation on the spheroids.

CHAPTER 4. EXPERIMENTAL PROTOCOLS FOR IMMUNO-ONCOLOGY ASSAYS USING DROPLET MICROFLUIDICS

Chapter 5

CTL accumulation on spheroids in microfluidic droplets

After image analysis, three stages in the cell-cell interactions can be identified: the CTL exploration of the gel, their accumulation on the spheroid, and the killing of B16 Ova-expressing cells. We will describe the motion of the CTLs in the gel and their accumulation in the current chapter. The killing of the melanoma cells will be described in chapter 6.

5.1 CTL migration in micro-device reproduces *in vivo* behavior

The CTL migration is tracked in the time lapse movies as the cells perform 3D migration within the gel or on the surface of the spheroid (see Fig. 5.1a). The recorded velocities display alternating periods of motility and arrest phases, as seen by the high and low velocities in figure 5.1b. This behavior, as well as the value of the velocities, correspond well to previously reported CTL velocities in collagen gels *in-vitro* [108] or within tissues *in-vivo* [3,3,37].

In order to evaluate the influence of the spheroid presence in the droplet on the motility of the CTLs, the migration statistics without a spheroid present are compared with the statistics in the presence of a spheroid during the first 500 minutes of an experiment. We restrict the analysis to cells that are not in contact with the spheroid. The displacement distributions (Fig. 5.1c) and the mean-square displacements (MSD) (Fig. 5.1d) do not show any significant difference between the two conditions. In both cases, the CTLs undergo super-diffusive random walks (Fig. 5.1d) with MSD ~ τ^{α} , where τ is the time between two observations and $\alpha = 1.6$, in agreement with what was reported *in-vitro* and *in-vivo* [37, 108]. This confirms that the microfluidic device provides a good environment to mimic CTL motility *in-vivo*.

After some time, one CTL comes in contact with the spheroid. This contact generally leads the T-cell to adhere to the spheroid and explore its surface over the course of a few hours (Fig. 5.2a,b). We select individual tracks with segments both on and off the spheroid to investigate more precisely the CTL motility change upon reaching the spheroid. We observe that the CTL behavior is strongly modified: They display



FIGURE 5.1: **a** (left) Representative image of CTLs with instantaneous velocity vectors inside Matrigel droplet. (right) CTL tracks in one droplet over 24 h, each color represents an individual cell track. The dashed black circle outlines the spheroid boundary. **b** Representative velocities as a function of time for three different T-cells. **c** Probability distribution of a cell to migrate by a given distance (Δr) during a fixed time step $\Delta t = 1$ minute (n = 67965 points without spheroid and n = 34072 individual points for CTLs in presence of the B16 spheroids). **d** Mean Square Displacement (MSD) of CTL migration with and without spheroids.

lower mean velocity (Fig. 5.2c) with a lower MSD exponent (Fig. 5.2c,d). The average MSD exponent goes from 1.4 when the cells move in the gel to 1.1 after the same cells have reached the spheroid.

The transition from a high-velocity, high- α regime in the gel to a low-velocity, low- α on the spheroid expressing the cognate antigen can be caused by multiple factors such as a change in the substrate mechanical properties, the adhesion to receptors on the B16 cells or the recognition of the cognate antigen on the spheroid surface. The motility pattern change validates experimentally the "exploration-exploitation" arbitrage that CTLs are subject to [36]: the CTLs adopt an exploration mode where they can scan large volumes for their cognate antigen, and once it is found they change their motility pattern to thoroughly detect all cells expressing the cognate antigen in the local region.

The CTL migration in the gel therefore recapitulates behaviors that have been reported In vivo [3,37,108,113], with the current data highlighting the switch in motility before and after the CTL contacts the spheroid surface.



FIGURE 5.2: **a** Time sequence showing the initial CTL approach and contact with a spheroid. **b** Track of a single CTL as it migrates in the matrigel and on the spheroid surface. Colormap represents the instantaneous velocity of the cell. **c** Average velocity and mean square displacement exponent (α) of cells migrating in the gel and on the spheroid. Each data point is the average velocity in a given droplet ($n_{\text{gel}} = 55$, $n_{\text{spheroid}} = 54$). **d** Mean-square displacement of cells migrating in the matrigel and on the spheroid. Bold and dashed lines represent the best fits for the MSD of CTLs on the spheroid and in the matrigel, with respective exponents of 1.1 and 1.4.

5.2 A positive feedback loop drives CTL accumulation on the spheroid

5.2.1 Qualitative description of the CTL accumulation on the spheroid

The CTLs migrate in the gel until they hit the spheroid. After the first CTL contact with the B16 cells, the arrival of successive CTLs leads to an accumulation of T cells on the spheroid in the case of the B16-OVA spheroids. We can monitor this accumulation at the single-droplet level (Fig. 5.3a) and at the chip level (Fig. 5.3b). This accumulation is also shown by computing the number of CTLs in each droplet that are present on the spheroid divided by the total number of CTLs in the droplet as a function of time (Fig. 5.3c). However the accumulation is not observed in the case of wild-type B16 cells, which do not express Ova.

At this stage an important question is whether this accumulation results from cells reaching the spheroid randomly, as they explore the droplet volume, or if the accumulation rate is enhanced due to cell-cell signaling. We address this question by analyzing the accumulation of the CTLs at the spheroid level. At this scale, the CTL accumulation is not homogeneous but stochastic, with CTLs both attaching and detaching over



FIGURE 5.3: **d** Number of CTLs as a function of time on two representative spheroids showing the detection of attachment/detachment events **b** Number of CTLs detected on each spheroid and **c** the fraction per droplet as a function of time. Each thin line represents a single tracked spheroid, in bold is the averaged value. In red is the accumulation for B16-OVA spheroids and in blue for B16 WT spheroids. [84 individual B16-OVA spheroids and 81 B16 WT spheroids tracked].

time (Fig. 5.3a, **Movie 2**, **Movie 4**).

5.2.2 The CTL migration in the droplet can be approximated with a random walk

Multiple models have been proposed to describe CTL motility [37]. However, the pertinence of any given description depends on the observation scale and the complexity of the description. Although it is known that Brownian motion only provides a rough estimation of the behavior of CTLs, we will show that is pertinent to describe their motility in the gel. Indeed, before reaching the spheroid, the cells have typically explored a large portion of the droplet, as seen in the tracked trajectories of Fig. 5.1a. In order to assess the pertinence of the Brownian motion (BM) model, we will investigate the key features of the CTL motion the gel droplets. A random walk can be characterized by the autocorrelation C of its observables (velocity, orientation etc.). For instance, the velocity autocorrelation is defined as $C_v(\tau) = \frac{\langle v_i(t) \cdot v_i(t-\tau) \rangle_{i,t}}{\sigma_z^2}$.

Persistent random walks, where the cell has a non-trivial memory time have nonnull autocorrelation times, meaning that the autocorrelation decreases slowly to zero (slowly with regard to the characteristic time of the system). On the other hand, Brownian motion has a null autocorrelation time, each step being independent from the previous ones. We show that the velocity and velocity angle autocorrelation plots are well fitted by a sum of exponentials, with the largest correlation time less than 7 minutes for both (Fig. 5.4a and b). During this time, the cells travel a maximum distance of the order of 20 μ m. Both the critical distance and the correlation time are small compared to the dimensions of the system (droplet radius of 200 μ m and first hitting time mean around 100min).

Since our inquiry is not on the properties of T-cell motion but on the general accumulation dynamics on the spheroid, we believe there is a strong argument in favor of using a minimalistic model to represent cell motion. Since the timescales and length scales of this accumulation process are much larger than the experimental auto-correlation times it is justified to describe the accumulation process of the T-cells on the spheroid using brownian motion as a model for CTL motility in the gel.



FIGURE 5.4: Velocity (\mathbf{a}) and direction (\mathbf{b}) auto-correlations for the cell motion in Matrigel droplets.

5.2.3 Estimating the attachment and detachment parameters $(\lambda_{in} \text{ and } \lambda_{out})$

Using the image analysis pipeline described above we are able to detect the positions of the spheroids and of the CTLs in the droplets. This information allows us to measure the number of CTLs that are attached to the spheroid at each instant. The statistics of attachment and detachment can then be used to detect whether the attraction of the cells takes place through an active process or whether it is due to random chance.

Model description

We consider a system of N independent CTLs undergoing Brownian-like motility in the droplet. Their motion is isotropic and the statistical properties are independent of both time and space. Their diffusion coefficient in the Matrigel is noted D. The spheroid itself is modeled as a sphere of diameter a and the droplet has a volume V. Each CTL can be in one of two states: in the Matrigel or on the spheroid (Fig. 5.5). We consider that when a CTL encounters the spheroid it immediately attaches to it and remains attached until a detachment event occurs.

Condamin et al. [114] show that the probability of hitting the target for the first time at time t is:

$$p(t) = \frac{4\pi aD}{V} \exp\left(-\frac{4\pi aDt}{V}\right),\tag{5.1}$$

provided that the target size a is small compared to the closed volume dimension (i.e. $a \ll V^{1/3}$). In our experiments, with droplets of characteristic size L, we typically have $\frac{a}{L} \approx 0.2$, but for practical reasons we consider the distribution of first contact-times p described in equation 5.1 to be a reasonable approximation.

Integrating equation 5.1 between 0 and Δt we get the probability f of a cell hitting and attaching to the spheroid during an interval Δt :

$$f(\Delta t) = 1 - \exp\left(-\frac{4\pi a D \Delta t}{V}\right).$$
(5.2)

Each CTL in the Matrigel has a given probability of hitting and attaching to the spheroid during a time Δt described in equation 5.2. Let us first consider the cells attaching the target over a given time-interval Δt . The Matrigel droplet contains

a total of N cells and the results in equation 5.2 tell us that for a single cell, the probability of attaching between times 0 and Δt is: $f(\Delta t) = 1 - \exp(-\lambda_{in}\Delta t)$ where $\lambda_{in} = \frac{4\pi aD}{V}$ is the hitting rate. The probability g of multiple independent cells hitting the target during Δt follows a binomial distribution with f as parameter. Let us call c the number of contacts during Δt .

$$g(c > 0 | \Delta t, N) = 1 - \exp\left(-N\Delta t\lambda_{in}\right).$$
(5.3)

For small values of $N\lambda_{in}\Delta t$, we see that this simplifies to: $g(c > 0|\Delta t, N) \approx N\lambda_{in}\Delta t + O(N\lambda_{in}\Delta t)^2$. In effect, the condition $N\lambda_{in}\Delta t \ll 1$ means that the probability of a contact event happening during Δt is very small, and that two or more events happening during the same time-frame is negligible. At this stage, we do not know if this hypothesis is valid, however we can verify that the hypothesis is valid by using the estimated value for $\lambda_{in} \approx 5 \cdot 10^{-3} \text{ min}^{-1}$ (Fig. 5.5). For N = 10, this gives $N\lambda_{in}\Delta t \approx 5 \cdot 10^{-2}$, which is negligible compared to 1.

We can reproduce the same reasoning for the probability of having one or more cells leaving the spheroid - except that the probability is now function of the number of CTLs on the spheroid n and not of the number of CTLs in the gel N.



FIGURE 5.5: **a** Schematic of the stochastic accumulation model: CTLs can switch from the gel to the spheroid with different probabilities. $p_{in}(\Delta t)$ (conv. $p_{out}(\Delta t)$) is the probability for a cell to attach to (conv. detach from) the spheroid during a time interval Δt . Counting attachment and detachment events in the experiments allows us to infer the rates λ_{in} and λ_{out} , see section 5.2.3 for details. **b** Estimates for the attachment rates (λ_{in}), detachment rates (λ_{out}) and affinity ratio ($\lambda_{in}/\lambda_{out}$) for B16-WT (blue) and B16-OVA (red) cells. The box plots are obtained by using a bootstrapping method with 50 repetitions as described in the methods. **c** The affinity ratio as a function of the number of CTLs attached to the spheroid for B16 WT and B16 Ova spheroids. **d** Normalized attachment rate λ_{in} (white) and detachment rate λ_{out} (red) as a function of the number of CTLs attached to the spheroid. λ_{in} and λ_{out} are normalized by their mean values for 0 and 1 CTL on the spheroid respectively.

Estimating λ_{in} and λ_{out} from the data

As one can see in the model above, the attachment probability depends on a single parameter λ_{in} . We can model the detachment dynamics of the CTLs from the spheroid in a similar manner defined by parameter λ_{out} . λ_{in} and λ_{out} are the inverse of a characteristic time, and tell us how frequently a cell hits the spheroid or leaves it once attached. To retrieve the parameters from the evolution of the number of CTLs attached, we count the hitting and leaving events, then we get a hitting or leaving probability g.

Working backwards from equation 5.3, we extract the estimated value of λ , λ , from g by dividing the value by the number N of CTLs in the ECM or on the spheroid:

$$\tilde{\lambda}\Delta t = -\frac{\log\left(1-\tilde{g}\right)}{N} \tag{5.4}$$

and experimentally we have $\tilde{g} = \frac{\sum \text{hits}}{\sum \text{timepoints}}$. We can also specifically calculate a value of the hitting rate given a number of CTLs on the spheroid. We do so by estimating \tilde{g}_n the probability of hitting given that there are *n* CTLs on target and using the same formula as in 5.4 but only counting hits and time-points where the spheroid has *n* CTLs on it.

This procedure enables us to calculate the hitting (λ_{in}) and leaving (λ_{out}) parameters of our model. To estimate the uncertainty of our estimations we conduct a bootstrapping scheme. From a subset of the experimental accumulation plots (see Fig. 5.3b) we count the attachment and detachment events and estimate the values of λ_{in} and λ_{out} . We repeat the procedure 50 times selecting for each iteration 70% of the total number of spheroids in each condition (84 B16-OVA and 81 B16-WT spheroids). This gives us a distribution of the estimated parameters λ_{in} and λ_{out} represented in figure 5.5. From this set of parameters we retrieve a distribution for each parameter, which enables us to calculate the mean and the variance of each.

For more simplicity, from this point onward, we designate the experimental estimation of the arrival and leaving rates with λ_{in} and λ_{out} .

5.2.4 Estimation of the arrival and leaving rates λ_{in} and λ_{out} from experimental data

Evolution of λ_{in} and λ_{out} as a function of spheroid phenotype

The value of λ_{in} is found to be significantly higher in the B16-OVA spheroids compared the B16-WT spheroids, while the opposite is true for λ_{out} (Fig. 5.5b). These measurements indicate that the arrival rate of CTLs increases when the cells composing the spheroid express the cognate antigen recognized by the CTLs and that they stay attached for longer periods of time. Therefore the accumulation of CTLs on the spheroids is mediated by two independent phenomena, first the increase in arrival frequency and second by the decreased leaving frequency. The net effect of the attachment/detachment dynamics can then be summarized by the affinity ratio, $\lambda_{in}/\lambda_{out}$, which accounts for the net effective accumulation of CTLs on target. This ratio is found to be significantly higher with B16-OVA spheroids when compared to the B16-WT spheroids (Fig. 5.5b).

λ_{in} and λ_{out} vary with the number of CTLs on the spheroid

Evidence of a positive feedback loop for the attraction among the CTLs can be obtained by calculating the change of the affinity ratio as a function of the number of cells present on the spheroid. Indeed, the depth of the experimental data allows us to obtain a value of $\lambda_{in}/\lambda_{out}$ before the first contact and then successively track the change of this ratio after every contact, in each droplet (Fig. 5.5c). Again, the data for B16-OVA show a significant difference with the WT case. More interestingly, the higher the number of CTLs present on the B16-OVA spheroid, the higher the ratio, and thereby, the faster the accumulation rate: this is a hallmark of a positive feedback. Focusing on the values of λ_{in} and λ_{out} , we see that the increase in the accumulation rate is driven by an increase in λ_{in} for a constant value of λ_{out} (Fig. 5.5d). It demonstrates that the accelerated accumulation is mediated by the increasing attraction of the CTLs to the spheroid, which is triggered from the very first CTL attached to the spheroid (Fig. 5.5d).

5.3 A Non-Markovian accumulation model

In the accumulation procedure modeled in section 5.2, we hypothesize that the process is non-Markovian (i.e. that only the immediately previous step influences the attraction of CTLs). This is a strong assumption, and simulating the accumulation of CTLs on spheroids with the inferred parameters λ_{in} and λ_{out} we see that the average accumulation profile is very different in the Markovian and experimental case (Fig. 5.6). In particular, the simulated system rapidly (less than 2h) reaches a stable average number of CTLs on the spheroid after which the average accumulation doesn't rise significantly. This result is expected from a Markovian model as the CTL distribution on the spheroids stays stable after the mixing time, yet it fails to replicate the long term dynamics of the system.



FIGURE 5.6: Average number of CTLs on the spheroid as a function of time from experimental (black) and simulated (blue) data.

One of the immediate issues with the memory-less model is that it fails to account for past signals emitted by cells on the spheroid. It is known that CTLs produce several chemokines upon recognition of the target antigen [4] and these signals diffuse in the droplet. Previous studies on diffusion in droplets demonstrate the quick emergence of chemokine gradients within its closed environment [79]. Therefore, a CTL previously in contact with the spheroid will still influence the probability of another CTL hitting the spheroid many minutes later.

To correct for this effect we introduce the cumulative number of CTLs on the spheroid at time $t, \kappa(t) = \int_0^t n(s) ds$ (where n(t) is the number of CTLs on the spheroid at time t). A value of $\kappa(t) = 100$ means that the spheroid has had one cell on it for a hundred minutes, or ten cells for ten minutes. This value measures the cumulative contact time between CTLs and the spheroid. We then can use this variable to infer the values of λ as a function of κ in the same manner as described in section 5.2. The values of λ_{in} and λ_{out} as a function of κ are represented in figure 5.7a-b. We see that the value of λ_{in} increases linearly from κ equal 0 to 400. Also, the initial value of λ_{in} is very similar for both WT and the OVA spheroids, suggesting that before the first contacts the phenotype of the spheroid doesn't influence the initial attraction of CTLs. On the other hand, the value of λ_{out} decreases initially with κ before remaining stable, suggesting that the dynamics between the first contacts is different from the dynamics thereafter. These trends are made more explicit after normalizing the values of λ_{in} and λ_{out} by their average for $\kappa = 0$. For instance, for $\kappa = 300$, the hitting rate λ_{in} in six times higher than for $\kappa = 0$ (Fig. 5.7c), revealing the increased hitting frequency with increasing values of κ . On the other hand, it is difficult to conclude on the evolution of λ_{in} and λ_{out} in the case of WT spheroids, as very few spheroids ever reach high values of κ .

The measured evolution of λ_{in} and λ_{out} as a function of κ demonstrates that a positive feedback loop drives the CTL accumulation on spheroids. The previous conclusions remain true, although this refined model accounts for the long-term dynamics of the signaling between T cells in the droplet.

The inference of the values of λ_{in} and λ_{out} enables us to simulate the accumulation process on the spheroid. At each time step and for each spheroid we calculate the probability of a CTL attaching knowing the current number of CTLs already on the spheroid and the past spheroid history. We then do a Bernoulli draw using this probability and repeat the process, generating accumulation histories for each of the simulated spheroids. The accumulation simulated using the non-Markovian parameters replicates the long-term accumulation (Fig. 5.7d - right) similarly to what is observed in experiments. On the other hand, the Markovian model fails to recapitulate this long-term pattern (Fig. 5.7d - left). Averaging the simulated spheroids, we see that the long-term behavior of the non-Markovian models fits closely to the experimentally obtained accumulation curves (Fig. 5.7e).

The accelerated accumulation at the spheroid is best explained by long-term attraction to the spheroid. This attraction is reinforced with each successive contact and for each additional CTL on the spheroid, demonstrating the accumulation of the CTLs is mediated by a collective effect.

CHAPTER 5. CTL ACCUMULATION ON SPHEROIDS IN MICROFLUIDIC DROPLETS



FIGURE 5.7: **a** Evolution of λ_{in} as a function of κ for WT (blue) and OVA (red) expressing spheroids. **b** Idem for λ_{out} **c** Normalized evolution of λ_{in} and λ_{out} as a function of κ **d** Accumulation curves on spheroids with N = 11 CTLs in the droplet, the non-Markovian simulation (red) closely matches experimental results (black), contrary to the Markovian simulation (blue) **e** Average number of CTLs on the spheroid as a function of time. The non-Markovian average (red) closely matches the experimental results (black).

Key take-aways

Results In this chapter we have demonstrated that T cell accumulation on the tumor spheroid is driven by a positive-feedback loop: the more T cells on the tumor, the higher the probability for a new T cell to come and attach to the spheroid.

Literature comparison These results confirm recently published observations [4] that show a CCR5-mediated swarming of T-cells *in-vitro*, as well as *in-vivo* studies that report the accumulation of T-cells on targets [115, 116]. We show that these effects occur even for cell populations consisting of a few individuals and that a single contact can trigger the beginning of the positive feedback loop.

Next steps In the following section we go beyond the CTL accumulation to address the relationship between the accumulation of CTLs and their capacity to kill the B16 spheroids.

Chapter 6

A quantitative description of specific tumoroid killing by CTLs in droplets

In the previous chapter we have demonstrated that the accumulation of the T cells on the target tumors is antigen-specific and that it accelerates with each supplementary T cell attached to the spheroid. In the current chapter we will investigate the behaviors of the T cells once they have reached the spheroid.

6.1 Killing of B16 cells by CTLs is heterogeneous

After focusing on the behavior of the T-cells we now turn to the response of the cancer spheroids upon CTL accumulation. The time-lapse movies allow us to identify individual cell death events in the spheroids at the molecular level by detecting the activation of Caspase 3/7, which provides an early marker of apoptosis [117] (Fig. 6.1).



FIGURE 6.1: Representative sequence showing CTL positions on the spheroid, caspase 3/7 fluorescent death marker, and B16 cell fragmentation. The white arrow at 3h indicates the appearance of capsase signal next to a CTL. At 9h it represents a fragmented dead cell.

In the bright-field image we observe instances of rapid shedding of cellular material

CHAPTER 6. A QUANTITATIVE DESCRIPTION OF SPECIFIC TUMOROID KILLING BY CTLS IN DROPLETS

and debris from the spheroids (Fig. 6.1), which we refer to as "spheroid fragmentation". Combining this information with the position of CTLs relative to the spheroid, it is possible to record a detailed chronology of the key events taking place in each droplet by timing successive CTL contacts with the spheroid and the apparition of fragmentation events and Caspase 3/7 signals, as shown in Fig. 6.2.



FIGURE 6.2: Representative chronology showing the key events for a given spheroid interaction with CTLs: Contact times of CTLs on spheroids, detection of caspase 3/7 signal, detection of fragmentation events.

We observe that the timing of the first Caspase event post-CTL contact is well correlated with the first fragmentation event (Fig. 6.3a), indicating that the two observations are closely related. For this reason, we will hereafter use brightfield images to quantify spheroid killing, which simplifies the analysis pipeline and the experimental protocols. Indeed, the Caspase markers are particularly difficult to use since they bleach rapidly.

The timing of these killing events is highly variable, ranging from a few hours to beyond 24 hours. For some spheroids no fragmentation or Caspase events are observed over the course of an experiment. In the following analysis we will label "successful killing" the cases when the first fragmentation event is observed before t=14 h. The statistics of such events are summarized in Fig. 6.3b, which shows that 44% ($N_{OVA} = 84$ spheroids) of the OVA expressing spheroids are successfully killed by the CTLs. This contrasts with the B16-WT spheroids, where we do not observe any fragmentation events (Fig. 6.3b, XXX **Movie 3**). As such, we verify that the spheroid fragmentation is specific to spheroids expression the OVA cognate antigen.



FIGURE 6.3: **a** Time of first caspase signal vs. first observation of cell fragmentation. **b** Percentage of WT (black) and OVA (red) spheroids that show at least one fragmentation event in under 14 h. N equals 54 and 84 spheroids respectively.

6.2 Statistical observations on the spheroid fragmentation efficiency

As shown in section 6.1, we can record heterogeneous spheroid outcomes in the same chip; some spheroids fragment very fast (complete destruction at 8 hours), whereas others are left unscathed at 14 hours (see Fig. 6.1d). SHOW TIMELAPSES AS EX-AMPLES. Since the secondary droplet contains variable CTL numbers dependent on the initial cell concentration, this leads to a range of CTL numbers in the main droplet after droplet fusion (see Fig. 4.7). In addition to the number of CTLs in the droplet N, we record several other features: the first to the fourth contact times (t_1 to t_4 , indicating the moment at which the number of cells on the spheroid goes above 1 to 4 cells), the spheroid projected area and the maximum number of CTLs on the spheroid within the 14 hours of the experiment duration n.



FIGURE 6.4: Spheroid fate as a function of the number of CTLs in the droplet and the maximum recorded number of CTLs on the spheroid. The more CTLs the lower the survival chances of the spheroid expressing the cognate antigen.

We search for the statistically relevant variables predicting spheroid fragmentation with a generalized linear model [118]. The total observation sample size is of 96 events. This test enables us to study the influence and statistical power of each variable on visible spheroid fragmentation at 14 hours. From table 6.1 we see that only two variables have p-value below 0.05: the total number of CTLs in the droplet ($p_n = 0.019$, $\beta_n = 1.84$) and the maximum number of CTLs on the spheroid during the time-lapse ($p_N = 0.028$, $\beta_N = 1.56$). These two measures are very correlated (see Fig. 6.4); a high number of CTLs in the droplet increases the chances of having a high number of CTLs on the spheroid. The regression coefficients β_n and β_N are positive in both cases, confirming the positive correlation between these variables and the fragmentation probability. Interestingly, the other variables do not significantly predict the final spheroid state despite varying levels of correlation with spheroid death.

We have learned that the CTL concentration positively impacts spheroid fragmentation. However, from this data we don't know the mechanism by which the increased CTL concentration influences the spheroid killing. Looking at the time-resolved dynamics of fragmentation occurrences, we will demonstrate that cooperation between CTLs drives the fragmentation of the spheroids.

CHAPTER 6. A QUANTITATIVE DESCRIPTION OF SPECIFIC TUMOROID KILLING BY CTLS IN DROPLETS

Variable	Coeff. β	std err	\mathbf{Z}	р	[0.025]	0.975]
n	1.84	0.786	2.342	0.019	0.300	3.381
N	1.5567	0.707	2.203	0.028	0.172	2.942
Area	-0.2175	0.320	-0.681	0.496	-0.844	0.409
t_1	-0.5405	0.999	-0.541	0.588	-2.498	1.417
t_2	1.1746	1.732	0.678	0.498	-2.221	4.570
t_3	-2.3092	2.635	-0.877	0.381	-7.473	2.854
t_4	1.4984	2.191	0.684	0.494	-2.795	5.792

TABLE 6.1: Generalized Linear Model (GLM) [118] results (Number of samples is 96) from the statsmodel API [119]. n is the maximum recorded number of CTLs on the spheroid and N is the number of T-cells detected in the droplet, Area is the spheroid surface and t_1 to t_4 are the first times at which 1 to 4 CTLs are detected on the spheroid surface.

6.3 Tumor spheroid killing involves collective effects

6.3.1 Qualitative indications of collective killing

We now consider the relationship between the spatiotemporal dynamics of the CTLs and the tumor spheroid outcomes (successful or unsuccessful killing). An indication of the relevance of this link is obtained first by observing that the CTL accumulation rates are faster on the spheroids that display fragmentation than in the opposite case, both in absolute numbers (Fig. 6.5a) and as a fraction of the total number of cells per droplet (Fig. 6.5b). This indicates that a faster accumulation is correlated with efficient killing.



FIGURE 6.5: **a** Number of CTLs on the spheroid as a function of spheroid fate and as a function of time after the first contact. Fragmenting spheroids (red) see faster accumulation than non-fragmenting spheroids (black) and WT spheroids (blue). **b** Fraction (i.e. number of CTLs on the spheroid normalized by the number of CTLs in the droplet) of CTLs on the spheroid as a function of time after contact. Colors same as in (**a**).

Moreover observing the CTLs on the spheroid reveals that fragmentation events are associated with the presence of several CTLs in the vicinity of the spheroid (Fig. 6.6a,b). This local effect is quantified by counting the number of T cells present in within a 30 μ m radius of the first cell fragmentation event, giving a mean number of 3.4 cells (median at 3 cells), with fragmentation very rarely occurring with only one cell present in the neighborhood of the fragmentation event (Fig. 6.6b). Indeed, CTLs sometimes appear to besiege a salient B16 cell, causing it to burst after a few minutes of attack. This suggests that the CTLs tend to cluster at particular sites on the spheroid and that their clustering enhances their ability to induce spheroid fragmentation (Fig. 6.6c).



FIGURE 6.6: **a** Two representative images showing CTL clustering on the spheroid during first fragmentation event. The white circles have a radius of 30μ m around fragmenting cell. **b** Distribution of the CTL numbers within a radius30µmaround the fragmentation areas (N=31). **c** Sketch summarizing the observed behavior: CTL (green) clustering leads to tumor cell killing and fragmentation(red).

6.3.2 Estimating the collective fragmentation rate Γ_{frag}

The above observations can be analyzed in greater depth to demonstrate that the cumulative effect of the CTLs enables the killing of tumor cells. If we consider a single time-interval of length Δt , the probability of a fragmentation event happening is $\Gamma_{\rm frag}(n)\Delta t$. Conversely, the probability of the spheroid remaining intact is $1-\Gamma_{\rm frag}(n)\Delta t$ (Fig. 6.7a). Over the course of the experiment, the fragmentation events can happen at different moments as is schematically represented in figure 6.7b.

