TESIS DE LA UNIVERSIDAD DE ZARAGOZA

2023 151

María Inés Godinho Gonçalves

Computational Modelling of Cancer Systems: From Individual to Collective Cell Behaviour

Director/es
García Aznar, José Manuel

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

COMPUTATIONAL MODELLING OF CANCER SYSTEMS: FROM INDIVIDUAL TO COLLECTIVE CELL BEHAVIOUR

Autor

María Inés Godinho Gonçalves

Director/es

García Aznar, José Manuel

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

Programa de Doctorado en Ingeniería Biomédica

2023

ESCUELA DE INGENIERÍA Y ARQUITECTURA

Universidad Zaragoza

Computational Modelling of Cancer Systems: From Individual to Collective Cell Behaviour

Inês G. Gonçalves

DOCTORAL PROGRAMME IN BIOMEDICAL ENGINEERING

Supervisor: José Manuel García-Aznar

March 2, 2023

Computational Modelling of Cancer Systems: From Individual to Collective Cell Behaviour

Inês G. Gonçalves

March 2, 2023

Abstract

Due to its complexity, cancer continues to be one of the leading causes of death worldwide. The creation of suitable preventative practices and innovative therapies is limited by the lack of understanding of the basic mechanisms that cause cancer. As such, new, more effective methods that advance our comprehension of cancer must be developed. Recent years have seen a rise in the use of computational modeling to explain biological processes that are expensive and challenging to explore in experimental settings. These methods enable the translation of biological mechanisms into mathematical equations and assumptions that can be evaluated using computer tools to produce new hypotheses. Moreover, computational frameworks are becoming more potent due to availability to high-throughput data and extensive processing capacity.

The global aim of this dissertation is to design and implement computational models of cancer, starting with simple and isolated behaviours and progressing towards more complex phenomena. Three specific research fields are addressed to achieve this general objective: (i) single-cell motility, (ii) tumour growth and (iii) pattern formation. In the first objective, we present a computational model to simulate individual cell motility that considers the mechanical and chemical properties of the microenvironment. Subsequently, we extend this framework to account for cell-cell interactions and reproduce the growth of multicellular tumour structures. Lastly, we integrate all the previously mentioned biological events and introduce cell differentiation as this framework's final building block to simulate the spatial pattern formation.

In addition, this dissertation discusses the relevance of integrating experimental data and computational methods to improve biological relevance and confirm model outputs. In particular, we show how calibration and optimization techniques can be used to consider empircal data in model design and validation. Qualitative and quantitative experimental results from both the literature and novel expriments are reproduced in this dissertation to showcase different approaches in data integration.

Overall, this dissertation provides a prime instance of how computer modelling can be used to analyze and understand complex problems in cancer biology. We explicitly show how model components may represent certain aspects of cancer biology that can be enhanced and reproduced using experimental data. Consequently, it is shown that complex behaviour like tumour growth and pattern formation result from the intricate interplay between model components.

Resumen

Debido a su complejidad, el cáncer sigue siendo una de las principales causas de muerte a nivel mundial. La creación de prácticas preventivas adecuadas y terapias innovadoras está limitada por la falta de comprensión de los mecanismos básicos que causan el cáncer. Como tal, se deben desarrollar métodos nuevos y más efectivos que avancen nuestra comprensión del cáncer. En los últimos años, se ha visto un aumento en el uso de modelos computacionales para explicar procesos biológicos que son costosos y difíciles de explorar en entornos experimentales. Estos métodos permiten la traducción de mecanismos biológicos en ecuaciones y suposiciones matemáticas que pueden evaluarse utilizando herramientas informáticas para producir nuevas hipótesis. Además, las tecnologías computacionales se están volviendo más potentes debido a la disponibilidad de datos y la amplia capacidad de procesamiento.

El objetivo global de esta tesis es diseñar e implementar modelos computacionales de cáncer, comenzando con comportamientos simples y aislados y progresando hacia fenómenos más complejos. Se abordan tres campos de investigación específicos para lograr este objetivo general: (i) motilidad unicelular, (ii) crecimiento tumoral y (iii) formación de patrones. En el primer objetivo, se presenta un modelo computacional para simular la motilidad celular individual que considera las propiedades mecánicas y químicas del microambiente. Posteriormente, este trabajo fue ampliado para tener en cuenta las interacciones célula-célula y reproducir el crecimiento de estructuras tumorales multicelulares. Por último, todos los eventos biológicos mencionados anteriormente fueron considerados y se añadió la diferenciación celular como el bloque de construcción final de esta tesis para simular la formación de patrones espaciales.

Además, esta tesis analiza la relevancia de integrar datos experimentales y métodos computacionales para mejorar la precisión biológica y confirmar los resultados del modelo. En particular, muestra cómo se pueden usar técnicas de calibración y optimización para considerar datos empíricos en el diseño y validación de modelos. Los resultados experimentales cualitativos y cuantitativos, tanto de la literatura como de nuevos experimentos, se reproducen en este artículo para mostrar diferentes enfoques en la integración de datos.

En general, esta tesis proporciona un modelo de cómo se pueden utilizar los métodos computacionales para analizar y comprender problemas complejos en la biología del cáncer. Demuestra explícitamente cómo los componentes del modelo pueden representar ciertos aspectos de la biología del cáncer, que pueden mejorarse y reproducirse utilizando datos experimentales. En consecuencia, los comportamientos complejos, como el crecimiento tumoral y la formación de patrones, resultan de la intrincada interacción entre los componentes del modelo.

À minha família.

"Mas ao menos fica da amargura do que nunca serei
A caligrafia rápida destes versos,
Pórtico partido para o Impossível."

Álvaro de Campos, "Tabacaria"

Contents

1 Introduction **1**
- 1.1 Context . . . 1
- 1.2 Motivation . . . 4
- 1.3 Objectives . . . 6
- 1.4 Document Structure . . . 8

2 Background **9**
- 2.1 Introduction . . . 9
- 2.2 Biological background . . . 10
  - 2.2.1 Cancer development . . . 10
  - 2.2.2 The hallmarks of cancer . . . 11
- 2.3 Computational oncology . . . 16
  - 2.3.1 Modelling cancer systems . . . 16
  - 2.3.2 Continuum vs discrete models . . . 17
  - 2.3.3 Centre-based models (CBMs) . . . 18
- 2.4 Summary . . . 23

3 Single-cell motility **25**
- 3.1 Introduction . . . 26
- 3.2 Model optimization and data integration: an overview . . . 29
  - 3.2.1 Microfluidic devices and migration assays . . . 30
  - 3.2.2 Model calibration routines and data-driven models . . . 31
- 3.3 Modelling 3D cell motility in collagen matrices of different densities . . . 32
  - 3.3.1 Study outline . . . 32
  - 3.3.2 Methodology . . . 33
  - 3.3.3 Results . . . 37
  - 3.3.4 Conclusions . . . 38
- 3.4 An improved data-driven model of chemically-regulated motility . . . 41
  - 3.4.1 Study outline . . . 41
  - 3.4.2 Methodology . . . 42
  - 3.4.3 Results . . . 49
  - 3.4.4 Conclusions . . . 56
- 3.5 Summary . . . 59

4 Tumour growth **61**
4.1 Introduction . . . 62
4.2 Traditional models of tumour growth . . . 63
4.3 Modelling tumour spheroid growth in collagen matrices . . . 66
4.3.1 Study outline . . . 66
4.3.2 Methodology . . . 67
4.3.3 Results . . . 71
4.3.4 Conclusions . . . 73
4.4 A review on computational models of cancer glucose metabolism . . . 75
4.4.1 Glucose metabolism in cancer cells . . . 75
4.4.2 Modelling cellular systems and glucose metabolism . . . 79
4.4.3 Future perspectives . . . 83
4.5 Summary . . . 84

5 Pattern formation **87**
5.1 Introduction . . . 88
5.2 Methodology . . . 90
5.2.1 Study design . . . 91
5.2.2 Neuron representation . . . 93
5.2.3 Cell biological events . . . 93
5.2.4 Cell mechanics and physical interactions . . . 95
5.3 Results . . . 99
5.3.1 Simplified model of Homer-Wright rosette formation . . . 99
5.3.2 Numerical sensitivity analysis studies . . . 100
5.3.3 Tissue-scale simulations . . . 102
5.4 Summary . . . 104

6 Conclusions **107**
6.1 Introduction . . . 107
6.2 Main conclusions . . . 108
6.2.1 Single cell motility . . . 108
6.2.2 Tumour growth . . . 109
6.2.3 Pattern formation . . . 110
6.3 Future work . . . 111
6.3.1 Improved representations of the physical microenvironment . . . 111
6.3.2 Integrated intracellular models . . . 112
6.3.3 Optimization and parameter identifiability . . . 113
6.4 Thesis contributions . . . 114
6.4.1 Journal publications . . . 114
6.4.2 Conferences and workshops . . . 114
6.4.3 Supervising and mentoring activities . . . 115
6.4.4 Open-source software . . . 116
6.4.5 Collaborations . . . 116
6.4.6 Acknowledgements . . . 117

A Conclusiones **119**
A.1 Introducción . . . 119
A.2 Conclusiones generales . . . 120
A.2.1 Migración celular . . . 120
A.2.2 Crecimiento tumoral . . . 121
A.2.3 Formación de patrones . . . 122
A.3 Trabajo futuro . . . 123
A.3.1 Modelos intracelulares . . . 124
A.3.2 Optimización y identificabilidad de parámetros . . . 125
A.4 Contribuciones . . . 126
A.4.1 Publicaciones . . . 126
A.4.2 Participación en congresos y workshops . . . 126
A.4.3 Dirección de trabajos . . . 127
A.4.4 Software libre . . . 128
A.4.5 Colaboraciones . . . 128
A.4.6 Agradecimientos . . . 130

B Supplementary data **131**
B.1 Single cell migration . . . 132
B.1.1 Study 1: Model identification for the locomotive forces generator function 132
B.1.2 Study 1: Sensitivity analysis . . . 132
B.1.3 Study 2: Relevance of the administration methods of chemotactic substances 134
B.1.4 Study 2: Direct comparison between 2.5 mg/mL and 4.0 mg/mL matrices 136
B.1.5 Study 2: Spatial organization and intercellular distance . . . 136
B.1.6 Study 2: Model predictions evaluation of the estimated cell persistence time 138
B.2 Tumour growth . . . 139
B.2.1 Automatic cluster classification and quantification . . . 139
B.2.2 Oxygen diffusion and consumption dynamics . . . 143
B.3 Pattern formation . . . 144
B.3.1 Rosette formation at distinct adhesion and differentiation levels . . . 144

C PhysiCOOL: An open-source optimization Python library for PhysiCell **147**
C.1 Introduction . . . 148
C.2 Statement of Need . . . 148
C.3 Implementation . . . 149
C.3.1 Configuration file parser . . . 149
C.3.2 Black-box models . . . 150
C.3.3 Multilevel parameter sweeps . . . 151
C.4 Examples . . . 152
C.4.1 Simple model of logistic growth . . . 152
C.4.2 PhysiCell chemotaxis model . . . 154
C.4.3 Connecting to third-party libraries . . . 154
C.5 Future directions . . . 156

References **157**

List of Figures

1.1 Pre-clinical models in cancer research. 2
1.2 Phenomenological and mechanistic modelling approaches. 3
1.3 A simplified version of the Hallmarks of Cancer. 5
1.4 Visual summary of the dissertation's research objectives. 7

2.1 Representation of the multistep tumourigenesis process. 11
2.2 The Hallmarks of Cancer 12
2.3 Representations of the relationship between experimental and computational research. 16
2.4 Cell-based modelling frameworks. 18
2.5 Biological and physical rules in centre-based models 19