At each time frame, the CTLs on the spheroid can either stay on the spheroid and do nothing, leave the spheroid or cause a fragmentation event. Conversely, CTLs in the gel can either stay in the gel or attach to the spheroid. We consider the killing to be independent of the attachment/detachment process and only to depend on the number of CTLs on the spheroid at any given time. From the point of view of the spheroid, we can therefore model the killing process as a Bernoulli process where the probability of seeing a fragmentation event only depends on the number n of CTLs on it at that time, and of Γ_{frag} . An implicit assumption of this model is that the process is Markovian: the probability of fragmenting during a time-interval only depends on the number of CTLs on the spheroid at the time and not on the previous CTL actions.

Let us call Γ_{frag} the probability that the *n* CTL(s) on the spheroid cause a fragmentation during a single time-interval Δt . We can infer Γ_{frag} as a function of the number of CTLs on the spheroid from the experimental data. The probability that the first fragmentation event happens after exactly *k* time-steps after the experiment start is a function of Γ_{frag} . The process is memory-less so this yields:

$$P(T_{\text{frag}} = k | n, \Gamma_{\text{frag}}) = \Gamma_{\text{frag}}(n_k) \prod_{i}^{k-1} \left(1 - \Gamma_{\text{frag}}(n_i)\right).$$
(6.1)



FIGURE 6.7: **CTL number and collective behavior determine probability of killing. a** During a time-interval Δt , a spheroid with *n* CTLs attached to it has a probability $\Gamma_{\text{frag}}\Delta t$ of fragmenting. **b** Fragmentation events can happen at different times and with a different number of CTLs on the spheroid (illustration plot) **c** The fragmentation rate Γ_{frag} can be modeled as the result of independent CTLs interacting with the spheroid, with an individual fragmentation rate per CTL worth ρ . Conversely, the fragmentation rate can be viewed as the result of a collaborative process. **d** Estimates of Γ_{frag} as a function of the number of CTLs on the spheroid *n*. Experimental measurements (N = 84, black dots) are fitted with an exponential (dashed red line) compared with the results of the independent CTL model (dashed blue line).

Since each well is independant, we can also estimate the probability of observing the set of fragmentation times $\{T_{\text{frag}}\}_i$:

$$P(\{T_{\text{frag}}\}_i | \{n\}_i, \Gamma_{\text{frag}}) = \prod_i P(T_{\text{frag},i} | n_i, \Gamma_{\text{frag}}).$$
(6.2)

Now we can use Maximum Likelihood estimation (ML) to extract the value of Γ_{frag} that maximizes the probability to observe the experimental fragmentation times given the experimentally observed time-series of n. In short, we calculate Γ_{frag} :

$$\Gamma_{\text{frag}} = \arg \max_{\Gamma_{\text{frag}}} P(\{T_{\text{frag}}\}_i | \{n\}_i, \Gamma).$$
(6.3)

After some algebra this yields the following value for $\Gamma_{\text{frag}}(n)$:

$$\Gamma_{\rm frag}(n) = \frac{N_{\rm Death,n}}{N_{\rm Death,n} + N_n},\tag{6.4}$$

where $N_{\text{Death},n}$ is the number of times we observe a fragmentation event happening with exactly *n* CTLs on the spheroid, and N_n the total number of times *n* CTLs are on the spheroid before any fragmentation event. We can verify that if there never is any death observed with *n* CTLs on the spheroid, then $N_{\text{Death},n} = 0$ and $\Gamma_{\text{frag}}(n) = 0$ too. Reversely, if there is a fragmentation event as soon as *n* cells are on the spheroid, then $N_n = 0$ and $N_{\text{Death},n} > 0$, which gives $\Gamma_{\text{frag}}(n) = 1$.

We have shown that we can infer the evolution of Γ_{frag} from the experimentally observed fragmentation times and as a function of the number of CTLs on the spheroid n, the number of fragmentation occurrences with exactly n CTLs on the spheroid $N_{\text{Death},n}$ and the number of instances with n CTLs on the spheroid before the first fragmentation event N_n : $\Gamma_{\text{frag}}(n) = \frac{N_{\text{Death},n}}{N_{\text{Death},n}+N_n}$.

We see that the fragmentation probability at each time-point increases with the number of CTLs on the spheroid. Indeed, the experimental evolution of $\Gamma_{\text{frag}}(n)$ is well-fitted by an exponential (Fig. 6.7d). In the next section, we will show that this fragmentation probability evolution is incompatible with a model where the CTLs are independent of each other and where the fragmentation of the spheroid is a memory-less process.

6.3.3 Probabilistic modeling of the independent killing of B16 cells on the spheroid by CTLs

A naive model of CTL killing behavior

In figure 6.7d we show that the probability of a spheroid to undergo a fragmentation depends on the number of CTLs on the spheroid surface. However, we do not know if this increase is the result of the accumulation of independent random events at each time-interval or if it is the signature of cooperation between T-cells. To address this question, we construct a model of spheroid fragmentation where the CTLs are independent of each other and of the spheroid state at the previous time-point (i.e. the system is Markovian). From this we can calculate the expected evolution of $\Gamma_{\rm frag}$ differs from the theoretical one, we can reject the hypothesis, demonstrating that the system either is non-Markovian and/or that the CTLs aren't independent of each other.

We consider that the probability for a given T-cell to cause fragmentation is independent of the presence of the other T-cells. In quantitative terms, the probability of a fragmentation event occurring within the independent cell hypothesis is given by:

$$\Gamma_{\rm frag}(n) = 1 - (1 - \rho)^n,$$
(6.5)

where the total number of CTLs on the spheroid is given by n and ρ is the probability of a single CTL to kill during a time-interval Δt . Fitting equation (6.5) to the experimental measurements we see that the model doesn't accurately reflect experimental results (Fig. 6.7d). This means that we can confidently reject the hypothesis that the CTLs are independent of each other.

Accounting for CTL heterogeneity

We can improve the model to demonstrate that CTL heterogeneity doesn't affect the result. Indeed, it has been hypothesized that the heterogeneous nature of the CTL population can affect the efficacy of tumor cell killing at the single cell level [11]. In order to test this hypothesis we now model the CTL population as composed of independent cells, but the probability of being a fragmentation-causing CTL k is now itself drawn from a probability distribution \mathcal{K} of probability density f. The average value of k, $\langle k \rangle$ is given by $\langle k \rangle = \int k f(k) dk$. We choose to not specify \mathcal{K} to preserve generality. Then equation 6.5 becomes:

$$\Gamma_{\text{frag}}(n) = 1 - \mathbb{E}\left[\prod_{i=1}^{n} (1-k_i)\right]$$
$$= 1 - \left(\int (1-k) f(k) dk\right)^n$$
$$= 1 - (1 - \langle k \rangle)^n.$$
 (6.6)

The result is similar to that in equation 6.5. Thus, the expected fragmentation probability is parabolic whatever the heterogeneity profile of the CTL population provided that we compute the statistics over a wide-enough sample. Therefore the initial CTL population heterogeneity cannot explain the fragmentation profile recovered experimentally.

The experimental measurements are compared with an independent probabilistic model, in contrast to the hypothesis of cooperating CTLs (Fig. 6.7c). This hypothesis leads to the parabolic probability of killing, as shown by the blue dashed line on figure 6.7d. The poor match with the experimental data, which is well-fitted by an exponentially increasing value of Γ_{frag} (red dashed line), demonstrates that a model of independent Markovian interactions between the CTLs and the target cells fails to capture the underlying mechanism of spheroid fragmentation.

6.4 Long and short-range interactions combine to determine probability of CTLs to kill a tumor spheroid

6.4.1 Simulating the spheroid fate

The probability for successful killing to occur for any particular spheroid can now be explained as a combination of the effect of CTL accumulation on the spheroid and their collaborative killing behavior. This probability is a function of the total number of CTLs in the vicinity of the spheroid, which also correlates with the maximum number observed on the spheroid (Fig 6.4). It can be modeled using the parameters computed above for the two collective behaviors (Fig 6.8a). We know that due to a positive feed-back loop, CTLs exhibit different accumulation rates depending on the number of CTLs on the target spheroid. We have estimated the experimental distribution of the hitting probabilities in section 5.2.3. We also know that a certain number of CTLs on the spheroid are necessary for a high killing probability (see Fig. 6.7).

Combining these two pieces of information we run an algorithm which reproduces the accumulation process and the spheroid fragmentation. We simulate the spheroid fate as a branching process where the evolution of the number of CTLs on the spheroid at each time point depends on the experimentally derived parameters λ_{in} and λ_{out} described in Fig. 5.3. Simultaneously, at each time-step the spheroid has a probability of fragmenting depending on the number of CTLs on it and Γ_{frag} as described in Fig. 6.7.

6.4.2 Comparison with experimental results

This computation is performed for a number of CTLs per droplet ranging from 0 to 20 cells and repeated 50 times. This allows us to first compare the experimental first fragmentation times (black dots) with the simulated ones (blue line). In both cases the first fragmentation time decreases with the number of CTLs in the droplet and the simulated values are close to the experimentally observed ones (Fig. 6.8c). We then compare the simulated fragmentation probability curve as a function of the number of CTLs in the droplet to the experimental data (Fig. 6.8d). Not only does the simulation confirm the key role of CTL concentration in initiating spheroid fragmentation, the close match between simulated and experimental measurements (with no free parameters) indicates that the spheroid fragmentation process is well described from this succession of two mechanisms: cooperative recruitment and collective killing at the spheroid site.

Key take-aways

Results We show that tumor cell killing by T cells is mediated by collective effects: the spheroid fragmentation probability increases exponentially with the number of T cells clustered on it. This recapitulates *in vitro* [10] and *in vivo* [11] T cell behavior reported in the literature.

Simulation Combining the collective accumulation and fragmentation effects in a simulation, we show that we can reproduce experimental properties such as the fragmentation probability of the spheroid as a function of the T cell concentration in the droplet and the first fragmentation times.
CHAPTER 6. A QUANTITATIVE DESCRIPTION OF SPECIFIC TUMOROID KILLING BY CTLS IN DROPLETS



FIGURE 6.8: Combining short and long-range interactions to predict spheroid fate. a The concentration-dependent killing is modeled as the result of two complementary mechanisms: long-distance cooperative attraction of CTLs to the target site and local killing cooperation on the spheroid. b The evolution of the spheroid fate in the droplets can be modeled as a branching process: at each time step the spheroid can fragment with a probability $\Gamma_{\rm frag}\Delta t$ or not, and CTLs can either attach to or detach from the spheroid. c Simulating the spheroid fate using the model in (c), we recapitulate the experimentally derived first-fragmentation times. d Experimental (black dots) and simulated (blue line) spheroid fragmentation probability as a function of the number of CTLs in the droplet. Shaded area represents the 95% confidence interval of simulated data.

Chapter 7 Conclusions

This work introduces a new paradigm for extracting biological information from *in-vitro* experiments, by treating the parallel realizations as "Monte-Carlo experiments" that reach different outcomes in a probabilistic way. This contrasts with existing microfluidic models for cancer-immune interactions, which treat each chip as a single experiment and use traditional biological measurement techniques [53]. By comparison the droplet format provides several unique features, including the ability to merge many droplet pairs at a well-defined time, thus providing a common starting time of the parallel experiments [120]. Moreover, the encapsulation within droplets allows the conditions in each of the parallel experiments to be well controlled, thus allowing for massively multiplexed experiments on a single device.

Here these technical advantages are associated with probabilistic modeling to infer key biological information about the ability of CTLs to sense and respond to the tumor, by relating the spatiotemporal dynamics of the CTLs with the outcome for the tumor spheroid. We find that the first CTL-tumor cell contact, which occurs randomly, triggers a positive feedback loop that leads to an accelerated accumulation of CTLs on the spheroid. Later, CTLs form clusters on the spheroid that enhance their ability to kill the target cells, leading to tumor rejection.

Several mechanisms may account for the collaborative CTL accumulation and killing. Chemokines that are both sensed and produced by T cells have the ability to drive their swarming behaviour [4]. Cooperative killing, in which multiple sublethal cytotoxic hits synergize to induce target cell killing, has been previously described in the context of viral infection [11] and tumor development [10]. Alternatively, initial CTL-tumor cell interactions may facilitate tumor destruction by increasing MHC class I expression through IFN- γ production and diffusion in the tumor microenvironment for example [3,121]. As illustrated here, our approach helps support and generate new hypotheses that can be subsequently dissected at the molecular level.

Looking ahead, the Multiscale Immuno-Oncology on-Chip System (MIOCS) can now be generalized to include several immune cell types and more realistic tumor models in each droplet. Here again the spatiotemporal resolution and *Monte-Carlo* approach will be fundamental to understand the causality of the interactions and the effect of 3D geometry. Finally, working with patient-derived organoids [7,122] will have important implications for personalized medicine.

Part II

Network-based tissue analysis

Chapter 8 Introduction

Networks are everywhere. There are physical networks, such as power grids, the web [123] or Internet routers. Cities can be represented as weighted networks where each street is an edge. There are social networks: telephone calls, Facebook, dating or sexual contact networks [124]. Biological networks such as food-webs, protein-interaction networks [125,126]. And today neural networks power much of the artificial intelligence revolution [127,128]. Networks come in all sorts of shapes and colors.

The versatility and success of networks as a way of representing real-world systems is due to the simplicity of the network description; they enable a minimalistic representation of complex systems [129]. They also allow for an efficient embedding of relational information and accommodate a large spectrum of attribute types. The mathematical framework that has emerged from network science allows for the use of tools coming from a large range of fields, from statistical physics [130] to applied mathematics and computer science. In the following paragraphs we will give a rapid overview of networks, what they are and some of the uses that they've found in biology.

8.1 Networks in biology

Advances in technologies for biological measurements have led to a significant increase in the amount and in the quality of data available to researchers. This has made it possible to adopt quantitative approaches to studying biological systems at an unprecedented scale, and networks have proven to be a powerful way to model and make sense of these new streams of data to study [131–133].

Large-scale maps of protein interaction networks and transcriptional regulatory networks have been constructed for a number of organisms [125, 126] (Fig. 9.2). Networks also play a fundamental role in the spreading of epidemics, as has been demonstrated repeatedly [124, 134, 135] due to the interplay between epidemic dynamics and network structure. For instance, the unequal distribution of the numbers of contacts within a population significantly reduce the epidemic threshold, making the outbreak of epidemics more likely in networks with such a structure. Figure 8.1 shows the propagation of a simulated disease on a small network (*personal unpublished data*): susceptible patients on isolated leaves are protected from infection by highly connected recovered patients. Similarly, these models also show the need to target nodes of high degree to improve the efficacy of immunization schemes [124].



FIGURE 8.1: Example of an epidemic propagating on an example graph. In red infected nodes, in green the susceptible ones and in blue are the recovered nodes. The epidemic propagation dynamic is influenced by the graph structure with isolated leaves less probable to be infected than highly-connected nodes.

8.2 Spatial networks in biology

For many problems in network science, the topology itself is important: for example to assess probable gene functions from interactions pattern [126,133] or to understand the structure of regulatory networks in various bacterial species [136]. In these networks, the spatial distribution of the nodes is often irrelevant, and this can lead to very unequal degree distributions, typically power laws [137]. Similar problems include the distribution of sexual contacts in social networks [124], or to the distribution of links of web pages [123].

However, for many biological networks, space is extremely relevant: cell attachment, communication and interaction obviously depend on distance. For example cellular differentiation in fly embryos depend on minute gradients in specific cytokines [138]. Similarly the spatial organization of mesenchymal bodies has been shown to depend on preferential attachment between different cell types [13]. In contrast with the previously described objects, these networks do not contain a lot of topological variation. Indeed, because of the dense structure of the tissue it is unlikely that a single node will have a large share of all connections with the majority of the nodes having only a few edges as is possible in small-world networks. We expect a homogeneous distribution of the number of edges per node, but with potentially heterogeneities in the distribution of node attributes within the network. Analyzing such biological networks requires the development of tools to generate the network representation and information relevant for biologists.

Luckily, improvements in imaging techniques have given access to new troves of data. Analysis of cell positions in complex tissues or of protein expression at singlecell resolution and in spatial contexts can reveal unique insights into the organization of tissues and organs. The recent improvements in biological imaging have generated spectacular results. For example, the precise distribution of cells in the bone marrow (BM) has been described by Coutu et al. [139,140]. Also using fluorescence microscopy, the development of the new CODEX technology [16, 141] has made it possible to simultaneously image close to 40 markers in tissues. Additionally, new methods such as spatial transcriptomics have enabled researchers to link genetic expression and spatial information in a single format [15, 142]. Example projects include deciphering the structure of tissues after seqFISH [143] or smFISH analysis [144]. The combined methods above have revealed new insights such as gene expression gradients [143] or the preferential interaction between specific cell-types in the tissue [145].

8.3 Why should a biologist work on spatial networks?

Tissue organization occurs on at least two levels: on the one hand macroscopic regions that are typically observed at the 100μ m to 1mm scale, then on the other hand individual cells are studied with techniques such as FACS, yielding single-cell information but without any spatial context. This limits the ability to capture the tissue-level processes that emerge from such spatial relationships and calls for a multiscale framework to integrate the data from new sources.

Indeed, cells interact on multiple length-scales through direct contact of surfacebound receptors and ligands and mechanical effects, as well through long-range mechanisms such as soluble factors. These communication events can't usually be directly observed but are critical to understand emergent phenomena in tissue niches. Combining the progress of cell imaging with a network-based approach to describe the tissue, we propose a novel framework to study biological tissues in 3D. This consists in building a "network twin" of the tissue and using this image to extract statistical information on the sample organization. Using networks to make sense of spatially determined data is already a classical feature of geospatial data science with long-established libraries such as PySAL and has recently been implemented on 2D microscopy images [146].

In the following chapters, we will describe a novel package, Griottes, to build a 3D network representation of biological tissues. We will then show how this package can be applied to analyze the distribution of hematopoietic cells in the fetal liver.

Chapter 9 Building Griottes

The condensed version of chapter 9 will soon be available as:

Ronteix G., Bonnet V., Sart S., Esposito E. and Baroud C. (2021). *Griottes: a* generalist tool for network generation from segmented tissue images.

9.1 A short introduction to graphs

A graph, or a network, is a set of objects which are related. Each object is described by a node (or vertex), and each node can have several attributes [137]. For example, when studying the social network in a school, each node can describe each student, and the attributes can be the student's gender, height or any other feature of interest. These nodes are connected by edges or links that represent the relationship between nodes. The edges can be directed representing an asymmetrical relationship (for example a paper that cites another paper, Fig. 9.1a), or undirected (such as a relationship between two people, Fig. 9.1b).

Let us first introduce the graph formalism that will be used extensively in the following chapter. A graph G can be described as a pair G = (V, E) where V and E represent the vertexes and the edges. The number of edges leading to each node *i* is called degree and written k_i . For example, in figure 9.1b, $k_i = 3$ whereas $k_l = 1$.



FIGURE 9.1: **a.** Example of a directed graph. The edges leading to and from the nodes are oriented, enabling the representation of non-symmetric relations between the nodes. **b.** Example of an undirected graph.

CHAPTER 9. BUILDING GRIOTTES

The study of the topology of real-world networks has grown considerably these past years with the formidable rise of computer networks and social networks. Understanding how nodes are connected can be of critical importance. For example, when modeling epidemics on networks the structure of the said network can have a key impact on the epidemic dynamics (Fig. 9.2c) [135, 147]. In the coming couple of lines we will describe examples of networks in biological systems.



FIGURE 9.2: **a.** A famous social network that even made it into the theatres. **b.** Saccharomyces protein interaction network (from [125]) **c.** The interconnection of airports across countries can affect the spreading of an epidemic [135]

The sheer number of use-cases, from the social sciences (Fig. 9.2a), to yeast metabolism (Fig. 9.2b) and epidemiology (Fig. 9.2c) reveals the versatility of the network representation. In the following section I will present the motivation for developing a graph-generating library from microscopy images.

9.2 Motivation

Characterization of the spatial context of cells and linking local phenotype to the environmental cell niche is of critical significance to link global tissue function to individual cell behavior. Cell organization occurs across scales, and new frameworks and tools are necessary to connect them. Most analysis conducted today, be it of the cell transcriptome through RNAseq or of the proteome through FACS for example, requires dissociating individual cells in the tissue, thereby losing the spatial context of individual cells [142]. To preserve this information, new methods have been proposed such as spatial transcriptomics. The spectacular promise of these technologies have not gone unnoticed and spatial transcriptomics was named "method of the year" 2020 by *Nature Methods* [15]. This has led to several of publications to quantify and make sense of the treasure-trove of genomic data that is produced by the new methods [144, 148].

On the other hand, microscopy techniques have made spectacular progress in the past decade. New techniques such as light sheet imaging have made it possible to image samples in 3D with single-cell precision [149]. And new methods for marking cells have made it possible to multiply the number of targets in each image [139, 140]. For example, the development of methods such as CODEX multiplexed imaging [16] have enabled the multiplexed imaging of up to 40 markers simultaneously (Fig. 9.3a-b). New techniques have also been developed to multiply the number of markers that could be imaged in 3D as shown in figure 9.3c-e, which has enabled researchers to precisely quantify the composition and structure of the bone marrow.

Yet, I believe that a tool is missing to bridge the gap between the acquisition of tissue images and their use for biological inquiries. Although the imaging and image segmentation techniques have improved, quantifying the tissue properties remains taxing. This calls for the creation of imaging and image analysis pipelines capable of extracting single-cell information and combining the phenotypic and the spatial information. Several teams have sought to develop such tools extracting phenotypic information from confocal images [145, 150] or from CODEX arrays [141]. However, these tools are integrated in complex programs and don't adapt well to different image types. Here we propose a standardized framework for extracting information from segmented images.

9.3 Program pipeline

For the purpose of analyzing tissue images, we have developed a network generation library called **Griottes**. This is a one-stop tool that extracts information from segmented cells (such as their position, fluorescence intensity levels etc.) and constructs the network image of a tissue.

The general structure of the image analysis pipeline is organised in two independent steps shown in green in figure 9.5. The following paragraphs describe the input/output properties as well as a broad outline of the operations conducted at each step. A more detailed explanation concerning the cell attribute measurement process and the network construction will be given in sections 9.4.1 and 9.4.2.

Input Griottes takes a labeled 2D or 3D image as an input (Fig. 9.4). The labeled image has at least one channel composed of a matrix filled with the area occupied by each cell filled with an integer specific to each cell. We chose to not integrate a cell segmentation method as the segmentation method strongly depends on the individual image characteristics and the framework already in use. This also makes it possible to analyze images coming from different sources, such as imaging mass cytometry where specific segmentation methods have been proposed already [151].



FIGURE 9.3: **a.** Presentation of the CODEX workflow (see [141]). **b.** Seven-color images of select tissue cores from patient samples using the CODEX methodology described in (a.). **c.** Multiplexed immunofluorescent staining of an adult mouse femur. Taken from [140] **d.** and **e.** staining and detection of rare cell types in 3D bone marrow images. Taken from [139].

Individual attribute measurements From the individual segmented cells, Griottes can extract the essential geometric properties of the cells such as their eccentricity, as well as their fluorescence properties. From this, Griottes generates a pandas DataFrame object [110] as an intermediate output containing all the individual cell information. This table contains the spatial position of every cell in the tissue, as well as the eventual fluorescence and geometrical features requested by the user. Generating the table from an image typically takes of the order 10 to 20 seconds for a 20GB-big image containing up to 50.000 cells.



FIGURE 9.4: Schematic of a labeled image. Pixels of value 0 represent the background, and different cells are labeled with distinct integers.

Network construction From the individual cell properties **Griottes** builds a spatially contextualized graph where each node can be populated with cell attributes depending on the will of the experimenter. The links connect cells that are deemed in contact with each other according to the heuristic of the chosen construction method. These construction methods can vary depending on the uses and will be described below.

It is possible to skip the first two steps and provide cell positions in the tissue to construct the network representation of the sample. This can be useful for data extracted from other scientific software such as Imaris.

Output The returned object is a graph where the previously measured properties are entered as node attributes, each node in the network mapping to a given cell in the tissue. Currently the preferred format is a NetworkX graph object [152] since it's one of the most popular graph manipulation libraries in the Python community. More formats will be added soon.

9.4 From image to network

9.4.1 Extracting single-cell information from the image

The most immediate information to access are the cell positions within the tissue. Beyond this feature, we can measure the geometric properties of the cells and their fluorescence levels. In the following paragraphs we describe how these properties are gathered in Griottes.

Geometrical property measurements This can be of particular interest as it is known that cell morphology can be a significant feature of tissues and their characterization can shed light on morphogenesis or mechanical constraints within the tissue (see Dolega et al. [153]). To be able to readily include the morphological information as a feature of the cells in the network, we added an option to measure the geometrical properties of the cell masks in Griottes. This includes extracting the mask volume, eccentricity and orientation in space.



FIGURE 9.5: Image analysis pipeline. From a 2D or 3D segmented image Griottes enables the user to extract single cell data and construct a NetworkX graph object. In red are the program outputs and in green the different steps in the protocol.

To measure the eccentricity and the orientation in space, we conduct a principle component analysis (PCA) of the pixels belonging to the cell mask. This generates a set of vectors $\{\mathbf{e}_i\}_i$ and eigenvalues $\{\mathcal{E}_i\}_i$ for *i* ranging from 2 to 3 depending on whether the image is 2D or 3D. From the principle component vector \mathbf{e}_1 we extract the cell orientation and from the eigenvalues we can estimate the eccentricity *e* in 2D and 3D (with equations 9.1 and 9.2 respectively).

$$e = \frac{\mathcal{E}_1}{\mathcal{E}_2} \tag{9.1}$$

$$e = \frac{\mathcal{E}_1}{\sqrt{\mathcal{E}_2^2 + \mathcal{E}_3^2}} \tag{9.2}$$

Fluorescence measurement techniques To characterize the cell type it is necessary to develop a protocol to measure the cell fluorescence. The ideal method would be to have access to the whole cell volume in order to precisely account for markers staining the membranes or the cytoplasm. This can be done by adding membrane markers in order to precisely distinguish individual cells from each other [16,140]. When a large number of stains are available, this method is particularly precise. However if the number of channels available is limited then consuming a marker to stain the membranes is experimentally costly as it limits the amount of information that is possible to capture. One possibility, chosen by Coutu et al. is to limit the analysis to a subset of cells [140]; in their case they decided to only study the c-Kit+ sub-population, where the c-Kit staining made the full segmentation of cellular membranes possible.

We developed three distinct methods to measure the fluorescent properties of the labeled cells. The fastest method relies on the **regionprops** routine from the Scikit



FIGURE 9.6: Different intensity measurement methods exist to extract information from the fluorescence images. The nuclei themselves can serve as masks (a), or one can measure the pixel intensity within a sphere of any given radius (b). A more sophisticated method involves creating the Voronoi tesselation of the image and extracting the pixel intensities within each individual voxel (c).

Image library [109]. It is particularly useful for larger images as it is fast and efficient, however we only access the fluorescence levels inside the labeled cell mask. It is therefore particularly reliable the case of images where the cell masks are built from membrane stainings and the full volume of the cell is recovered.

If on the other hand only the nucleus is stained, then only the information from the pixels inside the nucleus mask is taken into account as shown in figure 9.6a. Despite the theoretical limitations of only measuring the intensity inside the nuclei mask, in practice the results on experimental images are satisfying and make it possible to automatically classify cells in a way that fits with visual classification as shown in figure 9.7.



FIGURE 9.7: Multi-channel images of a fetal liver slice. The scattered dots show the positions of detected cells, their color represents the cell classification. A cell is detected as being positive (+) if the average intensity of the pixels within the mask are above a certain threshold. More details on the cell classification procedure are available in section 10.3.4.

We also developed alternative ways of measuring the fluorescence information beyond looking at the pixels within the nuclei mask: the second method relies on measuring the average intensity within a given radius from the cell center (see figure 9.6b.). Finally the third one builds a Voronoi tesselation of the volume and extracts the fluorescence from within the cell's voxel as illustrated in figure 9.6c. The analysis time is a serious limitation of the two latter methods as the images start to gain in size since

CHAPTER 9. BUILDING GRIOTTES

the algorithm is linear in the cell number. In practice, for images above a couple of hundred of MBs, the computation time is prohibitive.

9.4.2 Building the network

Once the information on the cell properties is extracted from the image, we build the connectivity graph binding neighboring cells together. Three methods are already implemented to build the network: the geometric graph method, the voronoi method and the neighborhood method.

Geometric graph construction To build a geometric graph we connect cells distant below a given radius [17]. It has been argued that this connection method is particularly relevant when cell-cell communication is driven by chemical cues [154]. However, as shown in figures 9.8a and 9.8c, this method is less reliable if cell-cell interactions are contact mediated as it can connect cells that have a third cell separating them.



FIGURE 9.8: Different methods for building a network lead to different network structures: **a.** Illustration of the rule to build a geometrical network. The cells within a certain radius of a given node are connected to it. **b.** The Voronoi network construction rule connects cells irrespective of their distance from each other but excludes certain connections based on geometrical constraints of the network. **c.** Combining the two rules leads to the construction of a network that is more physiologically relevant than in the two other methods. Connections between cells that are unrealistically long are eliminated (dotted blue line) and the connections of cells behind other cells are removed (dotted red line).

Voronoi graph construction In the second method we use a Delauney-triangulation based connection condition where links above a certain threshold are broken to avoid the physiological incoherence that the Delauney triangulation can generate in real tissues (see 9.8c). This is a robust state-of-the-art method widely used to build biologically relevant networks in many different contexts [146, 148, 154]. Although the Voronoi tessellation method is known to not perfectly replicate tissue morphology in 2D [155], it provides a reasonable connectivity estimation.

Contact-based graph construction We can use contacting neighbourhoods to build the graph. This method is particularly pertinent if the membrane masks are reliable, since they give access to the true cell-cell contacts. Contrary to the Delaunay

and geometric graphs described immediately above, this method is more robust when the cells have complex shapes.