3.1 Cell motility mechanisms in 3D fibrous matrices. 27
3.2 Microfluidic devices. 29
3.3 Forces acting on the single-cell motility model. 35
3.4 Simulation and calibration workflow for the single-cell motility model. 38
3.5 Experimental and simulated results for the individual migration setup. 39
3.6 Experimental-computational study design to characterize fibroblast motility in response to chemotactic gradients. 42
3.7 Representation of the forces considered in our computational model. 45
3.8 Simulation and optimization workflow for the chemotaxis model. 48
3.9 Experimental and computational results for the travelled distances in 2.5 mg/mL matrices over 4 days. 51
3.10 Experimental and computational results for the travelled distances in 2.5 mg/mL matrices over 9 days. 53
3.11 Experimental and computational results for the travelled distances in 4.0 mg/mL matrices over 4 days. 54

4.1 The effect of matrix density on tumour growth. 63
4.2 Continuum and discrete models of tumour growth. 64
4.3 Simulation and calibration workflow for the single-cell motility model. 70
4.4 Representation of cell positions after five simulated days of tumour growth. 71
4.5 Evolution of cluster area growth over five simulated days. 72
4.6 Cluster eccentricity after 7 days of growth. 74
4.7 Cell metabolism differences in healthy and cancer cells. 76
4.8 Internal organization of a tumour spheroid. 78

4.9 Examples of multiscale hybrid model results. 82

5.1 Simplified representations of different types of rosettes. 89
5.2 Simplified representations of the neuronal differentiation process in healthy tissues and neuroblastoma. 92
5.3 Neuron implementation and rules. 94
5.4 Representation of the mechanics module which considered agent interactions based on adhesion/repulsion dynamics and neurite spring forces. 96
5.5 Computational model results obtained for an Homer-Wright (HW) rosette formed under ideal conditions. 100
5.6 Computational results of HW rosette formation at different differentiation rates and cell-cell adhesion levels. 101
5.7 Histological records and simulation data for different stages of neuronal differentiation 103

B.1 Results for a simplified model of cell motility based on probability density distributions. 133
B.2 Comparison between a simplified model of cell motility based on experimental data and *PhysiCell* results. 133
B.3 Sensitivity analysis for a simplified model of cell motility based on experimental data. 134
B.4 Displacement values for migrating fibroblasts over 4 days when platet derived growth factor-BB (PDGF-BB) administration was ceased after 48 hours of experiments. 135
B.5 Displacement values measured for migrating fibroblasts over 4 days in 2.5 mg/mL and 4.0 mg/mL collagen matrices. 137
B.6 Experimental and computational results for the intercellular distances between fibroblasts seeded in single-chamber devices. 137
B.7 Distribution of the fibroblast persistence times quantified experimentally at 96 hours of experiments. 139
B.8 Effect of the DBSCAN parameters on cluster metrics. 141
B.9 Cluster area quantification approaches. 142
B.10 Differences between the experimental and computational values for cluster eccentricity. 142
B.11 Effect of the ratio between the major and minor axes of an ellipse on cluster eccentricity. 143
B.12 Oxygen levels over the simulation timespan. 144
B.13 Computational results of HW rosette formation at different differentiation rates and cell-cell adhesion levels. 145

C.1 *PhysiCOOL*'s contributions and advantages to the *PhysiCell* ecossystem 150
C.2 Model and optimization results for the logistic growth example. 153
C.3 Model and optimization results for the chemotaxis example. 155

List of Tables

3.1 Reference parameter values for the single-cell motility model. 34
3.2 Comparison between the experimental and computational velocities generated by the single-cell model. 40
3.3 Reference parameter values for the chemotaxis model. 46
3.4 Model optimization results for the experiments where fibroblasts were grown in 2.5 mg/mL collagen matrices over 4 days. 52
3.5 Model optimization results for the experiments where fibroblasts were grown in 4.0 mg/mL collagen matrices over 4 days. 56

4.1 Reference parameter values for the tumour growth model. 67
4.2 Parameter values for the post-processing of tumour growth modelling results. . . 69
4.3 Summary of the reviewed multiscale models. 81

5.1 Reference parameter values for the model. 90

B.1 Relative increase of the mean distance travelled by fibroblasts in response to different concentrations of PDGF-BB. 134

C.1 Parameter values used in the multilevel optimization examples. 152

Abbreviations

2D two-dimensional

3D three-dimensional

ABM agent-based model

ATP adenosine triphosphate

BC Bhattacharyya coefficient

BO Bayesian Optimization

CA cellular automata

CBM centre-based model

CPM cellular Potts model

DBSCAN density-based spatial clustering of applications with noise

ECM extracellular matrix

EMT eptithelial-mesenchymal transition

FBA flux balance analysis

FDG-PET [18F]-fluorodeoxyglucose-positron emission tomography

GFP green fluorescent protein

GUI graphical user interface

HPC high-performance computing

HW Homer-Wright

M2BE Mutiscale in Mechanical and Biological Engineering

MAPK mitogen-activated protein kinase

MC monolayer channel

MCTS multicellular tumour spheroid

NHDF normal human dermal fibroblasts

NSCLC non-small cell lung cancer

OC opposite channel

ODE ordinary differential equation

PDE partial differential equation

PDGF-BB platet derived growth factor-BB

PDMS polydimethylsiloxane

RB retinoblastoma

ROS reactive oxygen species

RTK receptor tyrosine kinase

SBML Systems Biology Markup Language

CHAPTER 1

Introduction

Contents

1.1 Context 1
1.2 Motivation 4
1.3 Objectives 6
1.4 Document Structure 8

1.1 Context

Cancer is a major public health problem and a leading cause of death worldwide (Bray et al., 2021; Siegel et al., 2022). In fact, it was estimated that 19.3 million new cancer cases occurred worldwide in 2020, leading to almost 10.0 million deaths (Sung et al., 2021), while, in Europe, there were 4.0 million new estimated cases and 1.9 million cancer-related deaths (Dyba et al., 2021). Coincidentally, the economic burden of cancer is substantial, rising to a total of €199 billion in Europe in 2018 (Hofmarcher et al., 2020), and it affects both patients and healthcare systems. Therefore, it is increasingly more relevant to develop cost-effective strategies to prevent and treat cancer at early stages of development. Early diagnosis tools have proved to be invaluable in decreasing cancer mortality (Siegel et al., 2022). However, the advances in cancer prevention have not been matched by the progress in therapy (Jubelin et al., 2022). Cancer comprises a group of diseases that are still not fully understood that they may develop from any kind of cells in the body (Cooper, 2000). Furthermore, tumours are highly heterogeneous and patient-specific, meaning that the same therapy may not have the same expected effect when applied to different tumours, even when they are of the same type (Karolak et al., 2018; Marusyk and Polyak, 2010).

There have been significant advances in recent years in developing experimental studies of cancer development that may improve the understanding of this disease (Costa et al., 2020). Experimental models aim to replicate complex behaviour with simplified frameworks, which can range in complexity, as shown in Fig 1.1. *In vivo* models, for example, rely on living organisms, allowing for high resemblance between experiments and actual biological behaviour. However,

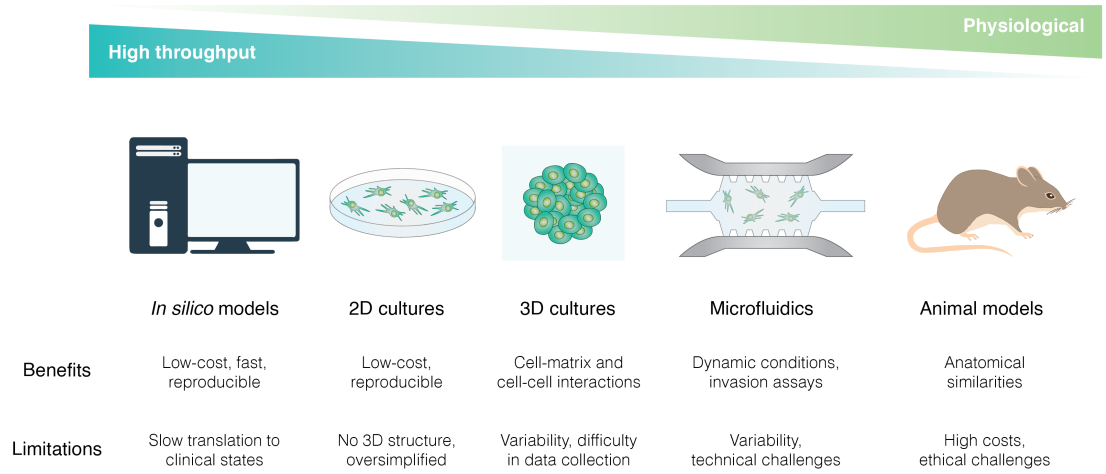

Figure 1.1: **Pre-clinical models in cancer research.** Benefits and limitations of *in silico*, *in vitro* and *in vivo* techniques employed to study cancer development. *In vitro* models are well-established as they are less expensive and faster than *in vivo* models, which are also known to raise ethical concerns. Despite being less biologically accurate, *in silico* models have been used consistently in combination with experimental assays to understand biological data and test hypotheses about its causality. Adapted from (De Pieri et al., 2021; Moya-Garcia et al., 2022).

they raise ethical questions and are difficult to quantify, given that tumours develop inside an organism. *In vitro* models, on the other hand, are performed outside of living systems (Hagiwara and Koh, 2020). Traditionally, *in vitro* models use two-dimensional (2D) substrates, yet this is now perceived as an oversimplification since cells in the human body are found in three-dimensional (3D) environments that modulate cellular behaviour and change in response to cell interactions (Jubelin et al., 2022). Remarkably, recent *in vitro* studies have been shifting their focus from 2D to 3D approaches since the latter capture cell behaviour in a manner that better resembles *in vivo* dynamics (Wu et al., 2014; Pedersen and Swartz, 2005). Nonetheless, experimental models are still expensive and time-consuming and often focus on specific events, thus failing to scale when trying to extrapolate their effect on a macroscopic view (Bekisz and Geris, 2020). As a result, there is a demand for instruments that enable more efficient investigation of biological systems.

Mathematics, specifically statistics, has been intensively used in biology and medicine to evaluate qualitative hypotheses through quantitative experimental and clinical data (Anderson and Quaranta, 2008; Pérez-García et al., 2016). For instance, statistical tests are used frequently to assess whether different environmental conditions and therapeutic agents affect tumour development. Advanced tools, such as mathematical modelling and computational simulations, are less likely to be found in the health sciences field, though, despite having a pivotal role in physics and engineering (Byrne, 2010; Victori and Buffa, 2018). Mathematical models are sometimes viewed as unduly simplistic by experimentalists and physicians who feel that cancer is a complicated disease that cannot be reduced to simple models (Gatenby and Maini, 2003). While it is true that