9.4.3 Comparing network construction methods



FIGURE 9.9: From the segmented cells (marked with a blue dot), it is possible to construct different tissue networks

In practice, the different network construction methods described in section 9.4.2 produce networks with very different properties. To illustrate these differences, we will apply the **Griottes** pipeline to an image of the zebrafish telencephalon generously provided by Nicolas Dray from the Zebrasfish and Neurogenetics unit at the Pasteur Institute. This region of the zebrafish brain possesses an epithelial-like structure, and it is hypothesized that the distribution of the cells and their local environment influences the fate of the cells.

The advantage of this data is that zebrafish lines that have fluorescently labeled ZO1 exist. This protein is membrane-bound and so it is possible to acquire high-quality images of the cell membranes. This will make it possible to compare the three network-construction methods above.

The geometric network connects all cells where the center of mass of the masks are separated by less than a critical distance. In this case, the masks for each cell cover the whole surface delineated by the membranes, but in images where the nuclei are stained the nodes are positioned in the center of the nuclei. This is a minor point, but can influence the network topology when cells have elongated shapes.

The properties of geometric networks is a standard topic in network science [17] and it has been mentioned as being relevant in systems where cell-cell communication is dominated by chemical signals [154]. Looking at the distribution of the links in figure 9.9, we see that a dense mesh of edges is established in regions of small sizes and that big cells are loosely connected. Given the structure of the tissue, this distribution

is unsatisfactory. Using the contact-based network as a reference, the visual conclusion is confirmed quantitatively by comparing the degree of each cell cell in the geometric network with its degree in the contact-based network (Fig. 9.10). The resulting distribution is particularly broad, with a standard deviation of 2.7 links, reflecting the realistically dense mesh of cell-cell contacts in dense areas and the absence of links between big cells in the tissue.



FIGURE 9.10: (top) Cell degree distribution for different network construction methods. (bottom) Distribution of degree differences between the contact-based and the Voronoi/geometric method.

An alternative is to construct the network thanks to the Delauney triangulation rule. The network produced is complementary to the Voronoi tessellation, which here mimics the cell boundaries. Comparing the Voronoi network to the contact-based one, we see few visible differences. The degree of each cell in the contact-based network k_{Contact} and in the Voronoi one k_{Voronoi} further confirms this: the distribution in figure 9.10 is quite narrow, with an average value of 0.3 links and a standard deviation of 1.3 links (i.e. $k_{\text{Voronoi}} \approx k_{\text{Contact}} \pm 1.3$ for k the node degree). Thus, although the Voronoi distribution doesn't perfectly replicate the individual environment of each cell, it provides a good proxy for it.

This is particularly useful for images where the cell membranes aren't accessible and where we rely on the nuclei staining to get access to cell positions. In this situation we must rely on either the geometric or the Voronoi network construction methods. By showing that the individual links generated by the Voronoi network-construction method resemble those generated by finding the cell-cell contacts, we prove the validity of the Voronoi construction method.

A feature of the network construction tool in **Griottes** is the possibility to remove links that are greater than a given critical distance. The question then becomes choosing the distance and understanding how it affects the resulting network structure. In figure 9.11, we plot the evolution of the average degree of a given cell in the network as a function of the maximum cell-cell distance. The contact-based network remains of the same degree whatever the distance since neighborhoods are built from direct cell-cell contact. What we see is that for small maximum distances (<12 μ m) the Voronoi and geometric networks look very much alike. However, as the maximum distance increases, the average degree in the geometric network increases too, reflecting the fact that two cells can be linked even if a cell is wedged between them (Fig. 9.8). The average degree in the Voronoi network on the other hand stabilizes and reaches an asymptotic value similar to the contact-based network. This quick illustration highlights the need to carefully select a pertinent maximum cell-cell contact length in order to construct a realistic network.



FIGURE 9.11: Degree as a function of the maximum cell-cell distance and the network construction method.

9.4.4 Edge weighting by contact area

In many problems, the cell-cell interaction area can be pertinent to understand tissue behavior since both chemical and mechanical signals are transmitted through membrane contacts. As such, it can be pertinent to include the contact surfaces between neighboring cells in the graph to increase the relevance of the downstream analysis. This feature can be transposed in the graph object output by weighting the edges that link two nodes, replicating *in-silico* the differing interaction strengths between cells observed *in-vivo*.

Griottes includes an option to weight edges as a function of the contact surface between two cells if the membrane masks are available. From an input image where the segmentation masks are constructed from the membranes, it is possible to calculate the contact surfaces between any two cells to construct a weighted graph as illustrated in figure 9.12. The option is currently only available in 2D images but will soon be implemented in 3 images too.



FIGURE 9.12: Weighted connectivity graph constructed from an image of neural stem cells in the zebrafish telencephalon. The node colors represent different cell types. Image from Dr. Nicolas Dray and Prof. Laure Bally-Cuif.

Key take-aways

Context Emerging technologies and improved imaging techniques increase the quality and quantity of tissue images. New tools for analyzing these images after segmentation are needed to extract relevant information from the experiments.

Results I have written **Griottes**, a Python library that facilitates the construction of the connectivity network of the tissue. This network representation captures essential cell features, including their phenotype, neighborhood etc., and builds a "network twin" of the imaged tissue. This graph object can be analyzed using conventional network-analysis tools, as well as the libraries presented in the appendix (see chapter ??). Taken together, **Griottes** greatly facilitates the quantitative analysis of tissue images.

9.5 Conclusion

Griottes is a tool designed to simplify image analysis pipelines and facilitate the extraction of information from segmented images. The core intuition behind the program is that the graph object is a convenient method to store cell positions and monitor the interactions between cells. In order to analyze the "network twin" of biological images of interest, I have also developed several minor software libraries (their utility will be illustrated in chapter 10). This puts the generated graph at the center of the subsequent analysis steps (Fig. 9.13) and Griottes becomes the tool that opens the door to new kinds of studies.



FIGURE 9.13: Schematic of a typical analysis pipeline. It has been designed to be composed of independent libraries that each conduct a single, clearly defined, task. In red are the output and input objects for each program, and in green are the programs themselves with their task description.

Beyond facilitating the analysis of image data, graphs also provide a very parsimonious method for storing the information of interest. Images are notoriously memorygreedy, roughly storing the intensity level of every single pixel in the image. The graph framework only keeps track of the features that the end-user wishes to keep. In practice, this means that we could quantitatively analyze and store data from images in the 10s of GB in size in dict-like objects taking up memory in the 100s of kBs.

In the near future, the library will be improved so that it can be easily installed on third party computers and used on any tissue image. In the meantime, I have started a project where **Griottes** plays a key role: the structural analysis of fetal liver development. The next chapter elaborates on this project.

Chapter 10

Structural analysis of the fetal liver as a hematopoietic expansion niche during development

This project is a collaboration between the Lymphocytes and Immunity lab at the Pasteur Institute and the Stem Cells in Regenerative Biology and Repair lab in the I3S (Instituto de Investigação e Inovação em Saúde) INEB (Instituto de Engenharia Biomédica), Universidade do Porto). In particular, I acknowledge the support of FCT/MCTES in the framework of the project "POCI-01-0145-FEDER-032656" through grants to P. Pinto-do-Ó.

10.1 Introduction

Hematopoietic Stem Cells (HSCs) are able to maintain homeostasis of the blood system during an individual's lifetime, producing all mature blood lineages (a process also named hematopoiesis). These include, among others, platelets, erythrocytes (red blood cells), and immune cells such as myeloid cells, T cells and B cells [156] (figure 10.1a.). Therefore, HSC transplantation has been a very successful cellular therapy in treating a large range of hematological problems such as auto-immune or genetic diseases [157]. Indeed, the capacity of HSCs to generate any mature adult hematopoietic cell type is particularly precious for individuals who have had their hematopoietic system ablated, for example following irradiation to treat leukemia [156]. However, the use of HSCs as a treatment has been hampered by the availability of the cells themselves. Two main sources exist today, either from the bone marrow (BM) of the patient or a compatible donor, or from the umbilical cord. However, in both cases, low cell numbers constitutes a critical limitation for the success of therapy, and the process of *ex vivo* HSC population expansion has yet to be understood.

In the adult, HSCs reside in the BM, in specific niches that support their maintenance and their differentiation. However, the expansion of the HSC population *in vivo* occurs during development in the fetal liver (FL). As such, understanding the niche in which the HSCs proliferate in the FL is of critical importance to identify the determinants leading to successful population expansion. This is the goal of the current project: develop quantitative methods to analyze the FL structure with single-cell



FIGURE 10.1: **a.** HSCs can differentiate into any mature blood lineages. This versatility makes HSC transplantation a widely used and very successful therapy against a large range of hematological diseases [156]. **b.** HSCs are thought to proliferate in specific niches in the FL, these niches are characterized by specific soluble factors and cell types present in the environment of the cells [157].

precision in order to describe the FL and the specific niches supporting HSC expansion and/or differentiation towards specific lineages during development.

10.2 Ontogeny of the hematopoietic system

10.2.1 Known ontogeny of the hematopoietic system

It is well-known that in the adult, the HSCs reside mainly in the BM [158], where they self-renew and eventually differentiate in response to stress. However, the HSCs are known to proliferate in the FL without losing their stemness during embryonic development [159]. Indeed, the presence of the HSCs in the adult BM is the result of a journey across multiple anatomic locations.

Hematopoietic cells emerge in three distinct waves, with the two first waves in mice originating from the yolk sack at embryonic day (E) 7.5 and E8.5 producing multiple types of hematopoietic primitive cells [157]. HSCs emerge from the third wave of hematopoiesis occuring between E9.5 and E11 in the aorta-gonads-mesonephros region (AGM) [160]. They then mature as they migrate to the FL where they proliferate, multiplying their numbers more than 30-fold [159]. At around E16.5 HSCs start to migrate to the BM where they are maintained through adulthood [157]. Figure 10.2 presents a time-line of hematopoietic process in mice and humans.

10.2.2 HSC localization in the fetal liver

The hypothesis that stem cells are regulated by their environment has been proposed by Schofield [161] and states that stem cell properties are maintained by the surrounding cells designated "niche". Since we know that the HSCs proliferate in the FL, the question now becomes which cell types and particular environmental niches are associated with this proliferation.



FIGURE 10.2: Ontogeny of the hematopoietic system in mouse and human. Embryonic hematopoiesis is established in three distinct waves, with HSCs emerging in the AGM region and migrate to the FL. There the HSCs expand and differentiate into mature blood lineages before migrating to the BM around E17. Notations: YS, yolk sac; P-Sp para-aortic splanchnopleura; FL, fetal liver; AGM, aorta-gonads-mesonephros; BM, bone marrow. Extracted from [157].

An important non-hematopoietic cell type in the FL at E11 is hepatoblasts. These cells are bipotent immature FL cells that differentiate into hepatocytes or cholangiocytes. Hepatoblasts can be identified in situ using different markers, namely HNF4a and E-Cadherin, although specific identification requires the combination of several supplementary markers [157]. Hematopietic progenitors (identified as c-Kit+ cells) are found to be associated with hepatoblasts at E14.5 [162], though the c-Kit+ cells could also represent erythroid progenitors which are the most frequent population in the FL at this stage [163].

Endothelial cells are also highly represented in the FL, surrounding the main blood vessels in the tissue. The first vessels to appear are the hepatic sinusoids, around E10, followed by the portal vessels after E10.5-E12.5 in the mouse. The endothelial cells express a large variety of markers according to vascular location and developmental stage [157], and have been shown to be present in proximity of HSCs at E11.5 (Fig. 10.3a) [164]. The liver parenchyma is surrounded by a single epithelial layer (mesothelium) composed of mesothelial cells (MCs). MCs proliferate during liver development and are known to express several growth factors essential for hepatic development.

Another liver population is that of perivascular cells. Their precise characterization remains a bone of contention, though several cellular sub-populations are known to play a key role in the environmental niche of HSCs during FL development. For instance, a population characterized by the expression of Nestin and NG2 was identified as periportal pericytes (Fig. 10.3b) [165] and reported to be in close association with HSCs. Similarly, a key HSC niche factor - stem cell factor (SCF) - was proven to be expressed by endothelial and perisinusoidal hepatic stellate cells, thereby creating a



FIGURE 10.3: **a.** High magnification (60x) HSC (green) surrounded by endothelial cells (red) at 78 hours post-fertilization in a zebrafish. Image taken from [164]. **b.** Immunofluorescence analysis of liver cryosections at postnatal day 0 (P0) and postnatal day 8 (P8). * designate portal vessels (delimited by closed arrows). The bile ducts are designated with open arrows. Taken from [165].

vascular HSC niche in the developing FL [166].

The precise establishment of HSC location in the liver has been hindered by the multi-marker assessment required (i.e. Lin/c-Kit/Sca1/CD150/CD48) to identify the cells (see table 10.1 for a description of the different markers). The use of HSC-specific reporter lines can circumvent this experimental limitation, and different lines have been used in the BM [167]. However, to-date, no validation of these lines was performed in the embryo.

10.2.3 Spatial description of the fetal liver niche

After expansion, HSCs migrate from the FL and start colonizing the BM at E16.5. This coincides with the progressive vascularisation of the BM at E15.5-E16.5. Even though it isn't possible to directly image the distribution of HSCs yet due to the complex combination of antibodies necessary to characterize the cells with certainty, it is possible to image and analyze the structure of the fetal liver during key development phases. By selecting a reduced number of markers staining for cells of interest, it is possible to derive an initial description of the fetal liver (Fig. 10.4). These images make it possible to study the evolution of the fetal liver niche as it expands and supports the differentiation and clonal expansion of the hematopoietic progenitors.

The Cumano and Pinto-do-Ó labs have worked to develop marking and imaging protocols to stain and measure progenitor cell distributions in the FL and to characterize their niche (Fig. 10.4). What I have sought to do is to develop robust methods to quantify cellular distributions in the FL to analyze the images produced by their team. In the next section I will describe the different methods used to characterize the fetal liver and the main results I have obtained.

Marker	Marker properties
c-Kit	Hematopoietic progenitor marker.
Gr-1	Myeloid cell marker.
CD45	Pan-hematopoietic marker, stains all
	hematopoietic cells, except erythro-
	cytes.
NG2	Perivascular cells of the periportal ves-
	sel system and mesothelial cells.
Ter119	Erythroid-specific marker.
CD150	Stains different hematopoietic cells,
	namely megakaryocytes, and some
	progenitor cells; used to identify
	hematopoietic stem cells.
CD48	Stains different hematopoietic lineages,
	inclusively more committed progenitors;
	used to identify hematopoietic stem
	cells by negative exclusion.
Lin	mature hematopoietic lineages (i.e.
	lymphoid, myeloid, erythroid and
	megakaryocytic lineages).
Sca-1	Hematopoietic stem and progenitor
	cells; also stains a subset of endothelial
	cells.

TABLE 10.1: List of markers commonly used in the current study

10.3 Network-based structural analysis of the fetal liver niche

10.3.1 Imaging of the fetal liver

Image origins

The fetal liver images used for the current study were provided by Cumano and Pintodo-Ó's team at the Pasteur Institute and i3S (University of Porto).

Protocol

E12.5 and E18.5 fetal livers (FLs) were collected directly into freshly prepared 2% PFA fixative buffer (Fischer Scientific) and kept at 4°C overnight under rotation. After washing in PBS, tissues were shortly-stored at 4°C until sectioning. Following FLs embedding in 4% Agarose (Lonza), 100 μ m sections were performed using a Leica VT1200S vibratome. Sections were permeabilized and blocked with PBS containing 0.05% Tween-20 (AppliChem GmbH) and 10% donkey serum (Sigma-Aldrich). This solution was also used to dilute primary and secondary antibodies. Sections were incubated with primary antibodies for 2 hours at room temperature (RT) and overnight at 4°C, and secondary antibodies for 2h at RT, both in the dark and with rotation. Between incubations and after secondary antibody incubation, sections were washed



FIGURE 10.4: **a.** Immunostaining of E12.5 liver cryosections with the hematopoietic marker CD45 (red) and the pericyte marker NG2 (green). **b.** Immunostaining of E14.5 liver cryosections with the cell adhesion molecule CD146 (red), the pericyte marker NG2 (green) and endothelial marker CD31 (blue). **c.** Immunostaining of E16.5 liver cryosections with the mesenchymal marker PDGFR α (red), endothelial marker CD31 (green) and mesothelial marker Gp38 (blue). **d.** Immunostaining of E16.5 liver cryosections with the cell adhesion molecule ALCAM (red) and the pericyte marker NG2 (green). Nuclei are shown in white. Scale bar 30 µm. Fields are representative of the tissue section. The images are generously provided by Cumano and Pinto-do-Ó's team at the Pasteur Institute and i3S (University of Porto).

repeatedly with PBS containing 0.05% Tween-20 and 3% NaCl (minimum $4\times$, 15 min). Sections were counterstained with DAPI (D9542, Sigma-Aldrich) for 2 hours at RT and kept overnight in RapiClear 1.52 (SunJin Lab) for optical clearing. Sections were mounted in RapiClear and imaged on a Leica TCS SP8 using a PL APO 63x /1.40 oil immersion lens (zoom factor 0.75). 10-20 µm tile scans were acquired at 400Hz, bidirectional mode, with z spacing of 0.3 µm, at 12-bits and 1024×1024 resolution.

10.3.2 Project goals

To characterize the FL Cumano and Pinto-do-Ó's team performed immunostainings using markers for different stromal populations (hepatoblasts, endothelial, perivascular, mesothelial, etc) and also for hematopoietic cells (megakaryocytes CD41+, B cells B220+, myeloid Gr-1+ or CD11b+, progenitor cells Kit+ and HSCs) to identify and characterize FL hematopoietic niches that drive HSCs expansion and differentiation.

From these data, it has been noticed that hematopoietic cells start to cluster in specific locations starting at E16.5 - blood vessels and in the mesothelium (*unpublished data*). The goal of this project is to use the quantitative framework described in chapter 9 to characterize the tissue during fetal liver maturation. In this thesis, we will

focus on the distribution of myeloid cells. This allows us to illustrate that the analysis pipeline and the network-based analysis framework can be used to characterize complex tissues. In the future, the analysis of lymphoid and progenitor cells will complete the picture.

10.3.3 Description of the fetal liver tissue characteristics

Description of E12.5 FL tissues

In figure 10.5 we present an example image from an E12.5 tissue slice. At this stage the liver is divided into four separate lobes, two symmetrical dorsal and ventral lobes and the liver is sufficiently small for a slice of it to be included in a single image as is clear in figure 10.5.

We highlight several regions a-c that can be viewed in figure 10.6. Focusing on the blood vessel in the center of the FL slice, we see that the contour of the vessel is already recognizable with a NG2+-rich border. Zooming-in we can distinguish individual cells (see Fig. 10.6a) and study the structure of the tissue. We also observe the presence of NG2+ cells at the mesothelium border (Fig. 10.6b) and their absence in the parenchyma (Fig. 10.6c).

Description of E18.5 FL tissues

Similar E18.5 tissue images were obtained and we recognize different tissue composition. Indeed, the concentration of Gr-1+ cells seems qualitatively higher than at E12.5 and we recognize different tissue structures around the vessels irrigating the FL. Indeed, we clearly recognize the portal vessels by the NG2+ cells lining the endothelium (Fig. 10.8c) [165], and the central veins by their absence (see Fig. 10.8c).

From these qualitative observations we can build a schematic representation of the E18.5 fetal liver (fig. 10.23). These landmarks of the fetal liver topology will be studied in depth in the following sections of the thesis.

The number of cells in the fetal liver approximately doubles every 24h between E12 and E18, both from cell division within the organ an from the influx of cells from the bloodstream. The tissue is now much to big to be contained in a single image with single-cell precision, and we are limited to the analysis of fetal liver regions.

The staining is homogeneous within the tissue

A common issue in tissue staining protocols is the uneven diffusion of antibodies within the sample than can bias the imaging experiments. Depending on the sample preparation protocol, this can lead to stronger signal on the edges of the tissue, or, in contrary, in the middle of the sample. It is therefore necessary to be prudent when interpreting fluorescence levels within the sample.

In the current data from E12.5 and E18.5 fetal liver slices, we notice that the distribution of CD45+-cells is homogeneous throughout the sample (Fig. 10.24 and Fig. 10.25). This is a valuable internal protocol, suggesting that the staining is homogeneous throughout the tissue and that variations in other cell type distributions reflect biological signals within the experiment.

CHAPTER 10. STRUCTURAL ANALYSIS OF THE FETAL LIVER AS A HEMATOPOIETIC EXPANSION NICHE DURING DEVELOPMENT



FIGURE 10.5: Section of a whole E12.5 fetal liver. The highlighted areas are examined in figure 10.6. Scalebar is $200\mu m$.

а



FIGURE 10.6: **a** Portal vein from an E12.5 fetal liver. The scalebar is $40\mu m$ in both images. **b** Mesothelium highlight from an E12.5 fetal liver. Notice the NG2 monolayer at the mesothelium border and the clusters of CD45+ cells in the tissue. The scalebar is $60\mu m$. **c** Parenchyma highlight from an E12.5 fetal liver. The scalebar is $60\mu m$.



FIGURE 10.7: Schematic description of the main features of the E18.5 fetal liver.



FIGURE 10.8: **a** E18.5 fetal liver image stained with DAPI (gray), CD45 (red), NG2 (green) and Gr-1 (cyan). Scalebar is $250\mu m$. **b** Tissue slice in the z direction, the clearing and staining mark cells efficiently in all three spatial dimension. **c** Highlight of the border of a central vessel irrigating the fetal liver. Scalebar is $60\mu m$. **d** and **e** Highlight of the border of a portal vein and of the tissue parenchyma. Scalebar is $60\mu m$.

The images are invariant along the z-axis

The images obtained are 3D confocal scans of the FL tissue. To analyze them in 3D we first need to ensure ourselves that the image properties are independent of the imaging depth. Indeed, due to optical effects, the image properties could differ depending on the section depth, impacting the validity of our analysis.



FIGURE 10.9: **a** FL image of the mesothelium at E18.5. **b-d** Tissue images in the (x0z) plane corresponding to the positions in panel (a). The scalebar is $15\mu m$. **e** Average intensity in the different imaging channels as a function of imaging depth.

We measure the average pixel intensity in each channel as a function of the imaging depth in figure 10.9e. We see that the relative variations in image intensity are small and depend on the channel, suggesting that the changes reflect changes in cell composition rather than physical limitations. This is confirmed by inspecting z-slices of the image in three different positions (see Fig. 10.9b-d). We can therefore analyze the images as volumes and investigate the tissue structure in 3D.

10.3.4 Processing pipeline of the fetal liver images

Quantitatively analyzing the fetal liver images is a complex task, but that can be easily separated into different phases. I developed am image analysis pipeline to construct the "network twin" of each imaged tissue (Fig. 10.10). This is a graph of the tissue, where each of the nodes represents a single cell whose cell-type is determined by the available markers to characterize it.

I've described how cell properties can be extracted from a segmented image and how to build a network using **Griottes** in chapter 9. I will therefore spend the following paragraphs on the image segmentation procedure and how the cells can be classified based on their measured fluorescence.

Single-cell segmentation in the tissue

In the past couple of years, several powerful cell segmentation libraries have emerged. For practical reasons CellPose [14] stood out since it was written in Python, callable

CHAPTER 10. STRUCTURAL ANALYSIS OF THE FETAL LIVER AS A HEMATOPOIETIC EXPANSION NICHE DURING DEVELOPMENT



FIGURE 10.10: Image analysis pipeline to construct the network image of fetal liver slices.

from the command line and that it was possible to use it on the high computing cluster (HPC) hosted on the Inception GPULab at the Pasteur Institute. I detail the image analysis pipeline to segment the images in the appendix (chapter ??). The segmentation accuracy is above 95% in the tissue, as can be verified in figure 10.11.

Segmenting large tissue images was particularly challenging since segmenting images using CellPose is particularly memory-consuming. The computations need to run on GPUs to be feasible in a reasonable amount of time (ca. 30 minutes for large images) and GPUs only have approximately 4GB of memory on the Pasteur HPC. In order to analyze large tissues (Fig. 10.5) I developed a small library called Saucisson to cut the images into smaller pieces easily. This allowed me to rapidly section a fetal liver image into segmentable constituents, send them to the HPC, segment them separately, import them back again and reconstruct the original image and the mask of the segmented cells. This procedure allowed me to successfully segment images that were up to 45GB big.

The main concerns during the segmentation are the correct attribution of labels to dividing cells and multi-lobated megakaryocytes (often identified as two different cells). Once the nuclei have been segmented, we can build the network representation of the tissue using **Griottes**.

As mentioned above, the image sizes were particularly significant, which led to a large number of cells being detected after the segmentation protocol (typically >33k cells in each image). However, before analyzing this data, the detected cells need to be categorized into phenotypically relevant cell types.

Classifying the cells

The original images have multiple fluorescence channels, and so after treatment we now have the fluorescent intensity information for each cell in the labeled image in up to 4 channels. If we're interested of looking at the phenotypic properties of the cells, we can classify each individual cell in the network as positive or negative with regard to each marker from the intensity measurements (Fig. 10.12a,b). This is very similar to gating populations when analyzing FACS results, except we conserve the spatial context of each individual cell. This leads to the identification of different populations within the tissue sample as can be seen in figure 10.12c.

We confirm the thresholds for classifying cells as positive and negative visually on representative samples (Fig. 10.12d,e). The detected cell types fit well to the visual classification, although it is difficult to assign a cell type to some (rare) edge cases within the image (see Fig. 10.12e). Combining the information from the different fluorescence channels makes it possible to describe the phenotype of each cell. The final



FIGURE 10.11: **a** The individual cells can be detected and segmented from the nuclei staining (DAPI, gray) using CellPose [14]. Each individual cell is highlighted in a different color for lisibility. Scale bar is $60\mu m$. **b** Highlight from panel (a). The segmentation precision is satisfactory with very few errors. Scale bar is $25\mu m$

cell classification is also inspected visually and the parameters are adjusted until the visual classification agrees with the detected one (Fig. 10.13).

At this stage it is worth mentioning the main caveat of this analysis method. That is with regard of the nature of the measurement itself. Cellular expression of the markers that are recognized by the staining antibodies may vary as a function of the cellular environment. This means that expression levels measured on dissociated cells in FACS can differ from those recorded using microscopy data. Furthermore, the sample preparation protocol before imaging may itself impact the ability of antibodies to bind to the samples, by for example changing the shape of membrane proteins that were targeted (this was the case for a Kit-binding antibody that we used early on in our experiments). We have therefore multiplied different measurement methods (FACS, imaging, PCR etc.) to add certainty to the biological message that is drawn from the data.

The end result of the procedure is the generation of a spatially accurate network representation of the FL. This then allows the investigators to analyze the tissue using network analysis tools.


FIGURE 10.12: **a** Intensity distribution of Kit signal among the cells $(N \approx 70 \cdot 10^3)$ of an E18.5 fetal liver. Cells above an intensity threshold of 4.9 (dashed red line) are considered Kit+. The thresholds selection method is described in the main text. **b** Intensity distribution of CD45 signal among the same cells. Any cell with an intensity above 4.5 is considered CD45+. **c** Intensity distribution of Kit and CD45 intensities in the image. **d**, **e** The intensity thresholds are chosen after careful visual in spection. Cells detected as Kit+ or CD45+ (red dots) and as Kit- or CD45- (blue dots) fit well with visual classification, although edge cases can be difficult to validate (white arrow).



FIGURE 10.13: Multi-channel images of a fetal liver slice. The scattered dots show the positions of detected cells, their color represents the cell classification. A cell is detected as being positive (+) if the average intensity of the pixels within the mask are above a certain threshold.

Key take-aways

Context The fetal liver is a crucial site for hematopoietic cell expansion and differentiation. Understanding the processes that occur in the organ during development is key to understanding the maturation of the hematpoietic system. Despite this, very little is known about fetal liver organization since it's a difficult tissue to study and to image.

Results I have shown that we can segment fetal liver images and build the network recapitulating the tissue structure. The cells can be accurately classified into different cell types as a function of their measured fluorescence.

Future steps I will now analyze the network to uncover the organization of the fetal liver and the distribution of myeloid cells.

10.4 Using Hidden Markov Random Fields to characterize tissues

This section is best read while listening to *The Latent Space* by Mouse on Mars.

10.4.1 Why use Hidden Markov Random Fields?

The goal is to be able to reproducibly define biologically relevant regions in tissues in an unsupervised manner and starting from the network produced by the Salade package (for more information see Appendix section ??). We define each region as defined by a specific mix of cell types. Importantly, this means that there can be more regions than there are cell types in the tissue. At the end of the segmentation procedure, we expect to retrieve a latent network with the same structure as the observed network, with smooth, homogeneous regions defined as a function of the composition of the observed network. Hidden Markov Random Fields (HMRF) have demonstrated their efficacy for image segmentation [168], and have more recently been used to characterize regions in 2D tissues [144]. Here, we propose a method to exploit the HMRF framework in the context of network-representation of tissues.

10.4.2 From Markov chains to Hidden Markov Random Fields

Wikipedia tells us that "a Markov chain is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event" ¹. Thus, the Markov chain is useful when we need to compute a probability for a sequence of observable events (Fig. 10.14a). This simple model has proven extremely fruitful in many fields, ranging from applied mathematics to physics and - of course - biology.

As these models have gained in complexity, the kind of applications Markov chainbased models can answer have progressively expanded. One such instance is when the event we are interested in is hidden. A hidden Markov model (HMM) allows us to look at both observed and hidden events that we think of as causal factors in our probabilistic model (Fig. 10.14b). The canonical example of HMMs is to decipher speech, but another interesting example is to predict genetic sequences [169].

The issue with many HMMs is that they do not take into account spatial information that can heavily influence the dynamics of the system. To solve this, the Hidden Markov Random Field has been introduced [168]. Here the spatial information is encoded through the mutual influences of neighboring sites that can be inferred through analysis of the observed network (Fig. 10.14c).

10.4.3 Presenting the HMRF algorithm

Networks derived from tissue images can contain several thousands of cells. Finding characteristic regions within the tissue can be interesting to describe its architecture.