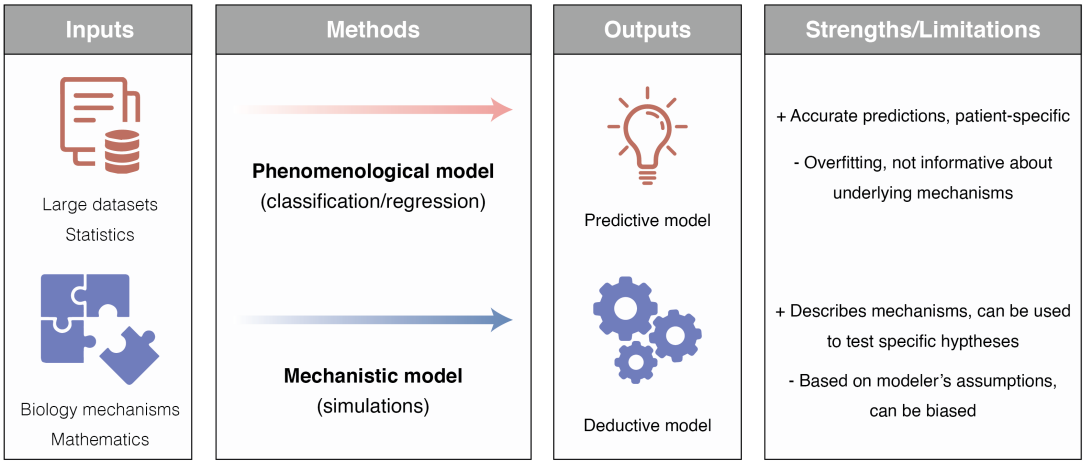

Figure 1.2: **Phenomenological and mechanistic modelling approaches.** Mechanistic models are based on generating new data from a set of assumptions rooted in biological mechanisms. Therefore, they are relevant to test biological hypotheses and when causality is paramount. Phenomenological models aim to reproduce data through statistics and/or machine learning approaches. Accordingly, they are very accurate yet fail to explain the mechanisms behind the biological data. Adapted from (Victori and Buffa, 2018; Baker et al., 2018).

phenomenological and statistical models can reproduce experimental data yet fail to explain it, mechanistic approaches can provide more information about the causes that lead to the biological behaviour (Transtrum and Qiu, 2016; Baker et al., 2018). A comparison between phenomenological and mechanistic methods is shown in Fig 1.2.

Phenomenological models are based on empirical observations, which they are able to reproduce and describe through statistics and machine learning approaches (Victori and Buffa, 2018; Baker et al., 2018). However, these predictions are based on results and are not directly derived from theory, providing no information about the mechanisms that lead to these outputs. On the other end, mechanistic models are based on causality and can be used to translate biological processes into mathematical equations and create model predictions based on concepts that can be interpreted by experimentalists (Craver, 2006). Consequently, the potential success of these methods to solve problems in computational oncology, i.e., the use of mathematical modelling and computational simulation techniques to study cancer, relies on significant knowledge of both the biological processes that occur in cancer and the mathematical and computational tools that may be applied to study them (Banwarth-Kuhn and Sindi, 2020; Rockne and Scott, 2019; Karolak et al., 2018). Henceforth, establishing collaborations between experimentalists and mathematicians or engineers is often required to produce meaningful models of biological systems.

Computational models allow for the investigation of topics that cannot be tested directly.. Furthermore, developers can integrate models into multiscale platforms to evaluate how events at different scales may interact and originate distinct behaviours (Rockne and Scott, 2019). Cell-based or agent-based models (ABMs) are a specific type of modelling approach that considers cells as individual agents that follow a set of rules that dictate how they behave and interact with other

agents and the surrounding microenvironment (Cannata et al., 2013). ABMs offer an overview of how population behaviour arises from single-cell rules where each cell or agent may act differently to reproduce cell heterogeneity (Montagud et al., 2021). In addition, ABMs are constructed in such a way that they may be coupled with models of varying scales. For example, intracellular models can be tested to study subcellular events like cell metabolism and regulatory networks (Rejniak and Anderson, 2011; Montagud et al., 2021), and reaction-diffusion models of the substances present in the cellular microenvironment may also be taken into account to model the cells' response to therapeutic chemical agents and modulate cell behaviour as a function of the local nutrient concentration (Macklin et al., 2012; Jagiella et al., 2016).

Multiscale ABMs are large and intricate computational models characterized by a substantial number of parameters. Therefore, these models are challenging to evaluate and calibrate (Hasenauer et al., 2015). Yet, the ever-growing amount of high-throughput experimental data can be used to empower data-driven optimization routines to calibrate computational models. Access to big data makes it possible to use appropriate statistical and machine-learning methods to inform mechanistic models. These new data-driven models may provide insights that conventional models based on prior assumptions may fail to achieve. Consequently, it is crucial to apply these model calibration and optimization methods during the model development process to assure biological validity.

1.2 Motivation

Mathematical models often assume that complex macroscopic biological behaviour arises from simple rules and interactions at the microscopic level (Banwarth-Kuhn and Sindi, 2020). In other words, it is possible to break down a complicated problem into a subset of small questions that, when solved, give rise to a global solution. This approach is commonly defined as *"bottom-up"* since it first considers small components of a system and progresses upwards. It must be noted that, in cancer biology, a model is not necessarily equal to the sum of its parts and that complex interactions between model components may be needed to be considered to replicate biological behaviour (Anderson and Quaranta, 2008). However, bottom-up approaches such as ABMs are a valuable starting point to unravel the mechanisms that drive cancer progression (Yu and Bagheri, 2020). Furthermore, ABMs can integrate stochastic modelling techniques to generate an additional complexity level where the model results are not solely directly defined by static and well-defined rules.

Despite the heterogeneous nature of cancer, some common trends among different cancer types can still be identified. This fact is invaluable in cancer modelling because it allows modellers to establish, at least, some of the basic building blocks that describe common tumour malignant behaviour. Hanahan and Weinberg (2000) developed a theoretical framework called *"The Hallmarks of Cancer"* that aims to summarize the enabling characteristics observed in cancer at a general level which promote malignant development. Moreover, new features have been proposed as *"emerging hallmarks"* recently, meaning that they favour malignant transformation but may not