¹Markov Chain, looked up on October 12, 2021.



FIGURE 10.14: **a** In a Markov chain, the probability distribution of next step c_i only depends on the state of the previous step c_{i-1} . **b** The Markov chain can itself be hidden, with an observable state at time *i* only depending on the hidden (or *latent*) state at the same timeinstance. **c** Instead of being directed in time, the latent state can be an arbitrary network in which the Markov property dictates that the latent state *i* only depends on the node's immediate neighbors.

This is a tough problem as these regions are seldom composed of a unique cell type (Fig. 10.16a). In order to find them we have developed an unsupervised regionalization algorithm that derives regions of homogeneous cell composition from a graph with nodes populated with categorical attributes. It is based on hidden Markov random fields and allows us to do this by defining each region in the latent space as having a unique cell composition.

Hidden Markov Random Fields work by generating regions in the latent space. Cells belonging to the same region tend to be clustered together since there is a cost for each link between two nodes belonging to different regions. There is also a cost if the cell type in the observed network is absent (or present in low concentrations) from the region it assigned to. This tends to generate regions of different cell compositions that can then be used to describe the tissue structure. For example, in an example tissue composed of four different cell types (Fig. 10.15), each region found by the (fictive) HMRF algorithm has a distinct cell composition. For further information, we detail the HMRF algorithm in section ?? of the appendix.

The algorithm output is a graph of the same structure as the input but with labeled regions with different cell compositions. Tissue areas far apart can be detected a having



FIGURE 10.15: A fictional result of the HMRF algorithm. The tissue is composed of two regions of differing composition. Cell types belonging to the same region are clustered together.

the same cell mix (Fig. 10.16b), enabling investigators to recognise physiologically similar regions. The composition of each region is immediately accessible (Fig. 10.16c). In figure 10.16d-g we show the convergence of μ for each cell type for different regions. We see that the parameters converge in less than a dozen epochs to a stable value with little change thereafter.

10.4.4 The fetal liver is irrigated by different vessel types

To illustrate the power of the HMRF algorithm, we analyze fetal liver slices to see if we can automatically distinguish physiologically relevant regions (fig. 10.16b). For the purpose of this analysis, we have asked the algorithm to detect four different regions in the tissue (this number is arbitrary but can be easily adjusted by the user). The composition of the different regions is shown in figure 10.16c highlights the characteristics of each one of these. The region 0 (named "parenchyma") stands out by the near-absence of any NG2+ and CD45-/Gr-1+ cells. On the other hand, three out of the five visible blood vessels irrigating the fetal liver are singled out in pink (named "NG2+ veins"). These regions (number 3) are characterized by high levels of NG2+ and CD45+/Gr-1+ cells. Interestingly, the unsupervised analysis differentiates between these morphologically similar regions in the tissue.

Complementary analysis where we manually select regions of interest shows that we can classify the vessels into two different types; the NG2+ portal veins (in blue) and the NG2- central vessels (in yellow) (Fig. 10.17b). These features are of particular interest as they have been associated with HSC presence [165]. A closer look at these highlighted regions show distinct differences between the tissue regions. The portal veins are rich in NG2+ cells, in particular in the immediate vicinity of the vessel border (Fig. 10.17c). In stark contrast, the cells constituting the central vessels express little or no NG2, but the region is rich in cells expressing CD45 and Gr1. Finally, looking at figure 10.17e, we see that the mesothelial epithelium seems rich in CD45+/Gr-1+ cells.

A quantitative analysis of the composition of these regions reveals stark differences between the different highlighted regions. In particular, we can see the numerous NG2+



FIGURE 10.16: **a.** Segmented tissue. Each dot represents a different cell with the color marking the specific cell type. Notice the in-homogeneous distribution of certain cell types in the tissue. **b.** Regions detected in the tissue by the Salad algorithm. Notice the different compositions of the blood vessel regions. **c.** Absolute cell compositions of the regions detected in (b). Each region corresponds to a specific composition. **d-g.** Evolution of the region composition mean μ in each cell type as a function of the iteration number. The parameters in the HMRF algorithm converge rapidly to their stationary values. (d) Gr-1-/CD45-/NG2-(e) Gr-1+/CD45+/NG2- (f) Gr-1-/CD45-/NG2+ (g) Gr-1-/CD45+/NG2-

cells composing the portal vein region (figure 10.17f.). Focusing only on the CD45 and Gr-1 signals among CD45+ cells, we see that the mesothelium, the portal veins and the central vessels are particularly enriched in CD45+/Gr-1+ myeloid cells compared to the rest of the tissue. The enriched presence of these cells around the different vessel types can be indicative of a niche promoting their proliferation or differentiation. Furthermore, the similar composition in double-positive cells between the portal veins and the central vessels supports the idea that the observed differentiation is independent on the vessel type and and the accumulation of these cells is driven by CXCL12 pro-



FIGURE 10.17: **a.** Fetal liver slice at E18.5. **b.** >30k cells are detected and classified in one of 5 cell types. Analysis of the tissue reveals three specific structures in the tissue: the mesothelium, NG2+ vessels (also called portal veins) and NG2- vessels (central vessels). **c-e.** Highlights of the three regions detected in the fetal liver slice. **f.** For each region we can precisely calculate the tissue composition. **g.** Focusing on the distribution of CD45+/Gr-1+ double positive cells among the CD45+ population, we see that the mesothelium is particularly rich in double-positives. **h.** The Gr-1 marker stains granulocytes that are one possible differentiation pathway for HSCs (panel from [170]).

duction by perivascular cells covering the blood vessels in the FL, regardless of their phenotype (*unpublished data* from the Cumano and Pinto-do-Ó labs).

10.5 Network-based description of cell niches

Understanding the niche that cells populate is crucial to understand the factors guiding cell differentiation, function and phenotype [157]. For instance, HSCs are thought to populate specific locations in the fetal liver such as close to the periportal pericytes [165]. In order to characterize the environment of individual cell types in the tissue, we can harness the graph-based representation of the tissue provided by **Griottes** to look at the local and regional tissue organization.

10.5.1 Investigating the immediate cell neighborhood

Motivation

To quantitatively describe the organization of cells in a tissue it is necessary to understand precisely how the cells are organized with regard to each other. By representing the tissue as a network of cells, we can go a step further and determine whether cells cluster together or avoid each other as a function of their phenotype. To do so, we focus on the nearest neighbors of each cell and compare the average local composition to the expected composition in the case of randomly distributed cell types via the permutation test developed by Schapiro et al. [171]. This method is particularly robust and enables us to quickly compare the relationships between an arbitrary number of cell types in the tissue.

Method

We start by calculating the distribution of neighbors of each cell. We denote i the index of the cell, $i \in [1, N]$, N being the number of cells. The distribution of neighbors is encoded in a vector of size M (M is the number of cell types):

$$\mathbf{n}_{\mathbf{i}}(m) = \frac{1}{N_i} \sum_{j \sim i} \delta(x_j, m), \qquad (10.1)$$

where N_i corresponds to the number of first neighbors of cell $i, j \sim i$ are the neighbors j of cell i, m is the cell type $(m \in [1, M])$ and x_j is the type of cell number j. Thus, for each cell type m, we get a vector, of size M, in which component m' gives the average proportion of neighbors of cell of type m' for a cell of type $m (\langle \mathbf{n}_i(m') \rangle_{i|x_i=m}, \text{ noted } \langle \mathbf{n}^m(m') \rangle).$

Next step is to compare this statistic obtained from the experimentally observed network to the same feature calculated on a random graph. To do so, we reshuffle randomly cell types within the network to preserve the structure of the tissue. We then compute the average distribution of neighbors and we do this multiple times (typically 10^3) so we can compute a probability that the experimental cell type distribution differs from a random one:

$$P(m, m') = \frac{\sum_{\text{permutations}} 1\left(\langle \mathbf{n}^m(m') \rangle_{\text{random}} \le \langle \mathbf{n}^m(m') \rangle_{\text{experiment}}\right)}{N_{\text{permutations}} + 1}.$$
 (10.2)

If your head hurts at this point, it's perfectly normal. The expression is complex so it is worth unpacking what is going on. If in the experimental network, the cells of type *m* are often neighbors with cells of type m', then the relative value of $\langle \mathbf{n}^m(m') \rangle_{\text{experiment}}$ will be high. This means that there is a high chance that for a given random network configuration, $\langle \mathbf{n}^m(m') \rangle_{\text{random}} \leq \langle \mathbf{n}^m(m') \rangle_{\text{experiment}}$ and therefore that for that particular configuration the indicator function will be worth 1. Repeating this procedure multiple times and extracting the average value of the indicator function enables us to assert whether or not the cell types attract or repulse each other significantly

For the purpose of this study, we consider that if $0.95 \leq P(m, m')$, then the cell types significantly attract each other. Conversely, if $P(m, m') \leq 0.05$, then the cell types repulse each other. This method makes it possible to summarize the relative distribution of cells in a single figure (Fig. 10.18c.).

Single cell neighborhood characterization in the fetal liver tissues



Neighbour cell phenotype

FIGURE 10.18: **a.** Segmented cell distribution from an E18.5 fetal liver slice. N = 6445 cells. **b.** Shuffled cell distribution on the network formed by the cells in the tissue shown in (a.). **c.** Interaction matrix between the different cell types of (a.), cells preferentially interact with themselves.

Deviations of the cellular environment of individual cells in tissues from the statistically expected environment can guide us towards specific interactions of interest [16, 171]. To identify specific niches we conducted permutation tests on data extracted from E18.5 fetal liver images. An example of the network before permutation is found in figure 10.18a. In contrast to examples of the permutated graph (Fig. 10.18b), we see no concentration of NG2+ cells around the portal vessels. These differences are reflected in the interaction matrix, where we see favorable interactions between NG2+ and CD45+ cells. This probably reflects the concentration of these different cell types around the vessels in the tissue. Conversely, we observe no preferential interaction between CD45-/GR-1-/NG2- cells, probably reflecting the relatively homogeneous composition of the parenchyma in E18.5 tissues.

Currently, we only have a limited number of cell types present in the tissue, limiting the explicative power of tools such as the permutation test explained immediately above. As we improve the input images, these tools will increase in relevance to interpret experimental data.

Improving the local tissue characterization

Currently network-based methodologies exploit the local cell environment composition to describe the cellular niches each cell is involved in. For instance, the i-niche method developed by Goltsev et al. [16] accounts for the average immediate neighbors of any given cell type. This means that the local tissue structure, (i.e. *where* any given neighbor will distributed in the neighbor network) and the cellular composition a couple of steps away in the network are ignored. An interesting improvement would be to use graph-based tools to maximize the number of environmental properties taken into account when characterizing the niche of any specific cell type.

10.5.2 Long-range spatial distribution properties

As mentioned in the previous paragraph, describing the immediate neighbors ignores any long-range clues in the tissue structure. However, understanding these interactions between cell types and the structure of the tissue is crucial to investigate cell-cell interactions in depth. This can be particularly significant when the biological phenomenon depends on the cooperation of multiple cell-types and the combination of mechanical and chemical cues.

Understanding the longer-range interactions between cell types embedded in a network is a familiar problem in network analysis. One answer to the problem is to use spatial statistics and adapt them to the network representation. For example, we can introduce the probability that a cell at a distance d from a cell of cell type m will be of the particular cell-type m'. We call this probability C(X, Y) the proximity parameter. We define L(i, d) the set of nodes at a distance of d from node i, I(X) the set of nodes of type X:

$$C(X,Y)(d) = \left\langle \frac{\sum_{j \in L(i,d)} \delta(X, \operatorname{Class}(j))}{\operatorname{Card}(L(i,d))} \right\rangle_{i \in I(X)}.$$
(10.3)

Using the proximity parameter makes it possible to describe the relationship between individual cell types, not only in the immediate neighborhood but also over longer distances. From the evolution of the proximity parameter as a function of d we infer the range of the interactions. In order to compare the experimentally obtained



FIGURE 10.19: **a.-d** Proximity parameter analysis of different cell types makes it possible to characterize cell distribution properties in the tissue or in specific tissue regions.

values, we calculate the proximity parameter for the same cell-type couple on a network with the same structural properties but where the cell types have been shuffled randomly. This provides a baseline comparison in order to interpret the data.

Looking at experimental results from the E18.5 fetal liver images previously described, we see that the proximity parameter for NG2+ cells with themselves reaches the value corresponding to a random distribution within a distance of d = 17. This is slightly larger than what is the case for CD45+/Gr-1+ cells where the critical distance is approximately d = 12. One way to interpret the data is remark that NG2+ cells tend to be clustered together in homogeneous but elongated structures (Fig. 10.17c).

More tellingly, we see that the normalized proximity parameter for NG2+ cells with regard to CD45+/Gr-1+ cells rapidly increases from 0 to ≈ 1.7 at d = 3 and that it then decreases to reach the random value at around d = 10 (Fig. 10.19). This reflects the proximity between CD45+/Gr-1+ cell clusters, often located around the vessels as we described in section 10.5.3.

Combining single cell level data in a network representation allows us to measure aggregate levels of interaction between cell types and to characterize more precisely the structure of the tissue. In the past paragraphs we have described some tools to analyze graphs generated by experimental image data, however, new methods can be readily developed depending on the question being investigated.

10.5.3 Border analysis reveals specific spatial organization around fetal liver features

It is well-known that HSCs, as well as myeloid and lymphoid progenitors, migrate from the fetal liver starting E16.5 to the bone marrow [157]. Although the specific details of this migration are not yet well-known, it is hypothesized that it is done through the liver vasculature. Knowing the precise tissue structure around the blood vessels and the mesothelium is therefore of great interest to understand the details of this cell migration.

To do so first requires determining a reference point from which calculate the structure of the tissue. We choose to use its borders - at the blood vessel or at the mesothelium border - as this reference. This first requires to detect the position of the border. We do this by calculating the distance between the blood vessel and each cell in the region. If the distance is below a given threshold we classify the cell as belonging to the tissue border. From this procedure we can label the cells detected as being on the tissue border, even for complex blood vessel geometries, as is shown in figure 10.20b. For each cell in the tissue region, we then consider as distance to the border d the shortest path from the cell to a cell labeled as being on the border (Fig. 10.20c). A cell on the border will have a distance to the border of 0, for the cell just behind it d = 1etc.

With this method we can now analyze the tissue structure close to the border. Pooling together the portal veins (N = 4) and the central vessels (N=3) and the mesothelium (N=4) from the two images of the E18.5 FL, we can analyze the distribution of certain specific cell types as a function of the distance to the tissue border. Due to the slight differences in tissue composition, we focus on the relative composition of each layer in a specific cell type. The relative composition is the composition of the layer in a given cell type (for example 22% CD45-/Gr-1-/NG2+) divided by the composition of the whole tissue (in this case 2% CD45-/Gr-1-/NG2+), this gives a value higher than 1 (it is worth 11 for the described CD45-/Gr-1-/NG2+ population) if this specific cell type is more represented in the layer than in the tissue as a whole. Conversely, if the relative composition is smaller than 1 it means that the cell type in question is rarer in this layer than in the tissue in general.

We observe that the distribution of CD45-/Gr-1-/NG2+ is maximal at the border with similar distributions between the mesothelium and the NG2- central vessels (Fig. 10.20d). On the other hand, NG2+ portal vessels significantly have significantly higher concentrations of NG2+ cells close to the border. At a distance of two edges from the border, the expression of CD45-/Gr-1-/NG2+ cells is lightly below 1 and has reached the concentration in the parenchyma.

Focusing now on CD45+/Gr-1+/NG2- myeloid cells, we see in figure 10.20e that the concentration of double positive cells in the mesothelium quickly falls to the parenchyma concentration after just 1 to 2 edges from the border. On the other hand, in both portal veins and central vessels, the relative expression of CD45+/Gr-1+/NG2-cells peaks at 9 times the tissue concentration and there are approximately 4 layers of cells between the observed peak and the parenchyma value. This suggests that the distribution in CD45+/Gr-1+/NG2- cells is very similar, both in amplitude and in



FIGURE 10.20: **a.** Portal vein from a fetal liver slice taken at E18.5. **b.** Using the borderdetection procedure, the cells on the vein edge are detected (red) and can be distinguished from the other cells from the tissue (blue). **c.** Schematic representation of the region next to the tissue edge. The cells are labeled as a function of the shortest distance to the border. Colors represent different cell types. **d-f.** Relative concentration of CD45-/Gr-1-/NG2+ cells, CD45+/Gr-1+/NG2- cells and CD45+/Gr-1+/NG2- cells as a function of the distance to the border for cells belonging to NG2+ veins (portal veins), NG2- veins (central vessels) and the mesothelium. **g.** Layer-based schematic representation of the tissue as a function of the distance from the border. Depending on the tissue type the cells structure theselves in a distinct manner.

structure, between the two types of vessels irrigating the fetal liver.

Finally, looking at the CD45+/Gr-1-/NG2- cells, we see that there is little over-

expression compared to the parenchyma expression levels and so we can conclude that there is no specific expression pattern of CD45+/Gr-1-/NG2- close to the border of the tissues (Fig. 10.20f). The concentration of CD45+ cells is constant throughout the tissue.

The analysis of the tissue over-expressions allows us to make a schematic representation of the tissue structure for the different kind of features we have singled out in the E18.5 fetal liver. The portal veins are schematically composed of a single layer rich in NG2+ cells, then of about 4 layers rich in CD45+/Gr-1+ cells before reaching the composition of the parenchyma. The NG2- vessels (or central vessels) are structured similarly, except that they do not present a monolayer rich in NG2+ cells and immediately show 4 layers rich in CD45+/Gr-1+ double positive cells before reaching the composition of the parenchyma. Finally, the mesothelium can be described as being composed of a monolayer rich in NG2+ and CD45+/Gr-1+ cells before reaching the composition of the parenchyma.

This reinforces the conclusion reached in section 10.4.4 that the composition in CD45+/Gr-1+ cells is very similar between portal veins and central vessels, except that leveraging the network description of the tissue and the structural analysis of the fetal liver enabled us to characterize the similarity with a new level of precision. Indeed, it showed that the structure and the amplitude of the cell compositions are similar between the two vessel types. In the case of migrating cells leaving the fetal liver, this suggests that that chemotactic signal attracting CD45+/Gr-1+ cells is similar in both vessel types.

The fetal liver undergoes drastic structural change during the development of the fetus with the progressive emergence of complex structures and the cell differentiation and maturation of its populations (see figure 10.21). In the next chapter, we will characterize and quantify the structural evolution of the FL and of its sub-components from FL images taken at E12.5 and E18.5.

Key take-aways

Network analysis protocols In this chapter, we have shown that the network-based representation allows us to implement specific tissue analysis methods: the HMRF regionalization algorithm, local neighborhood analysis and border analysis. These methods make it possible to characterize the tissue organization and highlight individual cell-cell interactions.

Results I have shown that we can segment fetal liver images and build the network recapitulating the tissue structure. The cells can be accurately classified into different cell types as a function of their measured fluorescence.

Future steps I will now analyze the network to uncover the organization of the fetal liver and the distribution of myeloid cells.

10.6 Quantifying the fetal liver maturation process

10.6.1 Characterizing the evolution of tissue composition



FIGURE 10.21: **a** From E10.5 until E15.5, hematopoietic cells colonize the FL that rapidly becomes the major fetal hematopoietic organ up . **b** The structuration of the FL occurs after E15.5 with several distinct processes such as the hepatoblast-to-cholangiocyte specification occuring around the portal vessels and the emergence of bile duct structures. **c** Mature liver architecture recognizable after birth. All panels are from [157].

During maturation, the fetal liver undergoes important structural change between E12.5 and E18.5 (Fig. 10.21). The organ roughly doubles in size every 24h during this period, a result of cell migration to the fetal liver and of frequent cell division. The tissue growth is accompanied by liver restructuring and cell differentiation [157]. We can distinguish three morphologically distinctive regions in the E12.5 and E18.5 fetal liver slices: the mesothelium bordering the tissues, the parenchyma and the different vessels irrigating the fetal liver. Using the Coloriage module (see section ?? for additional information), we select the cells from these regions to investigate the cell type distribution within the tissue. The E12.5 fetal liver tissue on a whole is composed of 86% NG2-/CD45-/Gr-1- cells, 7% NG2-/CD45+/Gr-1- cells, 6% NG2+/CD45-/Gr-1- and 1% NG2-/CD45+/Gr-1+ double positive cells. But the distribution of cells within the tissue is heterogeneous. For instance, the mesothelium is composed of close to 15% NG2+ cells but with less than 1% of CD45+/Gr-1+ cells. Conversely, the cells surrounding the vessels irrigating the FL have a higher chance of being CD45+ (Fig. 10.22a).

Comparing the evolution of cell types between E12.5 and E18.5 we see that on average there is a sharp 57% decrease in the fraction of NG2+ cells in the tissue (Fig. 10.22b), whereas the fraction of CD45+ and CD45+/Gr-1+ cells increases by 133% and 657% respectively (Fig. 10.22b). The change in cell composition reflects the progressive differentiation of cell sub-populations and structuring of the FL. In the coming sections we will describe the evolution of the FL mesothelium and of the vessel regions in greater detail.

The FL composition at E12.5 and E18.5 obtained by imaging and flow cytometry yield very similar results. For example, the fraction of CD45+/Gr-1+ cells is measured as being 7.3% in the FACS data and approximately 8% in the image-based analysis. This confirms the accuracy of the image-based approach to extract tissue composi-



FIGURE 10.22: **a** FL tissue composition at E12.5. **b** Evolution of the FL composition between E12.5 and E18.5. **c** FL composition obtained from flow cytometry measurements at E12.5 and E18.5.

tion data (Fig. 10.22c), but also assures us that few cells are lost during the FACS preparation procedure.

10.6.2 CD45+/Gr-1+ cells accumulate around fetal liver vessels during maturation

Between E12.5 and E18.5, the vessels irrigating the FL differentiate into portal vessels characterized by an NG2+ perivascular layer [165] and central veins surrounded by NG2- cells (Fig. 10.23) [157]. The differentiation of the cell types surrounding the vessels is reflected in the overall composition change in the FL, with the NG2-/CD45-/Gr-1- population decreasing and with significant increases in the CD45+/Gr-1+, CD45+ and NG2+ populations (Fig. 10.24a).



FIGURE 10.23: Schematic description of the main features of the E18.5 fetal liver.

After identifying cells on the vessel border following the methodology developed in section 10.5.3 and looking at the distribution of differentiated cells as a function of the distance from the border cells, we see in figure 10.24b that there is both at E12.5 and E18.5 an increase in the fraction of NG2+ cells close to the border. However, the overall population at E12.5 is richer in NG2+ cells, the over-expression in NG2+ cells close to the border is particularly stark at E18.5 with up to 30 times more NG2+ cells next to the vessel border than in the tissue on average.

A population of particular interest are the CD45+/Gr-1+ myeloid cells. This cell type is characterized by a low basal expression level at E12.5 and particularly so in the vicinity of the vessel border as shown in figure 10.24c (the fraction of double-positive cells is close to zero near the edge of the tissue). Indeed, these cells are known to proliferate at later stages in the fetal development and the evolution of the fraction of myeloid cells is comparable with what has been reported in the literature [157]. The distribution of these cell types is significantly different at E18.5, with high fractions of CD45+/Gr-1+ cells accumulating between 0 and 5 layers away from the tissue edge (see figure 10.24c).

We can therefore propose a schematic representation of the tissue structure in the vicinity of the vessel border in figure 10.24g. At E12.5, there is a localized overconcentration in NG2+ cells up to one edge away from the border. The tissue composition then rapidly reaches the parenchyma composition of the tissue with no specific pattern with regard to the Gr-1 and CD45 markers. On the other hand, at E18.5, we see that the environment around the FL vessels is highly organized with a double layer rich in NG2+ cells, followed by approximately four layers rich in CD45+/Gr-1+ cells. The emergence of this pattern is consistent with previous studies of the evolution of the FL [157]. Indeed, it has already been reported that granulocytes - that are



FIGURE 10.24: **a** Evolution of the composition of FL blood vessels between E12.5 and E18.5. **b** Composition of the vessel environment in NG2+ cells as a function of the distance from the border at E12.5 and E18.5 (top). The relative composition (fraction of NG2+ cells per layer divided by the average fraction of NG2+ cells in the tissue) as a function of the distance to the border (bottom). **c-d** Similar analysis as in panel (b) conducted on CD45+/Gr-1+ and CD45+ cells. **e-f** Composition of the FL at vessel border at E12.5 and at E18.5. textbfg Schematic representation of the vessel border composition in the FL at E12.5 and E18.5.

Gr-1+/CD45+ - are scattered throughout the tissue from E16.5, concentrating around central veins and in the periphery by E17.5 [172].

10.6.3 The FL mesothelium undergoes significant reorganization between E12.5 and E18.5

The mesothelium designates the outer layer of cells in the fetal liver (Fig. 10.23). The mesothelial cells proliferate during liver development and remain quiescent after birth

[157]. What we see during development is that the composition of the mesothelium undergoes several significant changes during maturation.



FIGURE 10.25: **a** Evolution of the composition of FL mesothelium between E12.5 and E18.5. **b** Composition of the vessel environment in NG2+ cells as a function of the distance from the border at E12.5 and E18.5 (top). The relative composition (fraction of NG2+ cells per layer divided by the average fraction of NG2+ cells in the tissue) as a function of the distance to the border (bottom). **c-d** Similar analysis as in panel (b) conducted on CD45+/Gr-1+ and CD45+ cells. **e-f** Composition of the FL at mesothelial border at E12.5 and at E18.5. textbfg Schematic representation of the mesothelial border composition in the FL at E12.5 and E18.5.

The percentage of NG2+ cells decreases dramatically from close to 15% of the cells at E12.5 to 1% of the cells at E18.5 (Fig. 10.25a). In contrast, there is a sharp 2380% increase of the concentration of CD45+/Gr-1+ cells in the tissues between E12.5 and E18.5, with the double positive cells organizing along the very first layers of the FL tissue at E18.5 as is shown in the composition fraction of figure 10.25c and illustrated

by the mesothelium image in figure 10.17e. This increase is accompanied by a 2380% rise in the concentration of CD45+ cells. In effect, these percentages highlight the fact that Gr-1+ cells and CD45+/Gr-1+ cells are nearly absent from the fetal liver tissue in the earlier stages of development.

Using the network-based approach it is possible to characterize the composition of the mesothelial region. What we show in figure 10.25b and c is that the NG2+ cells at E12.5 ad the CD45+/Gr-1+ cells at E18.5 are concentrated along the external layers of the mesothelium. Compared to the cellular organization around the blood vessels described in the previous section, the length-span of the CD45+/Gr-1+-rich tissues markedly shorter (approx. 1 layers compared to the 4 layers around the portal veins and the central vessels at E18.5).

This enables us to construct a schematized representation of the mesothelium at E12.5 and E18.5 in figure 10.25e. The outermost layers of the tissue are rich in NG2+ cells at E12.5, typically up to two cell layers in depth. Then the tissue composition reaches the composition of the parenchyma, as is clear in the relative composition plots in figure 10.25b-d. At E18.5 we can represent the tissue as a layer rich in NG2+ and CD45+/Gr-1+ double-positive cells before reverting to the average composition of the parenchyma.

The contrast between the very localized CD45+/Gr-1+ rich cell layer on the outer side of the mesothelium and the multiple layers surrounding the portal veins and central vessels irrigating the FL suggest different accumulation mechanisms probably influence the tissue organization. Indeed, it is already known that the CXCL12 cytokine plays a significant role in attracting HSCs to the FL [157]. The multi-layered accumulation around the portal veins and central vessels suggests that a similar mechanism is at stake.

The distribution of the CD45+/Gr-1- cells is also telling: both around vessels and in the mesothelium, it is not influenced significantly by the proximity to the tissue border. Indeed, the concentration of CD45+ cells is similar to that in the parenchyma (see figures 10.24d and 10.25d), but the overall fraction of CD45+ cells increases between E12.5 and E18.5. In contrast to the seemingly homogeneous distribution of CD45+ cells, the distribution of CD45+/Gr-1+ cells is heavily influenced by the proximity to tissue borders, suggesting different differentiation and accumulation mechanisms are at stake between these two populations. For instance, it could be possible that a contactmediated interaction guides the creation of the CD45+/Gr-1+-rich monolayer at the mesothelial border and that the CD45+/Gr-1+ cells accumulate around the vessels driven by chemotactic cues. Further investigation is however required to validate this possibility.

CHAPTER 10. STRUCTURAL ANALYSIS OF THE FETAL LIVER AS A HEMATOPOIETIC EXPANSION NICHE DURING DEVELOPMENT

Key take-aways

Border analysis I have developed a method to analyse the cell composition as a function of the distance to tissue borders. This technique highlights the evolution of the cell composition in regions of interest with time.

Perivascular structure Vessels irrigating the fetal liver see a late-stage (E18.5), multi-layer, accumulation of myeloid cells, irrespective of the vessel type (portal vein or central vessel).

Mesothelium structure The E18.5 mesothelium concentrates myeloid and NG2+ cells in a monolayer. The cell composition rapidly reverts to that of the parenchyma deeper within the tissue.

10.7 Conclusion

Although the fetal liver is an organ of great importance for the development of the hematopoietic system, the precise characterization of the tissue is lacking. This is due to the many technical difficulties preventing easy imaging of the organ. Thanks to the work from Cumano and Pinto-do-Ó's teams at the Pasteur Institute and i3S (University of Porto) we now have good quality images of the fetal liver. This enables the quantitative analysis of its structure thanks to **Griottes**. From this study we learn that there is a significant remodeling of the fetal liver composition between E12.5 and E18.5. Interestingly, myeloid cells - quasi-abscent at E12.5 - cluster around the blood vessels irrigating the fetal liver at E18.5. This is just a peek into the organization of the organ though, and much more remains to be studied.