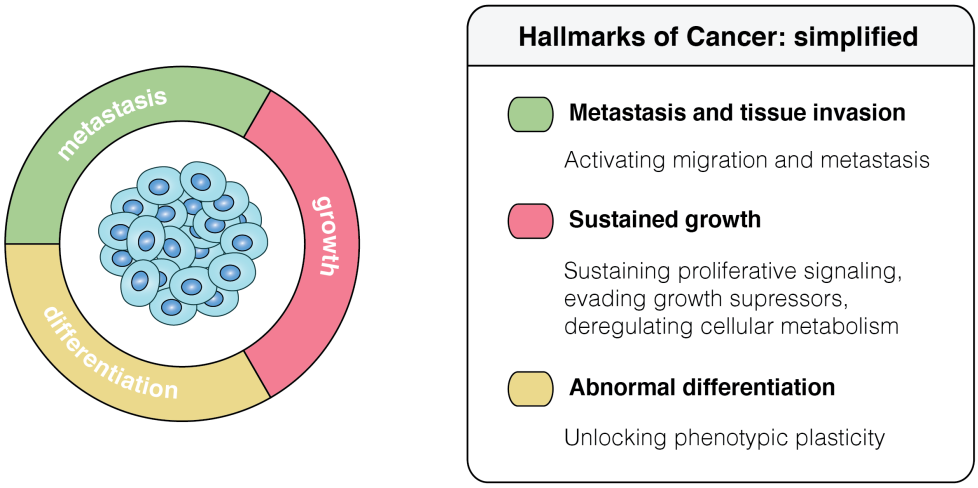

Figure 1.3: **A simplified version of the Hallmarks of Cancer.** Despite its complexity and heterogeneity, certain common characteristics can be found in cancer. In this dissertation, three main characteristics are explored, namely migration and metastasis, sustained growth and abnormal differentiation. Furthermore, we aim to explore how these different hallmarks may interact to generate specific biological behaviours.

be ubiquitous in cancer (Hanahan, 2022). Well-established hallmarks include, but are not limited to, characteristics that support unregulated and aberrant tumour growth and migration activation, which enables metastasis. Besides, phenotypic plasticity is an interesting hallmark, which alters the normal cell differentiation process and can lead to poorly differentiated cells. A simplified version of the Hallmarks of Cancer that includes the aforementioned characteristics is represented in Fig 1.3. Hence, it can be established that computational models of cancer should can take advantage of this framework as a learning mechanism to understand which components should be represented to reproduce at least some of these hallmarks.

Computational models can be overly focused on a single hallmark. For instance, several models have been designed to study migration activation and metastasis (González-Valverde and García-Aznar, 2018; Vargas et al., 2020), tumour growth (Gerlee and Anderson, 2007; Hyun and Macklin, 2013) or cell differentiation (Zubler, 2009; Colombi et al., 2020). Nevertheless, these models tend to assume that these aspects are isolated when in reality, biological events start at an individual level but evolve to more complex patterns when multiple hallmarks are considered. For example, migration and proliferation can be mutually exclusive, as stated by the *"Go or Grow"* hypothesis (Hatzikirou et al., 2012; Kim et al., 2014). In particular, this theory explains that migrating cells may have altered proliferation dynamics and arrest their cell cycle when migrating. Moreover, the emergence of specific spatial patterns in cell organization, such as rosettes observed in neuroblastoma, a type of paediatric cancer, are thought to arise from deregulated differentiation processes coupled to abnormal proliferation and differential adhesion between cells (Wippold and Perry, 2006; Duarte Campos et al., 2019). Accordingly, it is increasingly relevant to develop

computational models which reproduce multiple aspects of cell behaviour.

The mathematical biology community has developed several ABM-based frameworks that simultaneously consider events such as cell proliferation, death, migration and differentiation. Examples include *BioDynaMo* (Breitwieser et al., 2022), *Chaste* (Mirams et al., 2013), *CompuCell3D* (Swat et al., 2012), *Hybrid Automata Library* (Bravo et al., 2020), iDynoMiCs (Lardon et al., 2011), *Morpheus* (Starruß et al., 2014) and *PhysiCell* (Ghaffarizadeh et al., 2018). Most of these tools do not require extensive programming and mathematical modelling expertise and are accessible to biologists and clinicians who prioritize biological behaviour over mathematical formalism. Moreover, they are also optimized to take advantage of the increasingly available computational power. Besides, they are accessible and extensible, meaning that new users can build on previous model iterations and focus on creating new extensions to solve specific questions (Rockne et al., 2019; de Montigny et al., 2021; Letort et al., 2019).

Taking all of this into consideration, it is evident that computational biology is progressing towards integrated frameworks and the development of computational tools that can be utilised and shared by the community to build robust models. Therefore, this dissertation explores how these frameworks can be used to study cancer, starting from a simple single-cell model and progressing in complexity to reproduce collective behaviour. Moreover, model calibration is a topic of interest in this dissertation. In particular, manual and automatic workflows were developed to shown that quantitative and qualitative data can be employed in model calibration routines.

1.3 Objectives

The global research objective of this dissertation is to use computational models to reproduce and predict cell behaviours observed frenquently in cancer systems from a bottom-up perspective. Due to the high complexity of cancer diseases, three specific events were selected to be explored extensively: (i) individual cell migration, (ii) tumour growth and (iii) pattern formation dynamics, as shown in Fig 1.4. The order in which these processes were studied in this thesis is relevant as the questions explored in each study grow in complexity as these studies are presented, from individual to collective cell behaviour. Firstly, single cells are simulated, and their interactions with the surrounding extracellular matrix (ECM) are characterized to reproduce cell migration behaviour. Subsequently, cell-cell interactions are introduced to model tumour growth observations. Finally, both single-cell and cell-cell are considered and complemented by differentiation to identify how specific histological patterns arise from collective cell behaviour. Moreover, parameter exploration routines were employed to evaluate model behaviour, and model calibration techniques were applied throughout this dissertation when data was available to optimize the models developed.