In this chapter, we focused on the distribution of myeloid cells since they are an important precursor in the hematopoietic system. The next step is to study the distribution of hematopoietic progenitor cells and of lymphoid cells. At the time I am writing these lines, we have analyzed the very first images containing this information. It is my hope that we will gain a more exact overview of the maturation of the fetal liver in the coming months.

We have created several quantitative methods for describing the tissue structure of the FL, from the global cellular composition of the tissue down to the individual immediate environment of specific cell types. We have shown how the composition and the structure of the FL evolves during development, and highlighted the intricate organisation of NG2+ and CD45+/Gr-1+ myeloid cells at E18.5. However, the development of these methods was motivated by the search of HSCs in the FL and the evolution of their immediate environment during fetal maturation.

The main barrier to achieve this goal is the difficulty to successfully stain the FL with the successive antibodies necessary to stain phenotypic HSCs. At the moment at which I am typing these words, we haven't managed to reliably stain tissues with CD150, CD48, CD41, ESAM that would enrich >70% HSCs as CD150+ CD48- CD41-ESAM+, enabling the precise quantification of the HSC micro-environment at different time-points during the FL maturation. A possible alternative would be the analysis of reporter mice that have been used in the adult BM to label HSCs. This could significantly simplify the sample preparation phases thanks to the fluorescence of the HSCs in the liver tissue.

CHAPTER 10. STRUCTURAL ANALYSIS OF THE FETAL LIVER AS A HEMATOPOIETIC EXPANSION NICHE DURING DEVELOPMENT

Part III Bacterial rheotaxis

Chapter 11

Semi-analytical results on bacterial rheotaxis

This chapter will soon be submitted as:

Ronteix G., Lety-Stefanka A., Baroud C. and Amselem G. (2021). *Semi-analytical results on bacterial rheotaxis.*

The motility of microorganisms is a crucial biological function, enabling these microorganisms to explore their environment and to migrate to more favorable habitats [18, 173, 174]. Yet, many microorganisms such as bacteria or microalgae live in aqueous media, where their motility can be drastically affected by the surrounding fluid motion. The most obvious consequence of a surrounding flow on motility is that microorganisms are entrained by the flow. However, microorganisms can also travel perpendicular to the flow direction, due to the combined effects of fluid shear stress and of an asymmetry in the system.

The asymmetry can come from a directional stimulus. For example, the model microalgae *Chlamydomonas reinhardtii* is known to experience gyrotaxis: when the shear due to a downwards Poiseuille flow couples to the action of gravity, algae are focused to the center of the flow [175]. Likewise, when *C. reinhardtii* swimming in a Poiseuille flow are stimulated with a localized source of light, the combination of shear and phototaxis leads to their focusing to the centerline of the Poiseuille profile [176]. But the asymmetry can also come from the microorganism itself: when the helix-shaped spirochete *Leptospira biflexa* is placed in a Poiseuille flow, it drifts perpendicular to the streamlines, a result of the interaction between shear and the chirality of the helix [177]. Motile bacteria in a Poiseuille flow likewise experience a lift force perpendicular to the plane of the shear [178], a behavior termed rheotaxis and due to the interaction between three components: bacterial propulsion, shear, and the chiral geometry of the flagellar helix.

The founding experiment on bacterial rheotaxis was reported by Marcos et al. and consisted in flowing *Bacillus subtilis* bacteria in a Hele-Shaw microfluidic channel [178]. The shear profile of the Poiseuille flow in the channel led to bacterial rheotactic motion towards the side walls of the microchannel. The average rheotactic velocity was quantified experimentally as a function of shear rate, and experiments were in excellent agreement with numerical results. Mathijssen et al. [179] followed up on the study of

bacterial rheotaxis and analyzed the motion of the bacteria *Escherischia coli* in the vicinity of the solid walls of a microfluidic Hele-Shaw cell. They were able to observe and model four different regimes of bacterial swimming as a function of the applied shear rate. More recently, Jing et al. [180] determined experimentally the orientation and rheotactic velocity of chiral bacteria in shear flows, comparing them to numerical results. In these three studies [178–180], the influence of the shear rate on the rheotactic velocity was obtained by numerically solving the equations for swimming at low-Reynolds number, and fitting them to functional forms.

Here, we obtain semi-analytical formula for the rheotactic velocity of bacteria as a function of the shear rate. We begin by analyzing the case where bacteria follow purely deterministic trajectories in a shear flow. Bacterial trajectories are then the result of three compounding effects: (i) bacterial propulsion, which tends to make bacteria swim in a straight line; (ii) Jeffery orbits, due to the elongated shape of bacteria, which make the bacteria rotate in the plane of the shear; and (iii) rheotaxis, due to the chirality of the bacterial flagella, that tends to orient bacteria out of the plane of shear and makes them drift perpendicularly to this plane. In particular, we show that Jeffery effects and rheotaxis occur on two different time scales: the instantaneous dynamics of bacterial trajectories are dominated by Jeffery effects, while the long-term trajectories show a net drift due to rheotaxis. We recover similar dynamics to that obtained analytically very recently by Ishimoto [181].

In a second part, we take into account the effect of bacterial diffusivity on bacterial trajectories. While shear reorients bacteria perpendicularly to the shear plane, bacterial diffusion tends to randomize bacterial orientations. The analytical formula for the rheotactic velocity as a function of shear is plugged into a Fokker-Planck equation, leading to analytical results on the steady-state probability of reorientation as a function of the value of the shear. We can then define a Péclet number comparing the effects of diffusion and rheotaxis: $Pe = \frac{L\beta\dot{\gamma}}{D_z}$, where L is the size of the experimental setup perpendicular to the shear plane, D_z is the diffusion constant, β a numerical prefactor that depends on the bacterial geometry, and $\dot{\gamma}$ is the shear rate. The rheotaxis-induced bacterial population sorting increases with the Péclet number, reaching a plateau for $Pe \gg 1$

Theoretical predictions are compared with experimental data of bacterial rheotaxis in microfluidic droplets, where the steady-state distribution of bacteria experiencing rheotaxis is easier to monitor than in standard open channels. Experimental and theoretical results are in very good agreement.

11.1 Analytical results – deterministic case

11.1.1 Definition of the problem

Rheotaxis originates from an interplay between bacterial chirality, bacterial motility, and the presence of a shear flow. A bacteria is modelled as a spherical head of radius a attached to a helicoidal flagellum, see Fig. 11.1a. Call R the radius of the flagellum, l its total contour length , α its pitch, and n the number of turns it makes. These geometrical parameters are related through $2\pi Rn/l = \sin \alpha$. The location of a point on the helix is defined by its curvilinear coordinate s, with $0 \le s \le l$. Last, a right-handed helix (resp. left) has handedness $\xi = 1$ (resp. $\xi = -1$).

The orientation of the bacteria in the lab frame is defined by the two angles (θ, ψ) , where θ is the azimuthal angle of the bacteria in the *xy*-plane, and ψ is the angle the bacteria makes with the *xy* plane, see Fig. 11.1a. The center of the bacterial head is assumed to lie at the origin of the coordinate system. An external shear flow is imposed in the *xy*-plane, such that the fluid velocity at any point $\mathbf{r} = (r, \theta, \psi)$ is given by $\mathbf{v}_f(\mathbf{r}) = \dot{\gamma} y \mathbf{e}_x = \dot{\gamma} r \cos \psi \sin \theta \mathbf{e}_x$, where $\dot{\gamma}$ is the shear rate.



FIGURE 11.1: (a) The bacteria is modeled as spherical head attached to a helicoidal flagellum. A shear is imposed in the xy-plane, so that the external fluid flow is $\mathbf{v_f} = \dot{\gamma} y \mathbf{e_x}$. The bacterial orientation is defined by the angle θ in the shear plane, and by the pitch angle ψ perpendicular to the shear plane XXXXcorrect definition of ψ . The coordinate system attached to the bacteria is $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$, where $\mathbf{e_1}$ is along the main axis of the bacteria. (b, c) The motion of the bacteria orthogonally to its main axis $\mathbf{e_1}$ is coupled to its rotation through the term B in the motility matrix ζ . (b) A cell with a chiral helix will experience motion upon rotation. (c) For an achiral bacterial flagella, B = 0 and the rotation perpendicular to $\mathbf{e_1}$ doesn't induce any motion of the bacteria. (d) Definition of the orthogonal and parallel unit vectors e_{\perp} and e_{\parallel} along the flagellar helix.

The Reynolds number associated to a bacteria of size $\approx 10 \ \mu m$ swimming at a speed $\approx 20 \ \mu m/s$ in water, is Re $\approx 2 \times 10^{-4} \ll 1$. At this low Reynolds number, the force of the water on the spherical bacterial head is given by Stokes' formula: $\mathbf{F}_{\text{head}} = 6\pi\eta(\mathbf{v}_f(\mathbf{0}) - \mathbf{v}) = -6\pi\eta\mathbf{v}$, where η is the viscosity of the surrounding fluid and \mathbf{v} is the speed of the center of mass of the bacteria, taken to be at the center of the spherical head [182].

Resistive force theory allows to calculate the viscous force on the slender bacterial helix, by decomposing the friction force locally on the helix into two components: one locally parallel to the helix orientation, and one perpendicular to it [182]. The two friction coefficients are ζ_{\parallel} and ζ_{\perp} respectively and are given by:

$$\zeta_{\parallel} = \frac{4\pi\eta}{\log(4\lambda^2/e^2) - 1} \quad \text{and} \quad \zeta_{\perp} = \frac{8\pi\eta}{\log(4\lambda^2/e^2) + 1},$$
(11.1)

where λ is the wavelength of the helix and e is the half-thickness of the helix [182]. At a point P on the helix, defined by its coordinate \mathbf{r}_P , the helix velocity is $\mathbf{v}_h(\mathbf{r}_P)$ and the relative velocity between fluid and helix is $\mathbf{v}_{rel}(\mathbf{r}_P) = \mathbf{v}_f(\mathbf{r}_P) - \mathbf{v}_h(\mathbf{r}_P)$. We define $\mathbf{e}_{\parallel}(\mathbf{r}_P)$ and $\mathbf{e}_{\perp}(\mathbf{r}_P)$ the unit vectors that are tangent and perpendicular to the helix at point P. Resistive force theory gives the friction force locally:

$$\begin{aligned} \mathbf{f}_{\text{helix}}(\mathbf{r}_{P}) = & \zeta_{\perp} \left[\mathbf{v}_{\text{rel}}(\mathbf{r}_{P}) \cdot \mathbf{e}_{\perp} \right] \mathbf{e}_{\perp} + \zeta_{\parallel} \left[\mathbf{v}_{\text{rel}}(\mathbf{r}_{P}) \cdot \mathbf{e}_{\parallel} \right] \mathbf{e}_{\parallel} \\ = & \zeta_{\perp} \mathbf{v}_{\text{rel}}(\mathbf{r}_{P}) + \left(\zeta_{\parallel} - \zeta_{\perp} \right) \left[\mathbf{v}_{\text{rel}}(\mathbf{r}_{P}) \cdot \mathbf{e}_{\parallel} \right] \mathbf{e}_{\parallel}. \end{aligned}$$
(11.2)

The total force on the helix is then $\mathbf{F}_{\text{helix}} = \int_0^l \mathbf{f}(\mathbf{r}(s)) ds$.

Last, let us calculate the total torque on the bacteria. Call ω the rotation rate of the bacterial head, and $\Omega_{\rm f}$ the rotational of the fluid flow at the head center. The torque on the head is given by Faxen's second law: $\mathbf{T}_{\rm head} = 8\pi\eta a^3(\frac{\Omega_{\rm f}}{2} - \omega)$ [183]. The torque on the helix is $\mathbf{T}_{\rm helix} = \int_0^l \mathbf{r}(s) \times \mathbf{f}(\mathbf{r}(s)) ds$.

At low Reynolds number, inertia can be neglected, so that the sum of forces and sum of torques on the bacteria are zero at all times:

$$\int \mathbf{F}_{\text{head}} + \mathbf{F}_{\text{helix}} = 0 \tag{11.3a}$$

$$\left\{ \mathbf{T}_{\text{head}} + \mathbf{T}_{\text{helix}} = 0 \right. \tag{11.3b}$$

Equations 11.3a and 11.3b are a linear system of 6 equations with 6 unknowns, which can be rewritten in matrix form as $\zeta \mathbf{W} = \mathbf{A}$, where ζ is the 6 × 6 mobility matrix of the problem, $\mathbf{W} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)$ is a column vector with all the unknowns that we want to solve for, and \mathbf{A} is a forcing term with contributions from both the external flow and from the autonomous rotation of the helix [20].

This system proves extremely cumbersome to solve in the lab reference frame. Indeed, the helix orientation is characterized by 3 angles: the two angles defining the bacterial orientation (θ, ψ) , as well as a third angle ϕ corresponding to rotation of the helix around its axis, see Fig. 11.1. To go from the helix reference frame to the lab reference frame, one has to multiply the vectors $\mathbf{r}(s)$ and \mathbf{e}_{\parallel} by the Euler rotation

matrix
$$\mathbf{E} = R_{\phi}R_{\psi}R_{\theta}$$
, where $R_{\theta} = \begin{pmatrix} \cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1 \end{pmatrix}$, $R_{\psi} = \begin{pmatrix} \cos\psi & 0 & -\sin\psi\\ 0 & 1 & 0\\ \sin\psi & 0 & \cos\psi \end{pmatrix}$,
 $R_{\phi} = \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\phi & \sin\phi\\ 0 & -\sin\phi & \cos\phi \end{pmatrix}$, and trigonometric terms then pervade the motility matrix

ζ.

The trick is rather to solve the system of equations in the bacteria's reference frame. We define the orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, where \mathbf{e}_1 is along the bacterial axis, \mathbf{e}_2 is along θ , and \mathbf{e}_3 along ψ , see Fig. 11.1a. The motility matrix in the reference frame of the bacteria then has a simple expression:

$$\zeta = \begin{pmatrix} A_f & 0 & 0 & M & 0 & 0 \\ 0 & A & 0 & 0 & B & C \\ 0 & 0 & A & 0 & C & B \\ M & 0 & 0 & F & 0 & 0 \\ 0 & B & C & 0 & D & 0 \\ 0 & C & B & 0 & 0 & D \end{pmatrix},$$
(11.4)

where the terms are given by:

$$A_f = -6\pi\eta a + l\zeta_{\parallel} \left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}} \sin\left(\alpha\right)^2 \right)$$
(11.5)

$$M = \frac{1}{2\pi} l^2 \zeta_{\parallel} \left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}} \right) \cos\left(\alpha\right) \sin\left(\alpha\right)^2$$
(11.6)

$$A = -6\pi\eta a - \frac{1}{4}l\zeta_{\parallel} \left[1 + 3\frac{\zeta_{\perp}}{\zeta_{\parallel}} + \left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}} \right)\cos\left(2\alpha\right) \right]$$
(11.7)

$$B = \frac{1}{4\pi} l^2 \zeta_{\parallel} \left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}} \right) \cos\left(\alpha\right) \sin\left(\alpha\right)^2$$
(11.8)

$$C = \frac{1}{4}l^2 \zeta_{\parallel} \cos\left(\alpha\right) \left[-2\frac{\zeta_{\perp}}{\zeta_{\parallel}} + \left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}}\right) \sin\left(\alpha\right)\right]^2$$
(11.9)

$$F = -8\pi\eta a^3 - \frac{1}{2}lR^2\zeta_{\parallel} \left[1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}} + \left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}}\right)\cos\left(2\alpha\right)\right]$$
(11.10)

$$D = -8\pi\eta a^{3} - \frac{1}{24}l\zeta_{\parallel} \left[-12R^{2}\frac{\zeta_{\perp}}{\zeta_{\parallel}} + 4\left(3R^{2}\left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}}\right) - 2l^{2}\frac{\zeta_{\perp}}{\zeta_{\parallel}}\right)\cos\left(\alpha\right)^{2} + l^{2}\left(-1 + \frac{\zeta_{\perp}}{\zeta_{\parallel}}\right)\sin\left(2\alpha\right)^{2} \right]$$
(11.11)

11.1.2 Brief Observations

The motility matrix highlights key aspects of bacterial motility. The coupling between the bacteria rotation and the velocity along the bacteria's main axis \mathbf{e}_1 is given by M. We immediately verify that the bacteria doesn't undergo any motion when the helix is not chiral ($\alpha = 0$ or R = 0), or in the absence of a helix (l = 0). The asymmetry in the friction coefficients is likewise crucial for motion at low Reynolds number: M = 0when $\zeta_{\parallel} = \zeta_{\perp}$.

For a non-chiral geometry of the bacteria with a helix angle of $\alpha = 0$, the coupling between the motion along \mathbf{e}_1 and $\mathbf{\Omega}_1$ disappears. We also see that the coupling between \mathbf{e}_3 and $\mathbf{\Omega}_3$, as well as between \mathbf{e}_2 and $\mathbf{\Omega}_2$, is determined by the coefficient B (11.8) in the motility matrix. It is this coupling that causes the motion of the bacteria out of the plane it is rotating in, leading to a net motion orthogonally to its rotation plane (figure 11.1 b). Inversely, a radially symmetric object, equivalent to $\alpha = 0$, will have a coefficient B = 0, which leads to no movement out of the plane containing the rotation (Fig. 11.1c).

We use Mathematica to solve the 6×6 system of equations (11.3) in the reference frame of the bacteria (\mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3), giving us access to the velocity and angular velocity of the bacterial head in this reference frame. Then, by returning to the laboratory frame we access the velocity and angular velocity of the bacterial head in the lab frame. In particular, we obtain the rheotactic velocity of the bacteria in the z direction as a function of θ and ψ , coming from the shear on the helix in the xy plane.

The obtained expressions for the rheotactic velocity \dot{z} as well as for the angular velocities $\dot{\theta}$ and $\dot{\psi}$ are extremely complex; they can be accessed in the Mathematica notebook in the Supplementary Material. Plugging in numbers for the geometry of the bacteria allows to obtain semi-analytical results for the rheotactic velocity as a function of the shear rate.

11.2 Retrieving the instantaneous bacterial dynamics

11.2.1 Plugging in the geometric properties

Experimentally, the ratio between the parallel and perpendicular friction coefficients for a bacterial flagellum is $\zeta_{\parallel}/\zeta_{\perp} \approx 1.7$ [178]. Considering a bacterial head with $a \approx 1 \ \mu m$, a left-handed helix ($\chi = -1$) of length $l \approx 15 \ \mu m$, half-thickness $e \approx 240$ nm and wavelength $\lambda \approx 3 \ \mu m$, we find $\zeta_{\parallel} \approx 2\eta$, where $\eta = 10^{-3}$ Pa.s is the dynamic viscosity of water. The motility matrix then contains several coefficients which are the sum of two terms of the same order of magnitude: $6\pi\eta a \approx 2 \times 10^{-8} \text{ kg.s}^{-1}$ and a term of order $l\zeta_{\parallel} \approx 3 \times 10^{-8} \text{ kg.s}^{-1}$.

Replacing all geometric characteristics with their numerical values, and after rounding off to the two most important terms, the angular velocities and the rheotactic velocity of the bacteria in the lab frame are found to be:

$$\dot{\theta} = A_1 \dot{\gamma} + A_2 \dot{\gamma} \sin\left(\theta\right)^2 \tag{11.12}$$

$$\dot{\psi} = B_{\text{Jeff}} \dot{\gamma} \sin(2\psi) \sin(2\theta)$$

$$+ B_{\text{rheo}}\dot{\gamma}\cos(\psi)\cos(2\theta) \qquad (11.13)$$

$$\dot{z} = v_{bac,z} + C_{\text{Jeff}}\dot{\gamma}\sin(2\theta)\cos(\psi)\sin(2\psi) + C_{\text{rheo}}\dot{\gamma}\cos(2\theta)\cos(\psi)^{2}. \qquad (11.14)$$

The constants A_2 , B_{Jeff} and C_{Jeff} are non-zero only when the object studied is elongated. The constants B_{rheo} and C_{rheo} are non-zero when the object is chiral. Experimental values for the constants in Eq. (11.12) -(11.14) are $A_1 \approx 0.02$, $A_2 \approx 0.94$, $B_{\text{Jeff}} \approx 0.94$, $B_{\text{rheo}} \approx 4 \cdot 10^{-3}$, $C_{\text{Jeff}} \approx 9 \cdot 10^{-7}$ m and $C_{\text{rheo}} \approx -1.6 \cdot 10^{-8}$ m.

Studying these equations allows us to shed some light on the intricate dynamics of the bacteria. In the common case of an object with a radial symmetry axis (for instance an ellipsoid) subject to shear flow, the shear forces have the shear plane (xOy) as a symmetry plane. These forces being polar, we expect the effects to be symmetric with regard to the (xOy) plane according to the Curie principle [184], and one can verify that the Jeffery orbits don't undergo any net motion in the z-direction. The addition of a chiral helix to the object adds a new rheotactic term in the equation that breaks the symmetry of the problem, the helix itself not being symmetric with regard to the (xOy) plane. In the following paragraphs we will demonstrate the effect of these rheotactic terms on the bacterial dynamics.

The complexity of the motion of bacteria in shear flow previously required the use of fully numerical methods to derive bacterial dynamics [20, 179, 180]. By changing the reference frame we have derived the corresponding semi-analytical equations. To check the approximated results in Eq. (11.12)-(11.14), we ran numerical simulations of bacterial motion in a shear flow. Simulation results enabled to retrieve the angular dynamics as well as the rheotactic force on bacteria. Analytical and simulation results are in excellent agreement, as detailed below.

11.2.2 Bacterial rotation in the shear plane θ

The dynamics of θ described by Eq. (11.12) are characteristic of the motion of an elongated object in a shear flow at low Reynolds number, first studied by Jeffery [185].

In such a configuration, the object rotates in the plane of the shear, with a rotation rate proportional to the shear rate $\dot{\gamma}$. The constants A_1 and A_2 are related to the the effective aspect ratio $a_r = \sqrt{\frac{A_2+A_1}{A_1}}$ of the bacteria. Eq. (11.12) can in fact be rewritten as a function of a_r and $\dot{\gamma}$ [186]:

$$\dot{\theta} = \frac{\dot{\gamma}}{1+a_r^2} \left(\cos\left(\theta\right)^2 + a_r^2 \sin\left(\theta\right)^2 \right), \qquad (11.15)$$

The evolution of θ is then periodic with a period $T = \frac{2\pi}{\dot{\gamma}} \left(a_r + \frac{1}{a_r} \right)$, and given by [186]:

$$\tan\left(\theta\right) = a_r \tan\left(\frac{2\pi t}{T}\right). \tag{11.16}$$

The effective aspect ratio of the bacteria is $a_r = \sqrt{\frac{A_2+A_1}{A_1}} \approx 6.3$. Note that this rotation in the shear plane does not depend on the chirality of the object, but solely on its aspect ratio.

Results of numerical simulations of bacterial motion in a shear flow show excellent agreement with the analytical functional form given in Eq. (11.12), see Fig. 11.2a. An elongated object spends most of its time aligned with the shear ($\theta = 0$ and $\theta = \pm \pi$), independently of its chirality, see Fig. 11.2a and [185].

Note that the analytical formula obtained by Ishimoto for the dynamics of the orientation angle $\dot{\theta}$ of a helicoidal object in a shear flow has an additional term [181]; as calculated by Ishimoto however, this term is negligible for realistic bacterial geometries.



FIGURE 11.2: Angular dynamics of bacteria with a right-handed helix in a shear flow. Black dots: results of numerical integration of the equations for bacterial motion as used in [20, 179, 180]. Dashed red line: semi-analytical results. Dotted blue line: analytical result for an achiral elongated particle. (a) The evolution of the angle in the shear plane θ is independent of the chirality of a particle and is given solely by Jeffery dynamics, see Eq. (11.12). (b) An achiral particle experiences no tilt force out of the shear plane ($\dot{\psi} = 0$, blue dotted line), in contrast to a bacteria (black dots and red dashed line). (c) An achiral particle in the shear plane stays in the shear plane (dotted blue line), while a chiral particle experiences a rheotactic tilt (black dots and red dashed line).

11.2.3 Instantaneous dynamics of the pitch angle ψ

The dynamics of ψ , the angle made by the bacteria with the shear plane, result from two independent contributions: $\dot{\psi} = \dot{\psi}_{\text{Jeff}} + \dot{\psi}_{\text{rheo}}$, with $\dot{\psi}_{\text{Jeff}} = B_{\text{Jeff}}\dot{\gamma}\sin(2\psi)\sin(2\theta)$, and

 $\dot{\psi}_{\text{rheo}} = B_{\text{rheo}}\dot{\gamma}\cos(\psi)\cos(2\theta)$. The first term was already discussed by Jeffery [185], and is stems from the interaction between shear and an elongated object. It depends solely on the aspect ratio of the elongated particle through the constant B_{Jeff} . The second term in Eq. (11.13) is due to the chirality of the helix, and causes the bacteria to tilt away from the shear plane. For a realistic bacterial geometry, the instantaneous dynamics of ψ are dominated by Jeffery effects: $\dot{\psi}_{\text{Jeff}} \approx 0.94 \gg \dot{\psi}_{\text{rheo}} \approx 4 \cdot 10^{-3}$.

Numerical simulations of bacterial motion are again in excellent agreement with the semi-analytical results and highlight the specific dynamics caused by the rheotactic force. The dependency of $\dot{\psi}$ with ψ for a bacteria aligned with the shear ($\theta = 0$, the orientation in the shear plane in which bacteria spends most of its time) shows that the rheotactic term induces a bacterial tilt out of the shear plane, until the bacteria reaches an angle $\psi = \pm \pi/2$, see Fig. 11.2b. The angular dynamics of $\dot{\psi}$ also depends on the angle θ of the bacteria in the shear plane (see Fig. 11.2c). Note that for an achiral object, $\dot{\psi}(\psi = 0) = \dot{\psi}_{\text{Jeff}}(\psi = 0) = 0$: an achiral object in the shear plane stays in the shear plane. Likewise, an achiral object making an angle ψ with the shear plane, and aligned with the shear ($\theta = 0$) does not have a tendency to be tilted out of the shear plane: $\dot{\psi}_{\text{Jeff}}(\theta = 0) = 0$.

11.2.4 Instantaneous dynamics of the position z

The dynamics of z are described by Eq. (11.14) and depend on three contributions: the rotation of the helical flagella, which induces a bacterial velocity in z, $\dot{z}_{bac} = v_{bac,z}$; a term due to Jeffery orbits $\dot{z}_{\text{Jeff}} = C_{\text{Jeff}}\dot{\gamma}\sin(2\theta)\cos(\psi)\sin(2\psi)$; and a term due to the chirality of the helix $\dot{z}_{\text{rheo}} = C_{\text{rheo}}\dot{\gamma}\cos(2\theta)\cos(\psi)^2$. Note that $C_{\text{Jeff}} \gg C_{\text{rheo}}$, and so for most angles (θ, ψ) , the instantaneous dynamics of z are dominated by Jeffery effects. For a realistic bacterial geometry and a helix rotation rate of $\Omega_{\text{helix}} = 100 \text{ s}^{-1}$, $\dot{z}_{bac} \approx 6 \times 10^{-6} \text{ m.s}^{-1}$, which is of the same order as the Jeffery contribution. To clarify the interplay between the Jeffery and rheotactic effects, we assume for the time being that the bacteria is not self-propelled: $v_{bac,z} = 0$.

The relative contributions of the Jeffery and rheotactic effects are plotted in Fig. 11.3. For small values of the pitch angle ψ , the velocity in z is dominated by the rheotactic drift (Fig. 11.3a), while for $\psi > C_{\rm rheo}/C_{\rm Jeff} \approx 0.02$, the z-velocity is dominated by Jeffery effects (Fig. 11.3b). Note that the maximum drift velocity in z is an order of magnitude larger for $\psi = \pi/3$ (Fig. 11.3b) than for $\psi \approx 0$ (Fig. 11.3a), showing that most of the instantaneous motion perpendicular to the shear plane occurs for bacteria tilted with respect to the shear plane, and is due to Jeffery effects.

The influence of the pitch angle on the drift of a particle aligned with the shear $(\theta = 0)$ is shown in Fig. 11.3c. An achiral elongated particle aligned with the shear does not experience any drift along z (blue dotted line), while introducing chirality leads to an instantaneous rheotactic drift (red dashed line) whose value peaks at $\psi = 0$, see Fig. 11.3c.

To summarize, the instantaneous drift in z is due do rheotactic effects for $\psi \approx 0$ and to Jeffery effects otherwise. The maximum rheotactic drift occurs for $(\theta, \psi) = (0, 0)$.



FIGURE 11.3: Instantaneous rheotactic drift in z for a bacteria with a right-handed helix, as a function of the angle in the shear plane θ and the tilt angle ψ . Black dots: numerical resolution of the bacterial dynamics. Dashed red line: semi-analytical result for a bacteria. Dotted blue line: Jeffery contribution. (a) At $\psi = 0$, the rheotactic drift is dominated by rheotactic effects (red dashed line), while Jeffery effects do not contribute (blue dotted line). (b) At $\psi = \pi/3$, the rheotactic drift is dominated by Jeffery effects (blue dotted line), while rheotactic effects are negligible (red dashed line). (c) A bacteria aligned with the shear ($\theta = 0$) experiences a force that pushes it along -z.

11.3 Long-term bacterial dynamics

11.3.1 Analytical trajectories

The instantaneous bacterial dynamics tend to show that dynamics are dominated by Jeffery effects. Yet, Jeffery effects do not lead to an overall net motion in the plane perpendicular to the shear [187]. This stands in contrast to experimental results of bacterial rheotaxis in a shear flow [20, 179, 180], which empirically demonstrate a net movement of bacteria orthogonal to the shear plane. The apparent contradiction can be overcome when noticing that during a rotation of the bacteria in the shear plane, the Jeffery terms cancel out: $\int_0^T \dot{\psi}_{\text{Jeff}} dt = 0$, but that the – smaller in magnitude – rheotactic terms do not: $|\int_0^T \dot{\psi}_{\text{rheo}} dt| > 0$ (11.12). The rheotactic forces exerted on the bacterial flagella create a torque on the bacteria, leading to a progressive tilt in the pitch angle ψ (Fig. 11.4a). The evolution of the bacteria in z is thereby linked to its evolution in ψ , as demonstrated by the numerical integration of the instantaneous equations of motion which reveal a net drift in ψ and z, see Fig. 11.4b and Fig. 11.5a.