Research objective 1: Understanding how the extracellular microenvironment regulates single-cell motility

i Implementing and calibrating a computational model of single-cell motility which takes into account the impact of the mechanical properties of the surrounding microenvironment on cell

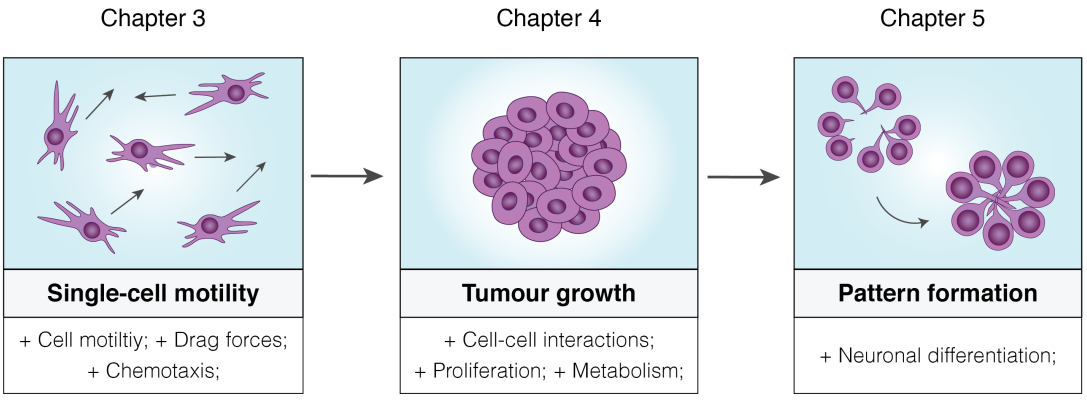

Figure 1.4: **Visual summary of the dissertation's research objectives.** This dissertation aims to study cancer biology through a bottom-up approach, adding new complexity levels as chapters are discussed. The first model on implementation focuses on simulating and calibrating individual cell motility. Subsequently, collective behaviour starts being considered, with the addition of cell-cell interactions and proliferation. Lastly, tissue organization is studied by introducing neuronal differentiation.

migration;

ii Extending the computational model to consider the effect of migration-inducing chemical substances on cell motility;

iii Implementing an automatic optimization pipeline to calibrate the extended model using new quantitative data and identify the model parameters that explain the observed *in vitro* behaviour;

Research objective 2: From single cell migration to collective organisation regulating tissue growth

i Reviewing computational models of tumour growth available in the literature with an emphasis on the role of metabolism reprogramming;

ii Extending the single-cell motility model to account for cell-cell interactions to study tumour formation and how it is modulated by mechanical forces imposed by the environment;

iii Calibrating the model above with qualitative data from spheroid growth assays performed in microfluidic devices;

Research objective 3: Modelling pattern formation as a differentiation process

i Implementing a model that introduces neuronal morphology and differentiation;

ii Identifying how differentiation stages and mechanical interactions regulate the formation of histological patterns, namely Homer-Wright rosettes;

1.4 Document Structure

This dissertation is divided into six chapters that guide the reader through the design, implementation and validation of cell-based computational models of increasing complexity that aim to unravel different hallmarks of cancer. **Chapter 1** contextualizes the motivation and objectives of this thesis and communicates the relevance of using computational modelling in cancer biology. **Chapter 2** extends on the basic background concepts to understand the biological and computational aspects explored in this dissertation, focusing on the hallmarks of cancer to be presented in the following chapters and on cell-based modelling techniques. Having established the main theoretical background for this dissertation, **Chapter 3** covers the implementation of the computational model that serves as the foundation work for the thesis. Specifically, it is addressed how a cell-based model may be used to study single-cell motility and the effect of mechanical and chemical factors on the cells' migration patterns. The use of experimental data to optimize computational models is also discussed, and the design of an optimization pipeline is described. Building on this work, **Chapter 4** extends the computational framework implemented in order to characterize how the combined effects of cell migration, cell-cell mechanics and proliferation modulate tumour growth. In addition, this chapter explores a new complexity level by reviewing the influence of altered cell metabolism in tumour growth and how metabolic reprogramming has been addressed in previous computational studies. **Chapter 5** adds the last degree of complexity to the computational model and presents a framework that includes all the biological phenomena studied in the previous chapters and introduces neuronal differentiation. This chapter assesses specifically how morphological patterns arise from the interplay between proliferation, differentiation and cell-cell interactions. Lastly, **Chapter 6** summarizes the results obtained throughout this dissertation by presenting its main conclusions and proposing possible future research lines to extend this work. Furthermore, this chapter lists the publications and contributions that resulted from this work.

CHAPTER 2

Background

Contents

2.1 Introduction **9**
2.2 Biological background **10**
2.2.1 Cancer development 10
2.2.2 The hallmarks of cancer 11
2.3 Computational oncology **16**
2.3.1 Modelling cancer systems 16
2.3.2 Continuum vs discrete models 17
2.3.3 Centre-based models (CBMs) 18
2.4 Summary **23**

2.1 Introduction

Cancer is a highly complex health condition which comprises over 200 disease types characterized by aberrant and excessive cell proliferation (Cooper, 2000; Pérez-García et al., 2016; Hassanpour and Dehghani, 2017). Healthy and differentiated cells are subject to several control mechanisms that regulate their life cycle and death. In cancer cells, though, these systems become deregulated due to gene mutations, causing abnormal growth and malignant features that result in the formation of a tumour and its subsequent progression (Weinberg, 2013). Due to its complexity, cancer is still not fully understood and frameworks that enable our comprehension of this disease are still needed, from conceptual models that explain cancer biology to practical applications through which new hypotheses can be generated and tested.

Mathematical models have been used in cancer biology for over 60 years (Byrne, 2010). Although traditional mathematical models can be solved manually, model developers have increasingly relied on computational tools as models became more complex and the access to computational resources increased over the years (Bekisz and Geris, 2020). The main goal of creating mathematical and computational representations of biological systems is to generate hypotheses

that can be tested, validated and guide the design of new experiments (Victori and Buffa, 2018; Enderling and Wolkenhauer, 2021). Accordingly, models should provide a level of simplification and abstraction to intricate biological phenomena, yet they should still be complex enough to depict the general behaviour of these systems accurately (Victori and Buffa, 2018; Transtrum and Qiu, 2016).