The evolution of the angle θ does not depend on the chirality of the object in the shear flow, and follows a periodic trajectory of period T, see Fig. 11.4b and Eq. (11.12). The pitch angle ψ also shows oscillations with time, with a period T/2, see Fig. 11.5a and Eq. (11.13). The trajectory of ψ depends on the chirality χ of the object: for a chiral particle ($\chi \neq 0$), the value of ψ displays a net drift superimposed to its oscillations (see blue line in Fig. 11.4b), while the trajectory of an achiral elongated particle ($\chi = 0$) does not display any overall drift (see red line in Fig. 11.4b and [187]). The long-term drift due to the chirality of the object is better seen when averaging the value of the pitch angle over several periods of rotation T, enabling to get rid of the oscillations due to Jeffery effects. Such a moving average of the pitch angle ψ over 3 periods of rotation T is shown in Fig. 11.5b, highlighting the net drift in ψ . The precise dynamics of the pitch angle ψ depend on the initial value of the angle ψ_0 , and in all



FIGURE 11.4: **a** The rheotactic force is exerted on the bacterial flagella, causing the bacteria to move in the z direction and exerting a torque tilting the bacteria and aligning it with the z-axis. **b** The evolution of θ and $\dot{\theta}$ is periodic with a period T given by Jeffery orbits (see main text).

cases the pitch angle converges to a steady-state value $\psi = \pm \pi/2$. For a chiral object in a shear flow with a rate $\dot{\gamma}$ such that $\chi \dot{\gamma} > 0$, the steady-state value is $\psi = \pi/2$ (see Fig. 11.5b). When $\chi \dot{\gamma} < 0$, the steady-state value of ψ is $\psi = -\pi/2$



FIGURE 11.5: **a** The rheotactic force causes a long-term drift of bacteria in z, in addition to the periodic Jeffery oscillations (blue line). An achiral particle shows solely periodic oscillations of z (red line). **b** Average position of the bacteria z (averaged over 7 periods of oscillation T). For a left-handed bacteria in a shear flow with a positive shear rate $\dot{\gamma}$, the position z decreases before stabilizing to a steady-state value that depends on the original orientation of the bacetria.

The position z of the particle likewise shows oscillations with a period T/2, for both chiral ($\chi \neq 0$) and achiral ($\chi = 0$) elongated particles, see Fig. 11.6b. The introduction of chirality leads to an overall drift of the particle along z, perpendicular to the shear plane, see Fig. 11.6b. As for the pitch angle, the long-term drift in z is better seen by averaging z over three periods of rotation T, to get rid of the oscillations due to Jeffery orbits. The position z evolves until stabilizing at a value that depends on the value

of the initial angle ψ_0 , see Fig. 11.7a. This stable position is reached when the chiral particle has stabilized itself at an angle $\psi = \pi/2$. There are then no more rheotactic effects: Eq (11.13) and (11.14) show that $\dot{\psi} \sim \cos \psi$ and $\dot{z} \sim \cos^2 \psi$, both of which are 0 for $\psi = \pm \pi/2$. The distance traveled in z then depends on the time needed for ψ to reach its stable-state $\psi = \pi/2$: the further away ψ_0 is from $\pi/2$, the longer the time to reach the stead-state in ψ and therefore the further the drift in z.



FIGURE 11.6: **a** Pitch angle ψ of the bacteria as a function of time and of the chirality of the helix for an initial position $(\theta_0, \psi_0) = (0, \pi/10)$. Achiral bacteria (red line, $\chi = 0$) experience periodic variations of the pitch angle but no average drift, as expected for particles undergoing Jeffery orbits. In contrast, chiral bacteria (blue line, $\chi \neq 0$) experience quasiperiodic variations of the pitch angle with a long-term drift. **b** ψ converges to the stable point $\psi_{inf} = \pi/2$ for chiral bacteria. The drift direction depends on the product $\dot{\gamma}$ and the chirality of the helix.

To represent the dynamics of the bacteria we project the values taken by (θ, ψ) on the unit sphere, such that $\tilde{x} = \cos(\psi)\cos(\theta)$, $\tilde{y} = \cos(\psi)\sin(\theta)$, $\tilde{z} = \sin(\psi)$. The trajectory taken by the bacterial orientation is represented in figure 11.7a: starting from $(\theta, \psi) = (0, \pi/10)$ the helix oscillates pseudo-periodically before stabilizing at $\psi = \pi/2$. This contrasts with the periodic oscillations in orientation of an achiral bacteria (Fig. 11.7b) that have been described by Jeffery [185].

Note that the dynamics described here reflect the behavior of non-motile bacteria, for which $v_{bac,z} = 0$. Adding self-propulsion leads to a long-term drift along the axis of the bacteria, which is perpendicular ($\psi = \pi/2$) to the shear plane, so that bacteria in a Poiseuille flow drift all the way to the channel sides [20, 179, 180].

11.3.2 The effect of diffusion

Rotational diffusion in (θ, ψ) can be neglected

Bacterial trajectories are not purely deterministic: acteria experience rotational diffusion, and their motion follows a run-and-tumble mechanism which leads to an effective


FIGURE 11.7: (**a** and **b**) Evolution of the orientation of a chiral (g) and achiral (h) bacteria as a function of time (from blue to red). An achiral bacteria reproduces periodic Jeffery oscillations.

randomization of their trajectories, characterized by an effective diffusion coefficient. The distribution $\rho(\theta, \psi, t)$ of the angles (θ, ψ) satisfies the following Fokker-Planck equation [186]:

$$\frac{\partial \rho}{\partial t} = \mathcal{D}_R \Delta \rho - \nabla \cdot (\dot{\omega} \rho), \qquad (11.17)$$

where \mathcal{D}_R is the effective diffusion coefficient of the bacteria and $\dot{\omega} = (\dot{\theta}, \dot{\psi})$ is their angular velocity. The rotational diffusion coefficient \mathcal{D}_R is given by the Stokes-Einstein relationship $\mathcal{D}_R = k_B T/D$, $k_B = 1.38 \times 10^{-23}$ SI is Boltzmann's constant, T = 293 K is the ambiant temperature, and D is the last diagonal term of the motility matrix ζ , see Eq. (11.4) [186]. We find $\mathcal{D}_R \approx 1.8 \text{ rad}^2 \cdot \text{s}^{-1}$, in good agreement with the previously obtained value $\mathcal{D}_R = 2.3 \pm 0.3 \text{ rad}^2 \cdot \text{s}^{-1}$ [188].

The angular rate of change $\dot{\psi}(\psi, \theta)$ is the sum of two terms, of orders of magnitude $B_{\text{rheo}}\dot{\gamma}$ and $B_{\text{Jeff}}\dot{\gamma}$. For a typical value of $\dot{\gamma} = 10 \text{ s}^{-1}$, we have $B_{\text{rheo}}\dot{\gamma} \ll \mathcal{D}_R$; rotational diffusion reorients bacteria much faster than rheotactic effects, and the rheotactic term can be neglected in front of the diffusive term in Eq. 11.17. In contrast, $B_{\text{Jeff}}\dot{\gamma} \gg \mathcal{D}_R$, and Jeffery effects dominate diffusion. In this case, the angular distribution $\rho(\theta, \psi)$ is well approximated by classical Jeffery orbits, and given by [186]:

$$\rho(\theta, \psi) = \frac{1}{a_r} \left[4\sin^2(\psi)(\sin^2(\theta) + a_r^{-2}\cos^2(\theta)) + \cos^2(\psi) \right]^{-3/2}.$$
 (11.18)

Note that the probability distribution $\rho(\theta, \psi)$ in the case where $B_{\text{Jeff}}\dot{\gamma} \approx \mathcal{D}_R$ is much more complex and can be found in [186]. For realistic bacterial geometries, the aspect ratio $a_r \gg 1$. The probability distribution $\rho(\theta, \psi)$ is then peaked around $\theta = 0$ and $\theta = \pi$, highlighting that the object tends to spend most of its time aligned with the shear (Fig. 11.8a). This result is in good agreement with previous experimental observations [179, 180].

Average rheotactic drift \dot{z} over a population of bacteria describing Jeffery orbits

The bacterial drift in the z direction, perpendicular to the shear plane, depends on the intrinsic bacterial propulsion along z, as well as Jeffery and rheotactic drifts. Bacterial



FIGURE 11.8: (a) Probability density map of the orientation of a bacteria in a shear flow, as predicted by Leal and Hinch [186]. (b) Drift of a bacteria (color-coded) due to Jeffery effects, as a function of bacterial orientation. (c) Drift of a bacteria (color-coded) due to rheotactic effects, as a function of bacterial orientation

propulsion is independent of the surrounding fluid velocity and is neglected in the following analysis. The instantaneous drift velocity \dot{z} then depends on the angles θ and ψ according to $\dot{z} = C_{\text{Jeff}}\dot{\gamma}\sin(2\theta)\cos(\psi)\sin(2\psi) + C_{\text{rheo}}\dot{\gamma}\cos(2\theta)\cos(\psi)^2$, see Eq. (11.14). The Jeffery term is plotted in the (θ, ψ) plane in Fig. 11.8b, highlighting that the sign of the drift velocity due to Jeffery effects depends on the sign of the pitch angle ψ . In contrast, the direction of the rheotactic drift does not depend on the drift angle ψ , see Fig. 11.8c. The average rheotactic speed is calculated by averaging the instantaneous speed $\dot{z}(\theta, \psi)$ over all bacterial orientations (θ, ψ) , weighted by the probability of a bacterial orientation $N(\theta, \psi)$:

$$\langle \dot{z} \rangle (\dot{\gamma}) = \int \dot{z}(\theta, \psi) N(\theta, \psi) \, d\psi \, d\theta,$$
 (11.19)

where $\dot{z}(\theta, \psi)$ is given by Eq. (11.14). This enables us to numerically estimate the integral (11.19):

$$\langle \dot{z} \rangle (\dot{\gamma}) = -8.1 \cdot 10^{-9} \dot{\gamma} \text{ m/s}$$
 (11.20)

The semi-analytical derivation of bacterial dynamics allows to recover the rheotactic speed of a bacteria and show that it is a linear function of the shear rate. The value of the numerical prefactor in Eq. (11.20) depends on the geometry of the bacteria in a non-trivial way, through the constants C_{Jeff} , C_{rheo} and the aspect ratio a_r . We call β this prefactor, so that $\langle \dot{z} \rangle (\dot{\gamma}) = \beta \dot{\gamma}$.

Adding diffusion in z

The above analysis is purely deterministic, where the speed \dot{z} is independent of space and time and depends only on the bacterial geometry and the surrounding shear rate. Yet, bacteria are subject to rotational diffusion, and their run-and-tumble motion leads to a randomization of their trajectories that can be characterized by an effective diffusion coefficient. To understand the effect of noise on bacterial rheotaxis, we call p(z)the probability density of finding a bacteria at z. This probability obeys the following conservation equation:

$$\frac{\partial p}{\partial t} = D_z \Delta p(z) + \nabla (\dot{z} p(z)).$$
(11.21)

 D_z is the translational diffusion coefficient of the bacteria in the z-axis.

In a bounded experimental setting, the probability flux is zero at the boundaries: $\partial p/\partial z = 0$ (non-penetration condition). Moreover, the rotation dynamics of the bacteria are significantly more rapid than the translation motion of the bacteria $(C_{\text{Jeff}}, C_{\text{rho}} \ll B_{\text{Jeff}}, A_1, A_2)$, so that the adiabatic approximation can be used, meaning we can replace \dot{z} with $\langle \dot{z} \rangle$ in equation (11.21). The average rheotactic speed from Eq. (11.19) is plugged into Eq. (11.21), allowing us to find the steady-state probability distribution along the z-axis as a function of $\dot{\gamma}$:

$$p(z) = \frac{e^{-z/L}}{L \, 1 - e^{-}},\tag{11.22}$$

where L is the size of the experimental system along z, and we have introduced the characteristic Péclet number $= \frac{L\langle \dot{z} \rangle}{D_z}$ as the key parameter defining the bacteria probability density along the z-axis in Eq. (11.22). When $\ll 1$, diffusion dominates rheotaxis and we can verify that the bacteria are uniformly distributed long z: $p(z) \sim \frac{1}{L}$. On the other hand, if $\gg 1$, then $p(z) \sim \delta(z, L)$ and we see that the probability distribution peaks at z = L: all bacteria experience rheotactic drift and accumulate at the boundary. The Péclet number depends on the shear rate $\dot{\gamma}$ through the average rheotactic speed \dot{z} : Pe = $\frac{L\beta\dot{\gamma}}{D_z}$, where β is a numerical prefactor depending on the geometry of the bacteria, see Eq. (11.19) and (11.20).

11.4 Experimental results

The fraction of bacteria undergoing rheotaxis as a function of shear was experimentally investigated using a droplet microfluidics setup. Bacteria were encapsulated in microfluidic droplets of radius $R \approx 75 \ \mu m$, which were immobilized in square grooves etched in the microfluidic channel ceiling. The grooves, of side 150 μm and depth 100 μm , acted as surface-energy anchors [71, 189]. Each droplet filled entirely its anchor, and bulged out in the main channel of height $L = 35 \ \mu m$, see Fig. 11.9a-b.

In the absence of an outer oil flow, bacteria swam freely within their droplets at a speed $v_{\rm bac} \approx 30 \ \mu {\rm m.s^{-1}}$, and the distribution of bacteria in a droplet was uniform, see the first snapshot in Fig. 11.9c. Imposing an oil flow in the microfluidic chamber led to a recirculation within the immobilized droplets [190]. The spatial distribution of bacteria within the droplets then changed drastically and became asymmetric: within 5 seconds after imposing the oil flow, bacteria segregated to one side of the droplet, see the middle snapshots in Fig. 11.9c for an average shear rate within the droplet $\dot{\gamma} \approx 2.8 \ {\rm s}^{-1}$. Reversing the flow direction led to a reversal of the bacterial segregation pattern, see the last snapshot in Fig. 11.9c for an average shear rate within the droplet $\dot{\gamma} \approx -2.8 \ {\rm s}^{-1}$. Probability density maps of the location of bacteria in the droplet highlight the segregation pattern in the presence of flow, with all bacteria being found in the top half (resp. bottom half) of the droplet for a shear rate $\dot{\gamma} \approx 2.8 \text{ s}^{-1}$ (resp. $\dot{\gamma} \approx -2.8 \text{ s}^{-1}$), see Fig. 11.9d.

The flow pattern within an anchored droplet or within a droplet moving in a Hele-Shaw cell is complex and three-dimensional [190, 191]. As a first approximation, the recirculation can be assumed to exist solely in the part of the droplet bulging out of the groove, in the main channel of height $L = 35 \ \mu m$; The recirculation can be neglected in the part of the droplet that is in the groove (data not shown). As a further simplification, for $0 \le z \le L$, the flow in the droplet was assumed to be two dimensional, in the xy plane, and such that the shear rate $\dot{\gamma}$ is equal and opposite in the two halves of the droplet, as indicated Fig. 11.9c [190, 191].

The bacterial asymmetry was quantified by calculating the difference in number of bacteria between top and bottom half of the droplet, and normalizing it by the total number of bacteria: calling N_+ the number of bacteria in the top half of the droplet (y < 0) and N_- the number of bacteria in the bottom half of the droplet (y < 0), the asymmetry was defined as $\alpha = \frac{N_+ - N_-}{N_+ + N_-}$. The asymmetry is then $\alpha = -1$ when all bacteria are in the lower half of the droplet, $\alpha = +1$ when they are all in the top half of the droplet, and $\alpha = 0$ when bacteria are at random positions in the droplet. The evolution of the asymmetry with time is shown in Fig. 11.10a for a mean oil speed of $u_{\rm oil} \approx 210 \ \mu {\rm m.s}^{-1}$, corresponding to a shear rate within the droplet of $\dot{\gamma} \approx u_{\rm oil}/R \approx 2.8 \ {\rm s}^{-1}$. In the absence of an oil flow $(t < 12 \ {\rm s})$, the asymmetry fluctuated around 0. Once the oil flow was applied, the asymmetry increased within 10 s to its steady-state value of ≈ 0.75 . The steady-state value of the asymmetry depended on the oil flow rate: the faster the oil speed $u_{\rm oil}$, the greater the shear rate $\dot{\gamma} \approx u_{\rm oil}/R$ in the droplet, and the greater the steady-state bacterial asymmetry, until reaching a plateau, see Fig. 11.10a. The asymmetry plateaus at $\alpha = 0.75 < 1$ because of the presence of bacteria close to the centerline of the droplet, that are entrained by secondary flows in the droplet.

To compare experimental and theoretical results, note that the two halves of the droplet have equal and opposite shear rates $\dot{\gamma}$. When the outer oil flowed from left to right, the Péclet number was therefore proportional to $\dot{\gamma}$ in the top half of the droplet (y > 0), and to $-\dot{\gamma}$ in the bottom half of the droplet (y < 0). As a first approximation, the flow speed along x in the droplet is given by $v_x = \dot{\gamma}(|y| - R/2)$. Calling Pe = $\beta \dot{\gamma} L/D_z$, the probability of finding a bacteria at a height z in the top-half (y > 0) of the droplet is then:

$$p_{y>0}(z) = \frac{\text{Pe}}{L} \frac{e^{-z\text{Pe}/L}}{1 - e^{-\text{Pe}}}$$
(11.23)

and the probability of finding a bacteria at a height z in the bottom-half of the droplet is:

$$p_{y<0}(z) = -\frac{\operatorname{Pe}}{L} \frac{e^{z\operatorname{Pe}/L}}{1 - e^{\operatorname{Pe}}}$$
 (11.24)

Calling Δh the depth of field of the microscope, the average asymmetry is then given by:

$$\left\langle \frac{N_{+} - N_{-}}{N_{+} + N_{-}} \right\rangle = \frac{\int_{L-\Delta h}^{L} p_{y>0}(z)dz - \int_{L-\Delta h}^{L} p_{y<0}(z)dz}{\int_{L-\Delta h}^{L} p_{y>0}(z)dz + \int_{L-\Delta h}^{L} p_{y<0}(z)dz}$$
(11.25)

$$\approx \frac{p_{y>0}(L) - p_{y<0}(L)}{p_{y>0}(L) + p_{y<0}(L)}$$
(11.26)

$$=\frac{2-e^{\rm Pe}-e^{-\rm Pe}}{e^{\rm Pe}-e^{-\rm Pe}}$$
(11.27)

A fit of the experimental data to Eq. (11.27) is shown in Fig. 11.10b. The Péclet number is given by $Pe = \beta \gamma L/D$, where the value of $\beta = 8.1 \times 10^{-9}$ m was previously calculated and $L = 35 \ \mu m$ is the main channel height, in which the recirculation exists. The fit gives an effective value of the diffusion coefficient $D_z = 2.6 \ \mu m^2/s$, a value significantly lower than the effective diffusion coefficient of a bacteria executing a runand-tumble motion ($\approx 100 \ \mu m^2/s$ [192]). This can probably be attributed to the fact that the flow is three-dimensional and more complex than assumed, so that the simple linear estimation of the shear rate used above is not correct. Moreover, recirculations in the third dimension will make bacteria move along z faster than by pure rheotaxis.

Comments

On the diffusion coefficient The discrepancy in the diffusion coefficient can have another explanation. It is possible that the mechanism suggested to explain the orientation of the bacteria is wrong. We are currently investigating the dynamics of ψ , the bacteria angle relative to the shear plane. A more precise estimation of this angle as a function of the shear stress could yield a more credible diffusion coefficient D_z .

11.5 Summary and outlook

In this article, we have provided a semi-analytical framework to calculate the trajectories of bacteria in a shear flow, which enabled us to compute the probability of bacteria experiencing rheotaxis as a function of the applied shear rate. Theoretical results on the fraction of bacteria displaying a rheotactic drift were favorably compared to experimental data of bacterial rheotaxis in microfluidic droplets.

Our theoretical results emphasize the two different time scales at play in bacterial rheotaxis: bacteria are elongated objects following Jeffery orbits in the shear plane, while at the same time experiencing a drift perpendicular to the shear plane (rheotaxis) due to their chiral helix. The instantaneous dynamics of bacterial trajectories are dominated by Jeffery effects (section 11.2), so that the angles that describe the bacterial orientation in the shear plane, θ , and perpendicular to the shear plane, ψ , as well as the position z of the bacteria all oscillate with a period T given by Jeffery (see Eq. 11.15 and Fig. 11.5a) [19, 186]. Jeffery effects alone do not lead to a drift perpendicular to the shear plane, which is instead due to the chirality of the bacterial helix and occurs on a longer time scale.



FIGURE 11.9: (a) Snapshot of the microfluidic device (bottom-view). The squares are holes in the channel ceiling, that act as surface-energy anchors in which droplets are created and immobilized. (b) Cross-sectional schematic of an individual droplet immobilized in its anchors, containing bacteria and surrounded by oil. The plane of observation is shown by the orange dashed line. (c) Left: Bacteria in a droplet in the absence of flow are randomly distributed. Middle: in the presence of an oil flow going from left to right ($u_{oil} = 2.6 \text{ mm/s}$), recirculation within the aqueous drops induces bacterial segregation in the top-half (y > 0) of the droplet. Right: in the presence of an oil flow going from right to left, bacteria segregate to the bottomhalf (y < 0) of the droplet ($u_{oil} = 2.6 \text{ mm/s}$). White arrows indicate the oil flow direction. (d) Probability density maps of bacterial presence, for an oil flow going from left to right ($u_{oil} = 2.6 \text{ mm/s}$), and from right to left ($u_{oil} = -2.6 \text{ mm/s}$).

The evolution of the angle θ in the shear plane is not affected by rheotactic effects, see Eq. (11.12). The dynamics of the pitch angle ψ however depends on the presence of the chiral helix, and the angle ψ slowly drifts, over a time scale much larger than the period of oscillations T, until reaching the fixed point $\psi = \pi/2$ (see Fig. 11.5b). The rheotactic drift in z, perpendicular to the shear plane, depends on the pitch angle through $\dot{z} \propto \cos(\psi)$, and stops for $\psi = \pi/2$ (see Fig. 11.6b). Once $\psi = \pi/2$ is reached, the drift observed experimentally in the z direction is purely due to intrinsic bacterial propulsion: bacteria are stably oriented perpendicular to the shear, and move along their axis of orientation.

The above picture is purely deterministic, and does not take into account the fact that bacteria are subject to rotational diffusion and exhibit a run-and-tumble motion that leads to a randomization of their trajectories [188,192], and counters the rheotactic effects. In a channel of finite size L perpendicular to the shear plane, the probability to find bacteria at a height z is found to depend on the applied shear rate through the Péclet number $Pe = \frac{L\dot{z}}{D_z}$, where D_z is an effective diffusion coefficient, see Eq. (11.22). For $Pe \ll 1$, bacteria are uniformly distributed along z, while for $Pe \gg 1$, all bacteria accumulate at the boundary z = L. The fraction of bacteria accumulating at the



FIGURE 11.10: (a) Asymmetric segregation of *B. subtilis* in a droplet. At the beginning of the experiment, bacteria swim freely in the droplet and there is no asymmetry. At t=10 s, an outer oil flow of mean speed $u_{\rm oil} = -2.6$ mm/s is applied, leading to bacterial segregation in the droplet. Black points: experimental data. Red line: fit to a sigmoid function. (b) The end-value of the asymmetry depends on the applied shear. Black points: experimental data. Red line: fit to a sigmoid function, enabling to extract an effective diffusion coefficient $D_z = 2.6 \ \mu m^2/s$.

boundary as a function of the applied shear rate was determined experimentally and fitted to the theoretical shape given by Eq. (11.27) with a single fit parameter D - z, see Fig. 11.10b. Experimental and theoretical results are in very good agreement and yield an effective diffusion coefficient of $D_z = 2.6 \ \mu m^2/s$.

While the work shown here focused on simple flagellar geometries and simple shear flows, the semi-analytical workflow can be readily extended to more complex problems which may not be amenable to fully analytical formulas. In particular, flows encountered by microorganisms in their natural environment are unsteady [193], introducing another characteristic time scale in the problem; the decoupling of time scales between Jeffery effects and rheotactic effects could then prove useful to simplify the analysis of rheotactic drift in these natural environments.

11.6 Materials and Methods

11.6.1 Bacteria

Experiments were performed with *Bacillus subtilis* strain GM2938, a kind gift from Dominique Le Coq (Micalis Institute, INRA, Université Paris-Sud). The day before an

experiment, cells from an LB-agar plate were resuspended in LB medium and grown overnight in an incubator set at 37°C, with a rotation rate of 180 rpm. On the morning of the experiment, 10 μ L of the overnight bacterial suspension was diluted into 2 mL LB medium (Sigma-Aldrich L3022, Merck, Germany), and bacteria were let grow in the incubator until they reached an optical density at 600 nm of OD₆₀₀ ≈ 0.2 (Ultrospec 100, Amersham Biosciences, UK). The bacterial suspension was then introduced in the microfluidic chip to form droplets as detailed below.

11.6.2 Microfluidics

The microfluidic device used was similar to the one described in [189], and consisted in a wide, 2D chamber $(0.5 \times 3 \text{ cm})$ connected to two inlets and one outlet. An array of square grooves, or anchors, was etched on the chamber ceiling. The chamber depth was 35 μ m, and the anchors had a side 150 μ m and height 100 μ m, see Fig. 11.9a-b for a sketch of the device. All devices were made out of PDMS and sealed on a glass microscopy slide by plasma bonding.

Before the beginning of experiments, microfluidic devices were rendered hydrophobic by surface treatment with Novec 1720 (3M), and prefilled with fluorinated oil (FC-40, 3M) containing 0.5% (w/w) surfactant (RAN Technologies). Droplets of bacterial suspension were created in two steps: first, an aqueous suspension of bacteria was flowed into the device, and then a gentle flow of oil was applied. This led to the breakup of droplets on the anchors, see [189,194] for details. A typical snapshot of a droplet formed by break-up is shown on Fig. 11.9c, where individual bacteria can be seen. As a result of the break-up process, droplets were immobilized in their respective anchors and surrounded by oil. Part IV Conclusion

Chapter 12

Conclusions and Future work

12.1 Immuno-oncology in droplets

Combining droplet microfluidics, tumor models, imaging pipelines and mathematical modeling enabled me to characterize T cell elimination of target spheroids. The results shed light on the collective effects (accumulation and killing) that underpin a successful immune response. We show that from the very first contacts, T cells signal to each other to guide other cells to the tumor, which increases the accumulation efficiency. From a technological standpoint, this work demonstrates that it is possible to build immuno-oncological models recapitulating complex *in vivo* behaviors in microfluidic droplets. We have developed the protocol to construct these models, add the T cells and monitor their interactions with the tumors. It has also led to the development of a set of image analysis tools that might prove to be useful in the future for any study implicating the tracking of motile cells inside droplets.

However, a lot of possible developments remain to be explored. In so far as the device is concerned, a major current barrier is the difficulty to selectively extract droplets from the chip for further downstream analysis (RNAseq, DNAseq etc.). It would be a huge leap forward to be able to connect the imaged spheroid fate within the droplet to other biologically relevant signals such as cytokine expression levels, metabolic information etc. In the shorter term, an interesting experiment would be to infer the droplet content in certain specific molecules of interest by adding antibody-covered beads to the droplets as has been used in Saint-Sardos et al. [79]. This would allow for the dynamic measurement of cytokine levels in the droplet as a function of time.

With this study, I showed that T cells communicate with each other upon contact with the spheroid, probably by secreting CCR5-binding cytokines [4], and that the signal is sensed by the T cells even at very low T cell numbers on the spheroid and so early on in the immune response. I also demonstrated the link between a successful early immune response and local T cell concentration.

In the future, it would be particularly interesting to monitor the initiation of the immune response *in-vivo* to compare the *in-vivo* and *in-vitro* accumulation dynamics and potentially identify the factors preventing T cells from migrating to the tumor site. It would also be exciting to combine the monitoring done in this project with the multiplexing capacity of the microfluidic device [12] to test the effect different cytokines and drugs at different concentrations of the efficacy of the immune response. Of par-

ticular interest would be the precise quantification of the effect of adding checkpoint inhibitors (such as anti-PD-L1 or anti-CTLA-4 antibodies) to the droplets to see if the T cell concentration threshold for successful spheroid fragmentation is lowered or not.

A significant limitation of the current tumor model is its relative simplicity. The immune reaction involves a large number of different cell types, which communicate with each other to modulate their responses. The simplistic model we have used here has the advantage of enabling us to validate the technology, but for clinical or pharmaceutical studies, it would be necessary to increase its complexity. This could be done in two different manners: either multiple cell types could be added to the droplets in accordance with the questions investigated. For instance, the literature describes the impact of tumor associated macrophages (TAMs) on the efficacy of the immune response [33]. Adding TAMs to the cell mix would enable investigators to model their influence *in-vitro*. Alternatively, one could make spheroids from cells derived from patient samples. This would make it possible to mimic the complexity of the tumor by including the whole patient cell diversity. On the other hand, cultivating patient derived samples is challenging and their procurement can be a significant barrier. If these difficulties can be overcome however, then the technology could be very promising, since patient derived organoids (PDOs) have already been proven to predict in-vivo responses to chemotherapeutic drugs [7].

In order to bridge these gaps and implement more advanced tumor models, Valentin Bonnet has started working for his PhD in the lab. Hopefully, he will be able to collaborate with a team of researchers from the Gustave Roussy Institute, where they will develop protocols for cultivate tumor samples from patients on the same microfluidic chip as the one described in this thesis.

12.2 Building a network representation of tissues

After developing new experimental models to test tumor rejection, I worked on a new theoretical framework to improve the understanding of biological tissues. Deploying graph-based analysis methods could greatly enrich the amount of information extracted from biological data. These methods have already started to appear to analyze spatial transcriptomics data [195] or data from multiplexed imaging [16]. I have written a program that could potential unlock the potential of spatially contextualized data analyses from any segmented image by creating a "network twin" of the original sample. A key advantage of the graph representation lies in its capacity to connect the scales from the single-cell all the way up to the whole tissue: instead of focusing solely on the immediate environment of cells, we can connect cell behavior with tissue-scale cell features. Designed for ease of use, the program can run from segmented images (be they 2D or 3D) and from a CSV containing the positions of input cells. It rapidly generates a graph object can analyzed using a wide range of available methods.

Initially a library written to analyze spheroid images generated in the lab, this work has attracted some interest from colleagues at the Pasteur Institute with some very interesting use-cases. For instance, Nicolas Dray from the Zebrafish Neurogenetics unit, wants to use the code to map neural stem cell activity and link behaviors to their local environment. I expect to get more feedback on the library once our work will have been presented outside the lab. This will probably guide updates to the library and inspire future developments.

12.3 A structural description of the fetal liver

We applied Griottes on fetal liver tissue sections to understand the organization of the organ during development. Thanks to the network framework, I was able to identify the structure of the tissue around portal vessels and central veins, revealing an important accumulation of GR1+/CD45+ cells at later time stages. The spatial quantification of liver tissues complements population-based approaches (cytokine production by RNAseq, cellular phenotype by FACS etc.) and sheds light on the development of this crucial organ.

This work demonstrates the potential that lies in combining innovative imaging approaches with new image and data analysis techniques. We were able to measure previously unattainable features that highlight new biological signals and can orient future understanding of liver development and hematopoietic stem cell (HSC) ontology. In the future, we hope to analyze the distribution of CD45+/Kit+ progenitor cells in the fetal liver at early stages. This population contains the HSCs and would therefore give us some clues on the factors driving HSC population expansion.

Other development possibilities are in the downstream analysis of the "network twin" of the tissue. One particularly intriguing idea would be to use graph neural networks to analyze the tissues. This has already been done successfully to study to explain part of the variance in gene expression by linking it to the immediate cell neighborhood [196]. In a similar vein, we could train a neural network to recognize niches containing HSCs (or any other cell type of interest). If the algorithm manages to differentiate HSC-containing environments to HSC-free ones, this means that a signal exists and that a feature of the network is statistically linked to HSC presence. In the future, this technique could be used to identify cell niches that aren't visible with the naked eye.

12.4 Rheotactic motion of bacteria in shear flow

Finally, I have also studied the motility of bacteria in microfluidic droplets. This led to the development of a semi-analytical framework to calculate the trajectories of bacteria in a shear flow, which enabled me to compute the probability of bacteria experiencing rheotaxis as a function of the applied shear rate. Theoretical results on the fraction of bacteria displaying a rheotactic drift were favorably compared to experimental data of bacterial rheotaxis in microfluidic droplets. The theoretical results emphasize the two different time scales at play in bacterial rheotaxis: bacteria are elongated objects following Jeffery orbits in the shear plane, while at the same time experiencing a drift perpendicular to the shear plane (rheotaxis) due to their chiral helix. It also highlighted the role of diffusion in reorienting the bacteria subject to rheotactic flow, which would otherwise stay oriented orthogonally to the shear plane.

While the work shown here focused on simple flagellar geometries and simple shear flows, the semi-analytical workflow can be readily extended to more complex problems which may not be amenable to fully analytical formulas. In particular, flows encountered by microorganisms in their natural environment are unsteady [193], introducing another characteristic time scale in the problem; the decoupling of time scales between Jeffery effects and rheotactic effects could then prove useful to simplify the analysis of rheotactic drift in these natural environments.

Bibliography

- P. W. Anderson. More is Different. Science (New York, N.Y.), 177(4047):393– 396, 1972.
- [2] Sadna Budhu, John D. Loike, Ashley Pandolfi, Soo Han, Geoffrey Catalano, Andrei Constantinescu, Raphael Clynes, and Samuel C. Silverstein. CD8+ T cell concentration determines their efficiency in killing cognate antigen-expressing syngeneic mammalian cells in vitro and in mouse tissues. *Journal of Experimental Medicine*, 207(1):223–235, 2010.
- [3] Ronan Thibaut, Pierre Bost, Idan Milo, Marine Cazaux, Fabrice Lemaître, Zacarias Garcia, Ido Amit, Béatrice Breart, Clémence Cornuot, Benno Schwikowski, and Philippe Bousso. Bystander IFN-γ activity promotes widespread and sustained cytokine signaling altering the tumor microenvironment. Nature Cancer, 1(3):302–314, 2020.
- [4] Jorge Luis Galeano Niño, Sophie V Pageon, Szun S Tay, Feyza Colakoglu, Daryan Kempe, Jack Hywood, Jessica K Mazalo, James Cremasco, Matt A Govendir, Laura F Dagley, Kenneth Hsu, Simone Rizzetto, Jerzy Zieba, Gregory Rice, Victoria Prior, Geraldine M O'Neill, Richard J Williams, David R Nisbet, Belinda Kramer, Andrew I Webb, Fabio Luciani, Mark N Read, and Maté Biro. Cytotoxic T cells swarm by homotypic chemokine signalling. *eLife*, 9, 10 2020.
- [5] Giulia Adriani, Andrea Pavesi, Anthony T. Tan, Antonio Bertoletti, Jean Paul Thiery, and Roger D. Kamm. Microfluidic models for adoptive cell-mediated cancer immunotherapies. *Drug Discovery Today*, 21(9):1472–1478, 2016.
- [6] Alexandra Boussommier-Calleja, Ran Li, Michelle B. Chen, Siew Cheng Wong, and Roger D. Kamm. Microfluidics: A New Tool for Modeling Cancer–Immune Interactions. *Trends in Cancer*, 2(1):6–19, 2016.
- [7] Russell W. Jenkins, Amir R. Aref, Patrick H. Lizotte, Elena Ivanova, Susanna Stinson, Chensheng W. Zhou, Michaela Bowden, Jiehui Deng, Hongye Liu, Diana Miao, Meng Xiao He, William Walker, Gao Zhang, Tian Tian, Chaoran Cheng, Zhi Wei, Sangeetha Palakurthi, Mark Bittinger, Hans Vitzthum, Jong Wook Kim, Ashley Merlino, Max Quinn, Chandrasekar Venkataramani, Joshua A. Kaplan, Andrew Portell, Prafulla C. Gokhale, Bart Phillips, Alicia Smart, Asaf Rotem, Robert E. Jones, Lauren Keogh, Maria Anguiano, Lance Stapleton, Zhiheng Jia, Michal Barzily-Rokni, Israel Cañadas, Tran C. Thai, Marc R. Hammond, Raven Vlahos, Eric S. Wang, Hua Zhang, Shuai Li, Glenn J. Hanna, Wei

Huang, Mai P. Hoang, Adriano Piris, Jean Pierre Eliane, Anat O. Stemmer-Rachamimov, Lisa Cameron, Mei Ju Su, Parin Shah, Benjamin Izar, Manisha Thakuria, Nicole R. LeBoeuf, Guilherme Rabinowits, Viswanath Gunda, Sareh Parangi, James M. Cleary, Brian C. Miller, Shunsuke Kitajima, Rohit Thummalapalli, Benchun Miao, Thanh U. Barbie, Vivek Sivathanu, Joshua Wong, William G. Richards, Raphael Bueno, Charles H. Yoon, Juan Miret, Meenhard Herlyn, Levi A. Garraway, Eliezer M. Van Allen, Gordon J. Freeman, Paul T. Kirschmeier, Jochen H. Lorch, Patrick A. Ott, F. Stephen Hodi, Keith T. Flaherty, Roger D. Kamm, Genevieve M. Boland, Kwok Kin Wong, David Dornan, Cloud Peter Paweletz, and David A. Barbie. Ex vivo profiling of PD-1 blockade using organotypic tumor spheroids. *Cancer Discovery*, 8(2):196–215, 2018.

- [8] Amir R. Aref, Marco Campisi, Elena Ivanova, Andrew Portell, Dalia Larios, Brandon P. Piel, Natasha Mathur, Chensheng Zhou, Raven Vlahos Coakley, Alan Bartels, Michaela Bowden, Zach Herbert, Sarah Hill, Sean Gilhooley, Jacob Carter, Israel Cañadas, Tran C. Thai, Shunsuke Kitajima, Valeria Chiono, Cloud P. Paweletz, David A. Barbie, Roger D. Kamm, and Russell W. Jenkins. 3D microfluidic: Ex vivo culture of organotypic tumor spheroids to model immune checkpoint blockade. *Lab on a Chip*, 18(20):3129–3143, 2018.
- [9] Sébastien Sart, Raphaël F.-X. Tomasi, Gabriel Amselem, and Charles N. Baroud. Multiscale cytometry and regulation of 3D cell cultures on a chip. *Nature communications*, 8(1):469, September 2017.
- [10] Bettina Weigelin, Annemieke Th. den Boer, Esther Wagena, Kelly Broen, Harry Dolstra, Rob J. de Boer, Carl G. Figdor, Johannes Textor, and Peter Friedl. Cytotoxic T cells are able to efficiently eliminate cancer cells by additive cytotoxicity. *Nature Communications*, 12(1):1–12, 2021.
- [11] Stephan Halle, Kirsten Anja Keyser, Felix Rolf Stahl, Andreas Busche, Anja Marquardt, Xiang Zheng, Melanie Galla, Vigo Heissmeyer, Katrin Heller, Jasmin Boelter, Karen Wagner, Yvonne Bischoff, Rieke Martens, Asolina Braun, Kathrin Werth, Alexey Uvarovskii, Harald Kempf, Michael Meyer-Hermann, Ramon Arens, Melanie Kremer, Gerd Sutter, Martin Messerle, and Reinhold Förster. In Vivo Killing Capacity of Cytotoxic T Cells Is Limited and Involves Dynamic Interactions and T Cell Cooperativity. *Immunity*, 44(2):233–245, 2016.
- [12] Raphaël F.X. Tomasi, Sébastien Sart, Tiphaine Champetier, and Charles N. Baroud. Individual Control and Quantification of 3D Spheroids in a High-Density Microfluidic Droplet Array. *Cell Reports*, 31(8), 2020.
- [13] Sébastien Sart, Raphaël F.X. Tomasi, Antoine Barizien, Gabriel Amselem, Ana Cumano, and Charles N. Baroud. Mapping the structure and biological functions within mesenchymal bodies using microfluidics. *Science Advances*, 6(10), 2020.
- [14] Carsen Stringer, Michalis Michaelos, and Marius Pachitariu. Cellpose : a generalist algorithm for cellular segmentation. pages 1–17, 2020.
- [15] Vivien Marx. Method of the Year: spatially resolved transcriptomics. Nature Methods, 18(1):9–14, 2021.

- [16] Yury Goltsev, Nikolay Samusik, Julia Kennedy-Darling, Salil Bhate, Matthew Hale, Gustavo Vazquez, Sarah Black, and Garry P. Nolan. Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell*, 174(4):968– 981, 2018.
- [17] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1-3):1–101, 2011.
- [18] E M Purcell. Life at Low Reynolds Number. 45(June 1976):3–11, 1977.
- [19] G. B. Jeffery and Louis Napoleon George Filon. The Motion of Ellipsoidal Particles Immersed in shear flow. Proc. R. Soc. A, 102(715), 1922.
- [20] Marcos, H. C. Fu, T. R. Powers, and R. Stocker. Bacterial rheotaxis. Proceedings of the National Academy of Sciences, 109(13):4780–4785, 2012.
- [21] Paul Ehrlich. Ueber den jetzigen Stand der Karzinomforschung. 1908.
- [22] Macfarlane Burnet. Cancer—a biological approach: Iii. viruses associated with neoplastic conditions. iv. practical applications. British medical journal, 1(5023):841, 1957.
- [23] Scott McComb, Aude Thiriot, Lakshmi Krishnan, and Felicity Stark. Introduction to the Immune System. In *The Variation and Adaptive Expression of Antibodies*, volume 1061, pages 1–20. 2013.
- [24] Nu Zhang and Michael J. Bevan. CD8+ T Cells: Foot Soldiers of the Immune System. *Immunity*, 35(2):161–168, 2011.
- [25] Johannes B. Huppa and Mark M. Davis. T-cell-antigen recognition and the immunological synapse. *Nature Reviews Immunology*, 3(12):973–983, 2003.
- [26] Caroline Robert. A decade of immune-checkpoint inhibitors in cancer therapy. Nature Communications, 11(1):10–12, 2020.
- [27] Michael L. Dustin. Modular design of immunological synapses and kinapses. Cold Spring Harbor perspectives in biology, 1(1), 2009.
- [28] Mohamed Trebak and Jean Pierre Kinet. Calcium signalling in T cells. Nature Reviews Immunology, 19(3):154–169, 2019.
- [29] Marine Cazaux, Capucine L. Grandjean, Fabrice Lemaître, Zacarias Garcia, Richard J. Beck, Idan Milo, Jérémy Postat, Joost B. Beltman, Eleanor J. Cheadle, and Philippe Bousso. Single-cell imaging of CAR T cell activity in vivo reveals extensive functional and anatomical heterogeneity. *Journal of Experimental Medicine*, 216(5):1038–1049, 2019.
- [30] Hélène Salmon, Katarzyna Franciszkiewicz, Diane Damotte, Pierre Validire, Alain Trautmann, Fathia Mami-chouaib, and Emmanuel Donnadieu. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. 122(3):899–910, 2012.
- [31] Roy Noy and Jeffrey W. Pollard. Tumor-Associated Macrophages: From Mechanisms to Therapy. *Immunity*, 41(1):49–61, 2014.

- [32] Alba Nicolas-boluda and Emmanuel Donnadieu. Comparative Immunology, Microbiology and Infectious Diseases Obstacles to T cell migration in the tumor microenvironment. *Comparative Immunology, Microbiology and Infectious Diseases*, 63(December 2018):22–30, 2019.
- [33] Elisa Peranzoni, Jean Lemoine, Lene Vimeux, Vincent Feuillet, Sarah Barrin, Chahrazade Kantari-Mimoun, Nadège Bercovici, Marion Guérin, Jérôme Biton, Hanane Ouakrim, Fabienne Régnier, Audrey Lupo, Marco Alifano, Diane Damotte, and Emmanuel Donnadieu. Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti-PD-1 treatment. Proceedings of the National Academy of Sciences of the United States of America, 115(17):E4041-E4050, 2018.
- [34] Idan Milo, Marie Bedora-Faure, Zacarias Garcia, Ronan Thibaut, Leïla Périé, Guy Shakhar, Ludovic Deriano, and Philippe Bousso. The immune system profoundly restricts intratumor genetic heterogeneity. *Science Immunology*, 3(29):eaat1435, 2018.
- [35] Daniel S. Chen and Ira Mellman. Oncology meets immunology: The cancerimmunity cycle. *Immunity*, 39(1):1–10, 2013.
- [36] Matthew F Krummel, Frederic Bartumeus, and Audrey Gérard. and mechanisms. Nature Publishing Group, 16(3):193–201, 2016.
- [37] Tajie H Harris, Edward J Banigan, David A Christian, Christoph Konradt, Elia D Tait Wojno, Kazumi Norose, Emma H Wilson, Beena John, Wolfgang Weninger, Andrew D Luster, Andrea J Liu, and Christopher A Hunter. Generalized Lévy walks and the role of chemokines in migration of effector CD8+ T cells. *Nature*, pages 3–7, 2012.
- [38] Edward J. Banigan, Tajie H. Harris, David A. Christian, Christopher A. Hunter, and Andrea J. Liu. Heterogeneous CD8+ T Cell Migration in the Lymph Node in the Absence of Inflammation Revealed by Quantitative Migration Analysis. *PLoS Computational Biology*, 11(2):1–20, 2015.
- [39] M M Van Buuren, J J Calis, T N Schumacher, M E Dudley, D C Roopenian, M Dupage, C Mazumdar, L M Schmidt, A F Cheung, T Jacks, M S Rooney, S A Shukla, C J Wu, G Getz, N Hacohen, T N Schumacher, C Kesmir, M M Van Buuren, D S Chen, I Mellman, E F Fritsch, N Hacohen, and C J Wu. Science-2015-Joyce-74-80. 348(6230), 2015.
- [40] Peter Friedl and Bettina Weigelin. Interstitial leukocyte migration and immune function. 9(9):960–969, 2008.
- [41] S. Ariotti, J. B. Beltman, R. Borsje, M. E. Hoekstra, W. P. Halford, J. B. A. G. Haanen, R. J. de Boer, and T. N. M. Schumacher. Subtle CXCR3-Dependent Chemotaxis of CTLs within Infected Tissue Allows Efficient Target Localization. *The Journal of Immunology*, 195(11):5285–5295, 2015.

- [42] Purnima Bhat, Graham Leggatt, Klaus I. Matthaei, and Ian H. Frazer. The kinematics of cytotoxic lymphocytes influence their ability to kill target cells. *PLoS ONE*, 9(5), 2014.
- [43] Joost B. Beltman, Athanasius F.M. Marée, and Rob J. De Boer. Analysing immune cell migration. *Nature Reviews Immunology*, 9(11):789–798, 2009.
- [44] Alexandra Sontheimer-Phelps, Bryan A. Hassell, and Donald E. Ingber. Modelling cancer in microfluidic human organs-on-chips. *Nature Reviews Cancer*, 19(2):65–81, February 2019.
- [45] Karen F. Chambers, Eman M.O. Mosaad, Pamela J. Russell, Judith A. Clements, and Michael R. Doran. 3D cultures of prostate cancer cells cultured in a novel high-throughput culture platform are more resistant to chemotherapeutics compared to cells cultured in monolayer. *PLoS ONE*, 9(11), 2014.
- [46] Wolf Herman Fridman, Franck Pagès, Catherine Sauts-Fridman, and Jrôme Galon. The immune contexture in human tumours: Impact on clinical outcome. *Nature Reviews Cancer*, 12(4):298–306, 2012.
- [47] Zilton Vasconcelos, Sabina Müller, Delphine Guipouy, Wong Yu, Claire Christophe, Sébastien Gadat, Salvatore Valitutti, and Loïc Dupré. Individual Human Cytotoxic T Lymphocytes Exhibit Intraclonal Heterogeneity during Sustained Killing. *Cell Reports*, 11(9):1474–1485, 2015.
- [48] Morgane Boulch, Capucine L. Grandjean, Marine Cazaux, and Philippe Bousso. Tumor Immunosurveillance and Immunotherapies: A Fresh Look from Intravital Imaging. *Trends in Immunology*, 40(11):1022–1034, 2019.
- [49] Béatrice Breart, Fabrice Lemaître, Susanna Celli, and Philippe Bousso. Twophoton imaging of intratumoral CD8+ T cell cytotoxic activity during adoptive T cell therapy in mice. *Journal of Clinical Investigation*, 118(4):1390–1397, 2008.
- [50] Francesco Marangoni, Thomas T. Murooka, Teresa Manzo, Edward Y. Kim, Esteban Carrizosa, Natalie M. Elpek, and Thorsten R. Mempel. The Transcription Factor NFAT Exhibits Signal Memory during Serial T Cell Interactions with Antigen-Presenting Cells. *Immunity*, 38(2):237–249, 2013.
- [51] Elena Agliari, Elena Biselli, Adele De Ninno, Giovanna Schiavoni, Lucia Gabriele, Anna Gerardino, Fabrizio Mattei, Adriano Barra, and Luca Businaro. Cancerdriven dynamics of immune cells in a microfluidic environment. *Scientific Reports*, 4:11–13, 2014.
- [52] Manu P. Kumar, Jinyan Du, Georgia Lagoudas, Yang Jiao, Andrew Sawyer, Daryl C. Drummond, Douglas A. Lauffenburger, and Andreas Raue. Analysis of Single-Cell RNA-Seq Identifies Cell-Cell Communication Associated with Tumor Characteristics. *Cell Reports*, 25(6):1458–1468, 2018.
- [53] Sarah E. Shelton, Huu Tuan Nguyen, David A. Barbie, and Roger D. Kamm. Engineering approaches for studying immune-tumor cell interactions and immunotherapy. *iScience*, 24(1):101985, January 2021.

- [54] Yotam E. Bar-Ephraim, Kai Kretzschmar, and Hans Clevers. Organoids in immunological research. *Nature Reviews Immunology*, 2019.
- [55] Theresa E Schnalzger, Marnix HP Groot, Congcong Zhang, Mohammed H Mosa, Birgitta E Michels, Jasmin Röder, Tahmineh Darvishi, Winfried S Wels, and Henner F Farin. 3D model for CAR -mediated cytotoxicity using patient-derived colorectal cancer organoids. *The EMBO Journal*, 38(12):1–15, 2019.
- [56] Dohyun Park, Kyungmin Son, Yunchan Hwang, Jihoon Ko, Younggyun Lee, Junsang Doh, and Noo Li Jeon. High-Throughput Microfluidic 3D Cytotoxicity Assay for Cancer Immunotherapy (CACI-IMPACT Platform). Frontiers in Immunology, 10(May):1–9, 2019.
- [57] Andrea Pavesi, Anthony T. Tan, Sarene Koh, Adeline Chia, Marta Colombo, Emanuele Antonecchia, Carlo Miccolis, Erica Ceccarello, Giulia Adriani, Manuela T. Raimondi, Roger D. Kamm, and Antonio Bertoletti. A 3D microfluidic model for preclinical evaluation of TCR-engineered T cells against solid tumors. JCI Insight, 2(12), 2017.
- [58] Marie Nguyen, Adele De Ninno, Arianna Mencattini, Fanny Mermet-Meillon, Giulia Fornabaio, Sophia S. Evans, Mélissande Cossutta, Yasmine Khira, Weijing Han, Philémon Sirven, Floriane Pelon, Davide Di Giuseppe, Francesca Romana Bertani, Annamaria Gerardino, Ayako Yamada, Stéphanie Descroix, Vassili Soumelis, Fatima Mechta-Grigoriou, Gérard Zalcman, Jacques Camonis, Eugenio Martinelli, Luca Businaro, and Maria Carla Parrini. Dissecting Effects of Anticancer Drugs and Cancer-Associated Fibroblasts by On-Chip Reconstitution of Immunocompetent Tumor Microenvironments. *Cell Reports*, 25(13):3884–3893, 2018.
- [59] Helen Song, Joshua D. Tice, and Rustem F. Ismagilov. A Microfluidic System for Controlling Reaction Networks in Time. Angewandte Chemie International Edition, 42(7):768–772, 2003. __eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.200390203.
- [60] Rémi Dangla, S Cagri Kayi, and Charles N Baroud. Droplet microfluidics driven by gradients of confinement. Proceedings of the National Academy of Sciences (PNAS), 110(3):853–858, January 2013.
- [61] André Lwoff and Antoinette Gutmann. Recherches sur un baccilus megatherium lysogène. Annales de l'institut Pasteur, 78:711–739, 1950.
- [62] A. Lwoff, R. Dulbecco, Marguerite Vogt, and Marguerite Lwoff. Kinetics of the Release of Poliomyelitis Virus from Single Cells. *Virology*, 1(1):128–39, 1955. Publisher: Baltimore.
- [63] G J NOSSAL. Antibody production by single cells. British journal of experimental pathology, 39(5):544–51, 1958. ISBN: 0007-1021.
- [64] B ROTMAN. Measurement of activity of single molecules of beta-Dgalactosidase. Proceedings of the National Academy of Sciences of the United States of America, 47:1981–1991, 1961. ISBN: 0027-8424.

- [65] David. Quere. Surfaces molles. page 94, 2013.
- [66] Bongseop Kwak, Yoohwan Lee, Jaehun Lee, Sungwon Lee, and Jiseok Lim. Mass fabrication of uniform sized 3D tumor spheroid using high-throughput microfluidic system. *Journal of Controlled Release*, 275(February):201–207, 2018.
- [67] Raphaël Tomasi. Multiscale cytometry of 3D cell cultures in microfluidic hydrogel arrays. PhD thesis, 2016. Thèse de doctorat dirigée par Baroud, Charles Biologie Université Paris-Saclay (ComUE) 2016.
- [68] Etienne Fradet. Monitoring biochemical reactions in stationary droplets. phdthesis, Ecole Polytechnique X, November 2013.
- [69] Jean-christophe Baret, Jean-christophe Baret Surfactants, Royal Society, and Jean-christophe Baret. Surfactants in droplet-based microfluidics. 12(3):422– 433, 2012.
- [70] Jean Christophe Baret, Felix Kleinschmidt, Abdeslam El Harrak, and Andrew D. Griffiths. Kinetic aspects of emulsion stabilization by surfactants: A microfluidic analysis. *Langmuir*, 25(11):6088–6093, 2009.
- [71] Rémi Dangla, Sungyon Lee, and Charles N. Baroud. Trapping microfluidic drops in wells of surface energy. *Physical Review Letters*, 107(12), 2011.
- [72] Paul Abbyad, Rémi Dangla, Antigoni Alexandrou, and Charles N Baroud. Rails and anchors: guiding and trapping droplet microreactors in two dimensions. *Lab Chip*, 11(5):813–821, 2011. ISBN: 1473-0197.
- [73] Sigrid A. Langhans. Three-dimensional in vitro cell culture models in drug discovery and drug repositioning. *Frontiers in Pharmacology*, 9(JAN):1–14, 2018.
- [74] Nupura S. Bhise, João Ribas, Vijayan Manoharan, Yu Shrike Zhang, Alessandro Polini, Solange Massa, Mehmet R. Dokmeci, and Ali Khademhosseini. Organon-a-chip platforms for studying drug delivery systems. *Journal of Controlled Release*, 190:82–93, 2014.
- [75] Linfen Yu, Michael C. W. Chen, and Karen C. Cheung. Droplet-based microfluidic system for multicellular tumor spheroid formation and anticancer drug testing. *Lab on a chip*, 10(18):2424–2432, September 2010. Place: England.
- [76] Sébastien Sart, Raphaël F.X. Tomasi, Antoine Barizien, Gabriel Amselem, Ana Cumano, and Charles N. Baroud. Mapping Structure and Biological Functions within Mesenchymal Bodies using Microfluidics. *bioRxiv*, 2019.
- [77] S Sarkar, N Cohen, P Sabhachandani, and T Konry. Phenotypic drug profiling in droplet microfluidics for better targeting of drug-resistant tumors. *Lab on a Chip*, page 10, 2015.
- [78] Bongseop Kwak, Yoohwan Lee, Jaehun Lee, Sungwon Lee, and Jiseok Lim. Mass fabrication of uniform sized 3D tumor spheroid using high-throughput microfluidic system. Journal of controlled release : official journal of the Controlled Release Society, 275:201–207, April 2018. Place: Netherlands.

- [79] Adrien Saint-Sardos, Sébastien Sart, Kevin Lippera, Elodie Brient-Litzler, Sébastien Michelin, Gabriel Amselem, and Charles N. Baroud. High-Throughput Measurements of Intra-Cellular and Secreted Cytokine from Single Spheroids Using Anchored Microfluidic Droplets. *Small (Weinheim an der Bergstrasse, Germany)*, 16(49):e2002303, December 2020. Place: Germany.
- [80] Patrick C. Sachs, Peter A. Mollica, and Robert D. Bruno. Tissue specific microenvironments: a key tool for tissue engineering and regenerative medicine. *Journal of Biological Engineering*, 11:34, November 2017.
- [81] Davide Vigetti, Martin Götte, Mauro S. G. Pavão, and Achilleas D. Theocharis. Cellular Microenvironment in Human Pathologies. *BioMed Research International*, 2013:e946958, November 2013. Publisher: Hindawi.
- [82] Adrian C. Shieh. Biomechanical Forces Shape the Tumor Microenvironment. Annals of Biomedical Engineering, 39(5):1379–1389, May 2011.
- [83] Xingchen Li and Jianliu Wang. Mechanical tumor microenvironment and transduction: cytoskeleton mediates cancer cell invasion and metastasis. *International Journal of Biological Sciences*, 16(12):2014–2028, April 2020.
- [84] Qiuping Liu, Qing Luo, Yang Ju, and Guanbin Song. Role of the mechanical microenvironment in cancer development and progression. *Cancer Biology & Medicine*, 17(2):282–292, May 2020.
- [85] Tatsuyuki Yoshii, Yingying Geng, Shelly Peyton, Arthur M. Mercurio, and Vincent M. Rotello. Biochemical and biomechanical drivers of cancer cell metastasis, drug response and nanomedicine. *Drug Discovery Today*, 21(9):1489–1494, September 2016.
- [86] Ming-Zhu Jin and Wei-Lin Jin. The updated landscape of tumor microenvironment and drug repurposing. Signal Transduction and Targeted Therapy, 5(1):166, December 2020.
- [87] Yao Yuan, Yu-Chen Jiang, Chong-Kui Sun, and Qian-Ming Chen. Role of the tumor microenvironment in tumor progression and the clinical applications (Review). Oncology Reports, 35(5):2499–2515, May 2016.
- [88] Xianjie Jiang, Jie Wang, Xiangying Deng, Fang Xiong, Shanshan Zhang, Zhaojian Gong, Xiayu Li, Ke Cao, Hao Deng, Yi He, Qianjin Liao, Bo Xiang, Ming Zhou, Can Guo, Zhaoyang Zeng, Guiyuan Li, Xiaoling Li, and Wei Xiong. The role of microenvironment in tumor angiogenesis. *Journal of Experimental & Clinical Cancer Research*, 39(1):204, December 2020.
- [89] Maonan Wang, Jingzhou Zhao, Lishen Zhang, Fang Wei, Yu Lian, Yingfeng Wu, Zhaojian Gong, Shanshan Zhang, Jianda Zhou, Ke Cao, Xiayu Li, Wei Xiong, Guiyuan Li, Zhaoyang Zeng, and Can Guo. Role of tumor microenvironment in tumorigenesis. *Journal of Cancer*, 8(5):761–773, 2017.
- [90] Yaolei Wang and Jinyi Wang. Mixed hydrogel bead-based tumor spheroid formation and anticancer drug testing. *The Analyst*, 139(10):2449–2458, May 2014. Place: England.

- [91] L. Yu, S. M. Grist, S. S. Nasseri, E. Cheng, Y.-C. E. Hwang, C. Ni, and K. C. Cheung. Core-shell hydrogel beads with extracellular matrix for tumor spheroid formation. *Biomicrofluidics*, 9(2):024118, March 2015.
- [92] Dongjin Lee and Chaenyung Cha. The Combined Effects of Co-Culture and Substrate Mechanics on 3D Tumor Spheroid Formation within Microgels Prepared via Flow-Focusing Microfluidic Fabrication. *Pharmaceutics*, 10(4), November 2018.
- [93] Dongjin Lee and Chaenyung Cha. Cell subtype-dependent formation of breast tumor spheroids and their variable responses to chemotherapeutics within microfluidics-generated 3D microgels with tunable mechanics. *Materials Science* and Engineering: C, 112:110932, July 2020.
- [94] Sabine Stoll, Jérôme Delon, Tilmann M. Brotz, and Ronald N. Germain. Dynamic imaging of T cell-dendritic cell interactions in lymph nodes. *Science*, 296(5574):1873–1876, 2002.
- [95] Giulia Pasqual, Aleksey Chudnovskiy, Jeroen M.J. Tas, Marianna Agudelo, Lawrence D. Schweitzer, Ang Cui, Nir Hacohen, and Gabriel D. Victora. Monitoring T cell-dendritic cell interactions in vivo by intercellular enzymatic labelling. *Nature*, 553(7689):496–500, 2018.
- [96] Tania Konry, Alexander Golberg, and Martin Yarmush. Live single cell functional phenotyping in droplet nano-liter reactors. *Scientific Reports*, 3(1):3179, December 2013.
- [97] Silvia Antona, Ilia Platzman, and Joachim P. Spatz. Droplet-Based Cytotoxicity Assay: Implementation of Time-Efficient Screening of Antitumor Activity of Natural Killer Cells. ACS Omega, 5(38):24674–24683, 2020.
- [98] Saheli Sarkar, Seamus McKenney, Pooja Sabhachandani, James Adler, Xiaozhe Hu, Dina Stroopinksy, Jacalyn Rosenblatt, David Avigan, and Tania Konry. Anti-myeloma activity and molecular logic operation by Natural Killer cells in microfluidic droplets. *Sensors and Actuators B: Chemical*, 282:580–589, March 2019.
- [99] Saheli Sarkar, Pooja Sabhachandani, Dashnamoorthy Ravi, Sayalee Potdar, Sneha Purvey, Afshin Beheshti, Andrew M. Evens, and Tania Konry. Dynamic Analysis of Human Natural Killer Cell Response at Single-Cell Resolution in B-Cell Non-Hodgkin Lymphoma. *Frontiers in Immunology*, 8:1736, December 2017.
- [100] Silvia Antona, Tobias Abele, Kevin Jahnke, Yannik Dreher, Kerstin Göpfrich, Ilia Platzman, and Joachim P. Spatz. Droplet-Based Combinatorial Assay for Cell Cytotoxicity and Cytokine Release Evaluation. Advanced Functional Materials, 30(46):1–9, 2020.
- [101] Aude I. Segaliny, Guideng Li, Lingshun Kong, Ci Ren, Xiaoming Chen, Jessica K. Wang, David Baltimore, Guikai Wu, and Weian Zhao. Functional TCR T cell

screening using single-cell droplet microfluidics. Lab on a Chip, 18(24):3733–3749, 2018.

- [102] Silvia Antona, Tobias Abele, Kevin Jahnke, Yannik Dreher, Kerstin Göpfrich, Ilia Platzman, and Joachim P. Spatz. Droplet-Based Combinatorial Assay for Cell Cytotoxicity and Cytokine Release Evaluation. Advanced Functional Materials, 30(46):2003479, November 2020.
- [103] Yuan Yuan, Julie Brouchon, J. Mauricio Calvo-Calle, Jing Xia, Li Sun, Xu Zhang, Kiera L. Clayton, Fangfu Ye, David A. Weitz, and John A. Heyman. Droplet encapsulation improves accuracy of immune cell cytokine capture assays. *Lab on* a Chip, 20(8):1513–1520, 2020.
- [104] Saheli Sarkar. T Cell Dynamic Activation and Functional Analysis in Nanoliter Droplet Microarray. Journal of Clinical & Cellular Immunology, 06(03), 2015.
- [105] Annabelle Gérard, Adam Woolfe, Guillaume Mottet, Marcel Reichen, Carlos Castrillon, Vera Menrath, Sami Ellouze, Adeline Poitou, Raphaël Doineau, Luis Briseno-Roa, Pablo Canales-Herrerias, Pascaline Mary, Gregory Rose, Charina Ortega, Matthieu Delincé, Sosthene Essono, Bin Jia, Bruno Iannascoli, Odile Richard-Le Goff, Roshan Kumar, Samantha N. Stewart, Yannick Pousse, Bingqing Shen, Kevin Grosselin, Baptiste Saudemont, Antoine Sautel-Caillé, Alexei Godina, Scott McNamara, Klaus Eyer, Gaël A. Millot, Jean Baudry, Patrick England, Clément Nizak, Allan Jensen, Andrew D. Griffiths, Pierre Bruhns, and Colin Brenan. High-throughput single-cell activity-based screening and sequencing of antibodies using droplet microfluidics. *Nature Biotechnology*, 38(6):715–721, June 2020.
- [106] Gustave Ronteix, Shreyansh Jain, Christelle Angely, Marine Cazaux, Roxana Khazen, Philippe Bousso, and Charles N. Baroud. A Multiscale Immuno-Oncology on-Chip System (MIOCS) establishes that collective T cell behaviors govern tumor regression. preprint, Biophysics, March 2021.
- [107] Yoav Manaster, Zohar Shipony, Anat Hutzler, Masha Kolesnikov, Camila Avivi, Bruria Shalmon, Iris Barshack, Michal J. Besser, Tali Feferman, and Guy Shakhar. Reduced CTL motility and activity in avascular tumor areas. *Cancer Immunology, Immunotherapy*, 68(8):1287–1301, 2019.
- [108] Andrea Imle, Peter Kumberger, Nikolas D. Schnellbächer, Jana Fehr, Paola Carrillo-Bustamante, Janez Ales, Philip Schmidt, Christian Ritter, William J. Godinez, Barbara Müller, Karl Rohr, Fred A. Hamprecht, Ulrich S. Schwarz, Frederik Graw, and Oliver T. Fackler. Experimental and computational analyses reveal that environmental restrictions shape HIV-1 spread in 3D cultures. *Nature Communications*, 10(1), 2019.
- [109] Stéfan Van Der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. Scikit-image: Image processing in python. *PeerJ*, 2014(1):1–18, 2014.
- [110] Wes McKinney. Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference, 1697900(Scipy):51-56, 2010.

- [111] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020.
- [112] Dan Allan, Casper van der Wel, Nathan Keim, Thomas A Caswell, Devin Wieker, Ruben Verweij, Chaz Reid, Thierry, Lars Grueter, Kieran Ramos, apiszcz, zoeith, Rebecca W Perry, François Boulogne, Prashant Sinha, pfigliozzi, Nicolas Bruot, Leonardo Uieda, Jan Katins, Hadrien Mary, and Aron Ahmadia. softmatter/trackpy: Trackpy v0.4.2. 10 2019.
- [113] G. Matthew Fricke, Kenneth A. Letendre, Melanie E. Moses, and Judy L. Cannon. Persistence and Adaptation in Immunity: T Cells Balance the Extent and Thoroughness of Search. *PLoS Computational Biology*, 12(3):1–23, 2016.
- [114] S. Condamin, O. Bénichou, and M. Moreau. Random walks and Brownian motion: A method of computation for first-passage times and related quantities in confined geometries. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 75(2):1–20, 2007.
- [115] Silvia Ariotti, Joost B Beltman, Rianne Borsje, E Mirjam, William P Halford, John B A G Haanen, and J Rob. Subtle CXCR3-Dependent Chemotaxis of CTLs within Infected Tissue Allows Efficient Target Localization. 2019.
- [116] Ian A. Cockburn, Rogerio Amino, Reka K. Kelemen, Scot C. Kuo, Sze Wah Tse, Andrea Radtke, Laura Mac-Daniel, Vitaly V. Ganusov, Fidel Zavala, and Robert Ménard. In vivo imaging of CD8+ T cell-mediated elimination of malaria liver stages. Proceedings of the National Academy of Sciences of the United States of America, 110(22):9090–9095, 2013.
- [117] Matteo Scabini, Fabio Stellari, Paolo Cappella, Sara Rizzitano, Gemma Texido, and Enrico Pesenti. In vivo imaging of early stage apoptosis by measuring realtime caspase-3/7 activation. Apoptosis, 16(2):198–207, 2011.
- [118] Peter McCullagh. Generalized linear models.
- [119] Skipper Seabold and Josef Perktold. Statsmodels: Econometric and Statistical Modeling with Python. PROC. OF THE 9th PYTHON IN SCIENCE CONF, (January 2010):57, 2010.
- [120] Raphaël F.-X. Tomasi, Sébastien Sart, Tiphaine Champetier, and Charles N. Baroud. Individual Control and Quantification of 3D Spheroids in a High-Density Microfluidic Droplet Array. *Cell reports*, 31(8):107670, May 2020.
- [121] Mirjam E. Hoekstra, Laura Bornes, Feline E. Dijkgraaf, Daisy Philips, Iris N. Pardieck, Mireille Toebes, Daniela S. Thommen, Jacco van Rheenen, and Ton N.M. Schumacher. Long-distance modulation of bystander tumor cells by CD8+ Tcell-secreted IFN-γ. Nature Cancer, 1(3):291–301, 2020.

- [122] David Tuveson and Hans Clevers. Cancer modeling meets human organoid technology. Science, 364(6444):952–955, 2019.
- [123] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Diameter of the World-Wide Web. Nature, 401(6749):130–131, 9 1999.
- [124] Fredrik Liljeros, Christofer R. Edling, Luís A. Nunes Amaral, H. Eugene Stanley, and Yvonne Åberg. The web of human sexual contacts. *Nature*, 411(6840):907– 908, 6 2001.
- [125] Andreas Wagner. The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes. *Molecular Biology and Evolution*, 18(7):1283–1292, 2001.
- [126] Haretsugu Hishigaki, Kenta Nakai, Toshihide Ono, Akira Tanigami, and Toshihisa Takagi. Assessment of prediction accuracy of protein function from proteinprotein interaction data. *Yeast*, 18(6):523–531, 2001.
- [127] Clément Roussel, Simona Cocco, and Rémi Monasson. Barriers and dynamical paths in alternating gibbs sampling of restricted boltzmann machines. *Physical Review E*, 104(3):034109, 2021.
- [128] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 5(4):115–133, 12 1943.
- [129] Réka Albert and Albert László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002.
- [130] Hippolyte Verdier, Maxime Duval, Francois Laurent, Alhassan Cassé, Christian L. Vestergaard, and Jean Baptiste Masson. Learning physical properties of anomalous random walks using graph neural networks. *Journal of Physics A: Mathematical and Theoretical*, 54(23):1–21, 2021.
- [131] O. Mason and M. Verwoerd. Graph theory and networks in biology. IET Systems Biology, 1(2):89–119, 2007.
- [132] U. Alon. Biological networks: The tinkerer as an engineer. Science, 301(5641):1866–1867, 2003.
- [133] Dennis Bray. Molecular Networks : The Top-Down View. 301(September):1864– 1866, 2003.
- [134] Stephen Eubank, Hasan Guclu, V. S.Anil Kumar, Madhav V. Marathe, Aravind Srinivasan, Zoltán Toroczkai, and Nan Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.
- [135] Vittoria Colizza, Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. The role of the airline transportation network in the. *Pnas*, 103(7):2015–2020, 2006.

- [136] M. Madan Babu, Nicholas M. Luscombe, L. Aravind, Mark Gerstein, and Sarah A. Teichmann. Structure and evolution of transcriptional regulatory networks. *Current Opinion in Structural Biology*, 14(3):283–291, 2004.
- [137] Jean Philippe Bouchaud and Marc Mezard. Complex systems. In Complex systems. 2013.
- [138] Mariela D. Petkova, Gašper Tkačik, William Bialek, Eric F. Wieschaus, and Thomas Gregor. Optimal Decoding of Cellular Identities in a Genetic Network. *Cell*, 176(4):844–855, 2019.
- [139] Daniel L. Coutu, Konstantinos D. Kokkaliaris, Leo Kunz, and Timm Schroeder. Three-dimensional map of nonhematopoietic bone and bone-marrow cells and molecules. *Nature Biotechnology*, 35(12):1202–1210, 2017.
- [140] Daniel L. Coutu, Konstantinos D. Kokkaliaris, Leo Kunz, and Timm Schroeder. Multicolor quantitative confocal imaging cytometry. *Nature Methods*, 15(1):39–46, 2018.
- [141] Christian M. Schürch, Salil S. Bhate, Graham L. Barlow, Darci J. Phillips, Luca Noti, Inti Zlobec, Pauline Chu, Sarah Black, Janos Demeter, David R. McIlwain, Nikolay Samusik, Yury Goltsev, and Garry P. Nolan. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell, 182(5):1341–1359, 2020.
- [142] Nicola Crosetto, Magda Bienko, and Alexander van Oudenaarden. Spatially resolved transcriptomics and beyond. *Nature Reviews Genetics*, 16(1):57–66, 1 2015.
- [143] Daniel Edsgärd, Per Johnsson, and Rickard Sandberg. Identification of spatial expression trends in single-cell gene expression data. *Nature Methods*, 15(5):339– 342, 2018.
- [144] Qian Zhu, Sheel Shah, Ruben Dries, Long Cai, and Guo Cheng Yuan. Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence in situ hybridization data. *Nature Biotechnology*, 36(12):1183– 1190, 2018.
- [145] Caleb R. Stoltzfus, Jakub Filipek, Benjamin H. Gern, Brandy E. Olin, Joseph M. Leal, Yajun Wu, Miranda R. Lyons-Cohen, Jessica Y. Huang, Clarissa L. Paz-Stoltzfus, Courtney R. Plumlee, Thomas Pöschinger, Kevin B. Urdahl, Mario Perro, and Michael Y. Gerner. CytoMAP: A Spatial Analysis Toolbox Reveals Features of Myeloid Cell Organization in Lymphoid Tissues. *Cell Reports*, 31(3), 2020.
- [146] Alexis Coullomb and Vera Pancaldi. Tysserand Fast reconstruction of spatial networks from bioimages. *bioRxiv*, pages 1–4, 2020.
- [147] Cristopher Moore and M. E. J. Newman. Epidemics and percolation in random networks. *Physical Review E*, 61(5):5678–5682, 2000.

- [148] Brendan F Miller, Dhananjay Bambah-mukku, Catherine Dulac, Xiaowei Zhuang, Jean Fan, and Cellular Biology. Characterizing spatial gene expression heterogeneity in spatially resolved single-cell transcriptomics data with nonuniform cellular densities.
- [149] Rory M. Power and Jan Huisken. A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nature Methods*, 14(4):360–373, 2017.
- [150] Caleb R Stoltzfus, Ramya Sivakumar, Leo Kunz, B E Olin Pope, Dario Speziale, Roberto Adelfio, Marina Bacac, Sara Colombetti, Mario Perro, and Michael Y Gerner. Multi-Parameter Quantitative Imaging of Tumor Microenvironments Reveals Perivascular Immune Niches Associated with Anti-Tumor Immunity Cocorresponding authors :. pages 1–39, 2021.
- [151] Matthijs J. D. Baars, Neeraj Sinha, Mojtaba Amini, Annelies Pieterman-Bos, Stephanie van Dam, Maroussia M. P. Ganpat, Miangela M. Laclé, Bas Oldenburg, and Yvonne Vercoulen. MATISSE: a method for improved single cell segmentation in imaging mass cytometry. *BMC Biology*, 19(1):99, 12 2021.
- [152] A A Hagberg, D A Schult, and P J Swart. Exploring network structure, dynamics, and function using NetworkX. 7th Python in Science Conference (SciPy 2008), (SciPy):11–15, 2008.
- [153] M. E. Dolega, M. Delarue, F. Ingremeau, J. Prost, A. Delon, and G. Cappello. Cell-like pressure sensors reveal increase of mechanical stress towards the core of multicellular spheroids under compression. *Nature Communications*, 8(May 2016), 2017.
- [154] France Rose, Luca Rappez, Sergio H. Triana, Theodore Alexandrov, and Auguste Genovesio. PySpacell: A Python Package for Spatial Analysis of Cell Images. *Cytometry Part A*, 97(3):288–295, 2020.
- [155] Sara Kaliman, Christina Jayachandran, Florian Rehfeldt, and Ana Suncana Smith. Limits of applicability of the voronoi tessellation determined by centers of cell nuclei to epithelium morphology. *Frontiers in Physiology*, 7(NOV), 2016.
- [156] Adam C. Wilkinson, Kyomi J. Igarashi, and Hiromitsu Nakauchi. Haematopoietic stem cell self-renewal in vivo and ex vivo. *Nature Reviews Genetics*, 21(9):541– 554, 2020.
- [157] Francisca Soares-da Silva, Márcia Peixoto, Ana Cumano, and Perpetua Pinto-do Ó. Crosstalk Between the Hepatic and Hematopoietic Systems During Embryonic Development. Frontiers in Cell and Developmental Biology, 8(July):1–20, 2020.
- [158] Samuel H. Cheshier, Sean J. Morrison, Xinsheng Liao, and Irving L. Weissman. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proceedings of the National Academy of Sciences of the United States* of America, 96(6):3120–3125, 1999.

- [159] Hideo Ema and Hiromitsu Nakauchi. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood*, 95(7):2284–2288, 2000.
- [160] Ana Cumano, Francoise Dieterian-Lievre, and Isabelle Godin. Lymphoid Potential, Probed before Circulation in Mouse, Is Restricted to Caudal Intraembryonic Splanchnopleura. *Cell*, 86:907–916, 1996.
- [161] R. Schofield. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells, 4(1-2):7–25, 1978.
- [162] Daisuke Sugiyama, Kasem Kulkeaw, and Chiyo Mizuochi. TGF-beta-1 upregulates extra-cellular matrix production in mouse hepatoblasts. *Mechanisms* of *Development*, 130(2-3):195–206, 2013.
- [163] Francisca Soares-da Silva, Laina Freyer, Ramy Elsaid, Odile Burlen-Defranoux, Lorea Iturri, Odile Sismeiro, Perpétua Pinto-do Ó, Elisa Gomez-Perdiguero, and Ana Cumano. Yolk sac, but not hematopoietic stem cell-derived progenitors, sustain erythropoiesis throughout murine embryonic life. Journal of Experimental Medicine, 218(4):e20201729, April 2021.
- [164] Owen J. Tamplin, Ellen M. Durand, Logan A. Carr, Sarah J. Childs, Elliott J. Hagedorn, Pulin Li, Amanda D. Yzaguirre, Nancy A. Speck, and Leonard I. Zon. Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell*, 160(1-2):241–252, 2015.
- [165] Jalal A. Khan, Avital Mendelson, Yuya Kunisaki, Alexander Birbrair, Yan Kou, Anna Arnal-Estapé, Sandra Pinho, Paul Ciero, Fumio Nakahara, Avi Ma'ayan, Aviv Bergman, Miriam Merad, and Paul S. Frenette. Fetal liver hematopoietic stem cell niches associate with portal vessels. *Science*, 351(6269):176–180, 2016.
- [166] Yeojin Lee, Juliana Leslie, Ying Yang, and Lei Ding. Hepatic stellate and endothelial cells maintain hematopoietic stem cells in the developing liver. The Journal of experimental medicine, 218(3), 2021.
- [167] Constantina Christodoulou, Joel A. Spencer, Shu Chi A. Yeh, Raphaël Turcotte, Konstantinos D. Kokkaliaris, Riccardo Panero, Azucena Ramos, Guoji Guo, Negar Seyedhassantehrani, Tatiana V. Esipova, Sergei A. Vinogradov, Sarah Rudzinskas, Yi Zhang, Archibald S. Perkins, Stuart H. Orkin, Raffaele A. Calogero, Timm Schroeder, Charles P. Lin, and Fernando D. Camargo. Liveanimal imaging of native haematopoietic stem and progenitor cells. *Nature*, 578(7794):278–283, 2020.
- [168] Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectationmaximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, 2001.
- [169] Sean R Eddy. What is a hidden Markov model? Statistical models called hidden Markov models are a recurring theme in computational biology. What are hidden Markov models, and why are they so useful for so many different problems? 5 I

E Start End Figure 1 A toy HMM for. *Nature Biotechnology*, 22(10):1315–1316, 2004.

- [170] H. Cho, K. Ayers, L. de Pills, Y. H. Kuo, J. Park, A. Radunskaya, and R. C. Rockne. Modelling acute myeloid leukaemia in a continuum of differentiation states. *Letters in Biomathematics*, 5(sup1):S69–S98, 2018.
- [171] Denis Schapiro, Hartland W. Jackson, Swetha Raghuraman, Jana R. Fischer, Vito R.T. Zanotelli, Daniel Schulz, Charlotte Giesen, Raúl Catena, Zsuzsanna Varga, and Bernd Bodenmiller. HistoCAT: Analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature Methods*, 14(9):873–876, 2017.
- [172] Jackline De Paula Ayres-Silva, Pedro Paulo De Abreu Manso, Mariana Rietmann Da Cunha Madeira, Marcelo Pelajo-Machado, and Henrique Leonel Lenzi. Sequential morphological characteristics of murine fetal liver hematopoietic microenvironment in Swiss Webster mice. *Cell and Tissue Research*, 344(3):455– 469, 2011.
- [173] Eric Lauga and Thomas R. Powers. The hydrodynamics of swimming microorganisms. *Reports on Progress in Physics*, 72(9), 2009.
- [174] Vasily Kantsler, Jörn Dunkel, Martyn Blayney, and Raymond E Goldstein. Rheotaxis facilitates upstream navigation of mammalian sperm cells. *eLife*, 3:1–12, 2014.
- [175] John O. Kessler. Hydrodynamic focusing of motile algal cells. Nature, 313(5999):218–220, jan 1985.
- [176] Xabel Garcia, Salima Rafaï, and Philippe Peyla. Light Control of the Flow of Phototactic Microswimmer Suspensions. *Physical Review Letters*, 110(13):138106, mar 2013.
- [177] Marcos, Henry C Fu, Thomas R Powers, and Roman Stocker. Separation of microscale chiral objects by shear flow. *Physical Review Letters*, 102(15):1–4, 2009.
- [178] Marcos, H. C. Fu, T. R. Powers, and R. Stocker. Bacterial rheotaxis. Proceedings of the National Academy of Sciences, 109(13):4780–4785, 2012.
- [179] Arnold J.T.M. Mathijssen, Nuris Figueroa-Morales, Gaspard Junot, Eric Clément, Anke Lindner, and Andreas Zöttl. Oscillatory surface rheotaxis of swimming E. coli bacteria. *Nature Communications*, 10(1):7–9, 2019.
- [180] Guangyin Jing, Andreas Zöttl, Éric Clément, and Anke Lindner. Chiralityinduced bacterial rheotaxis in bulk shear flows. *Science Advances*, 6(28), 2020.
- [181] Kenta Ishimoto. Helicoidal particles and swimmers in a flow at low Reynolds number. *Journal of Fluid Mechanics*, 2020.
- [182] Stephen Childress. Mechanics of Swimming and Flying. Cambridge Studies in Mathematical Biology. Cambridge University Press, 1981.

- [183] Hilding Faxen. Der Widerstand gegen die Bewegung einer starren Kugel in einer zähen Flüssigkeit, die zwischen zwei parallelen ebenen Wänden eingeschlossen ist.
- [184] Pierre Curie. Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique. J. Phys. Theor. Appl., 3(1), 1894.
- [185] G. B. Jeffery. The Motion of Ellipsoidal Particles Immersed in a Viscous Fluid. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 102(715):161–179, nov 1922.
- [186] L. G. Leal and E. J. Hinch. The effect of weak Brownian rotations on particles in shear flow. *Journal of Fluid Mechanics*, 46(4):685–703, 1971.
- [187] E. J. Hinch and L. G. Leal. Rotation of small non-axisymmetric particles in a simple shear flow. *Journal of Fluid Mechanics*, 92(3):591–607, 1979.
- [188] Jonathan Saragosti, Pascal Silberzan, and Axel Buguin. Modeling E. coli tumbles by rotational diffusion. implications for chemotaxis. *PLoS ONE*, 7(4), 2012.
- [189] Gabriel Amselem, Cyprien Guermonprez, Benoît Drogue, Sébastien Michelin, and Charles N. Baroud. Universal microfluidic platform for bioassays in anchored droplets. Lab on a Chip, 16(21):4200–4211, 2016.
- [190] Sungyon Lee, François Gallaire, and Charles N. Baroud. Interface-induced recirculation within a stationary microfluidic drop. *Soft Matter*, 8(41):10750, 2012.
- [191] Yue Ling, Jose Maria Fullana, Stéphane Popinet, and Christophe Josserand. Droplet migration in a Hele-Shaw cell: Effect of the lubrication film on the droplet dynamics. *Physics of Fluids*, 28(6), 2016.
- [192] Howard C. Berg. Random Walks in Biology. Princeton University Press, nov 1993.
- [193] Roman Stocker. Marine Microbes See a Sea of Gradients. Science, 338(6107):628– 633, 2012.
- [194] Gabriel Amselem, P. T. Brun, François Gallaire, and Charles N. Baroud. Breaking Anchored Droplets in a Microfluidic Hele-Shaw Cell. *Physical Review Applied*, 3(5):1–5, 2015.
- [195] Giovanni Palla, Hannah Spitzer, Michal Klein, David Fischer, Anna Christina Schaar, Louis Benedikt Kuemmerle, Sergei Rybakov, Ignacio L Ibarra, Olle Holmberg, Isaac Virshup, Mohammad Lotfollahi, Sabrina Richter, and Fabian J. Theis. Squidpy: a scalable framework for spatial single cell analysis. *bioRxiv*, page 2021.02.19.431994, 2021.
- [196] David S. Fischer, Anna C. Schaar, and Fabian J. Theis. Learning cell communication from spatial graphs of cells. *bioRxiv*, page 2021.07.11.451750, 2021.



Titre : Inférence des interactions cellule-cellule à partir de l'analyse quantitative d'images de microscopie

Mots clés : biologie quantitative, microfluidique, cancer, développement embryonnaire, imagerie.

Résumé : La régulation et le rejet éventuel des tumeurs par le système immunitaire est le résultat de multiples réseaux de régulation, influençant à la fois le comportement des cellules cancéreuses et immunitaires. Pour simuler ces effets complexes in-vitro, j'ai conçu une puce microfluidique permettant de confronter des sphéroïdes de mélanome à de multiples cellules T et d'observer les interactions qui en résultent avec une haute résolution spatio-temporelle et sur de longues périodes de temps. En utilisant de l'analyse d'images avancée combinée à des modèles mathématiques, je démontre qu'une boucle de rétroaction positive conduit l'accumulation de cellules T sur la tumeur, ayant pour conséquence une fragmentation accrue des sphéroïdes.

De nombreux processus biologiques sont le résultat d'interactions entre de multiples types de cellules, en particulier au cours du développement. Le foie fœtal est le lieu de la maturation et de l'expansion du système hématopoïétique. Pour l'étudier, j'ai développé des outils pour interpréter et quantifier de nouvelles images de l'organe, permettant la construction d'un "réseau jumeau" de chaque foie fœtal. Cette méthode permet de combiner les échelles unicellulaire et de l'organe dans une seule analyse, révélant l'accumulation de cellules myéloïdes autour des vaisseaux sanguins irriguant le foie fœtal aux derniers stades du développement de l'organe.

Title : Inferring cell-cell interactions from quantitative analysis of microscopy images

Keywords : quantitative biology, microfluidics, cancer, embryonic development, imaging

Abstract : The regulation and eventual rejection of tumours by the immune system is the result of multiple positive and negative regulation networks, influencing both the behaviour of the cancerous and immune cells. To mimic these complex effects in-vitro, I designed a microfluidic assay to challenge melanoma tumour spheroids with multiple T cells and observe the resulting interactions with high spatiotemporal resolution over long periods of time. Using advanced image analysis combined with mathematical modelling I demonstrate that a positive feedback loop drives T cell accumulation to the tumour site, leading to enhanced spheroid fragmentation. This study shows that it is possible to recapitulate complex antagonistic behaviours in-vitro, which paves the way for the elaboration of more sophisticated protocols, involving for example a more complex tumour micro-environment.

Many biological processes are the result of complex interactions between cell types, particularly so during development. The foetal liver is the locus of the maturation and expansion of the hematopoietic system, yet little is known about its structure and organisation. I developed tools to interpret and quantify imaging data from the fatal liver, enabling the construction of a "network twin" of each organ. This method makes it possible to combine the single-cell scale and the organ scale in the analysis, revealing the accumulation of myeloid cells around the blood vessels irrigating the foetal liver at the final stages of organ development. In the future, this technique will make it possible to analyse precisely the environmental niches of cell types of interest in a quantitative manner. This in turn could help us understand the developmental steps of crucial cell types such as hematopoietic stem cells.