This chapter aims to provide readers with a global overview of the theoretical background that serves as the foundation of this dissertation. Given the multidisciplinary nature of the models featured in this work, biological and computational concepts are addressed to provide readers with the fundamental knowledge to understand these models and their results. Henceforth, Section 2.2 presents an introduction to cancer biology, exploring the general mechanisms that lead to cancer development and highlighting the specific biological events studied in this PhD thesis. In addition, section 2.3 describes how mathematical and computational techniques have been increasingly used in biology and medicine to understand biological data and formulate hypotheses to explain this data. General concepts on the use of mathematics to recreate biological systems are introduced and the specific modelling frameworks employed in this dissertation are explored in more detail.

2.2 Biological background

2.2.1 Cancer development

Cancer is a genetic disease that originates from mutations in cancer driver genes, namely oncogenes, tumour-suppressor genes and stability genes (Vogelstein and Kinzler, 2004; Vogelstein et al., 2013). Oncogenes derive from proto-oncogenes which regulate normal cell proliferation and promote cell division under controlled conditions (Kontomanolis et al., 2020). However, when activated by mutations, oncogenes cause aberrant cell proliferation that evades the control mechanisms present in normal cells (Shortt and Johnstone, 2012). Conversely, tumour-suppressor genes are generally responsible for inhibiting cell proliferation and avoiding tumour formation (Cooper, 2000). Therefore, gene mutations in cancer lead to the inactivation or loss of tumour-suppressor genes, resulting in sustained growth and apoptosis evasion (Payne and Kemp, 2005). Lastly, stability genes, also known as DNA repair genes, assure that genetic alterations are minimal in healthy cells (Vogelstein and Kinzler, 2004; Jeggo et al., 2016). Henceforth, their inactivation enables cancer cells to maintain altered genomes that favour malignant behaviour.

Mutations are essential in tumourigenesis, a multistage process through which healthy cells continuously acquire malignant tumour characteristics. Tumourigenesis can be divided into three stages: (i) initiation, (ii) promotion and (iii) progression (Weston and Harris, 2003), as shown in Fig 2.1. During initiation, single progenitor cancer cells acquire mutations (e.g., oncogene activation or tumour suppressor inhibition) that confer them advantageous phenotypes (Cooper, 2000; Chaffer and Weinberg, 2015). DNA damage in cancer progenitor cells may arise from endogenous and exogenous agents, such as reactive oxygen species (ROS) and radiation (Barnes et al., 2018). If the mutated cells' control mechanisms fail to repair DNA damage or induce cell death, cells

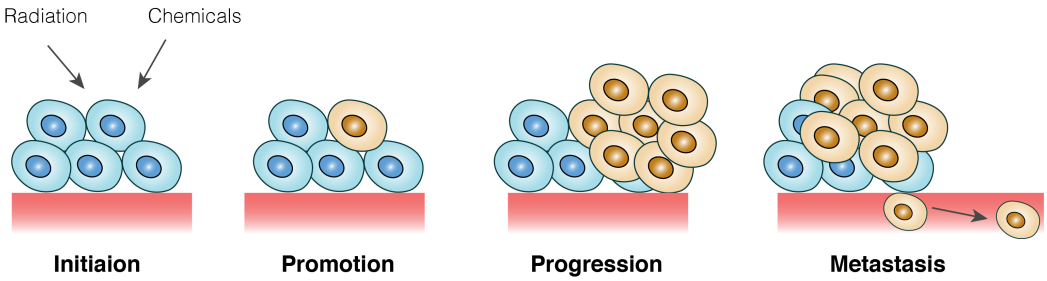

Figure 2.1: **Representation of the multistep tumourigenesis process.** Tumourigenesis is a process through which normal and healthy cells progressively acquire mutations that result in malignant adaptations, e.g., sustained proliferation and invasion activation, and make them selectively more apt to survive.

become initiated and proliferate, transmitting malignant properties to their descendants. Furthermore, additional gene mutations accumulate as cancer cells divide during the promotion stage, making them more resistant to the control mechanisms that should arrest unregulated proliferation and induce cell death (Sarkar et al., 2013; Cao, 2017). In addition, environmental cues can further drive tumour promotion (Laconi et al., 2008).

Lastly, cancer cells may acquire the ability to escape the primary tumour and invade other tissues in the final stage of tumour progression, creating secondary tumours, also referred to as metastases (Sarkar et al., 2013). The sequence that defines the steps that tumour cells undergo when they escape the primary tissue and colonize other parts of the human body is commonly referred to as the *invasion-metastasis cascade* (Valastyan and Weinberg, 2011). The metastatic cascade begins with the invasion of the surrounding microenvironment by tumour cells, which first must become dissociated from their original tissue by breaking cell-cell and cell-matrix adhesions (van Zijl et al., 2011). Subsequently, cancer cells enter blood vessels through which they circulate until they reach another tissue (Morgan-Parkes, 1995). At this point, individual cancer cells invade the new tissue and attempt to colonize it through the formation of secondary tumours, which are regulated by the physical properties of the new tissue (Massagué and Obenauf, 2016). Accordingly, it can be concluded that the malignant characteristics acquired by cancer cells during tumourigenesis are essential for cancer promotion and progression.

2.2.2 The hallmarks of cancer

Cancer is complex and challenging to understand due to its complexity and heterogeneity. With this idea in mind, Hanahan and Weinberg (2000) introduced an original conceptual framework named "The Hallmarks of Cancer" that enumerates functional capabilities shared among tumour types that differentiate them from healthy cells and that enable the formation of malignant tumours. The hallmarks of cancer originally consisted of six features: (i) resisting cell death, (ii) sustaining proliferative signalling, (iii) evading growth suppressors, (iv) activating invasion and metastasis, (v) enabling replicative immortality, and (vi) inducing or accessing vasculature (Hanahan and

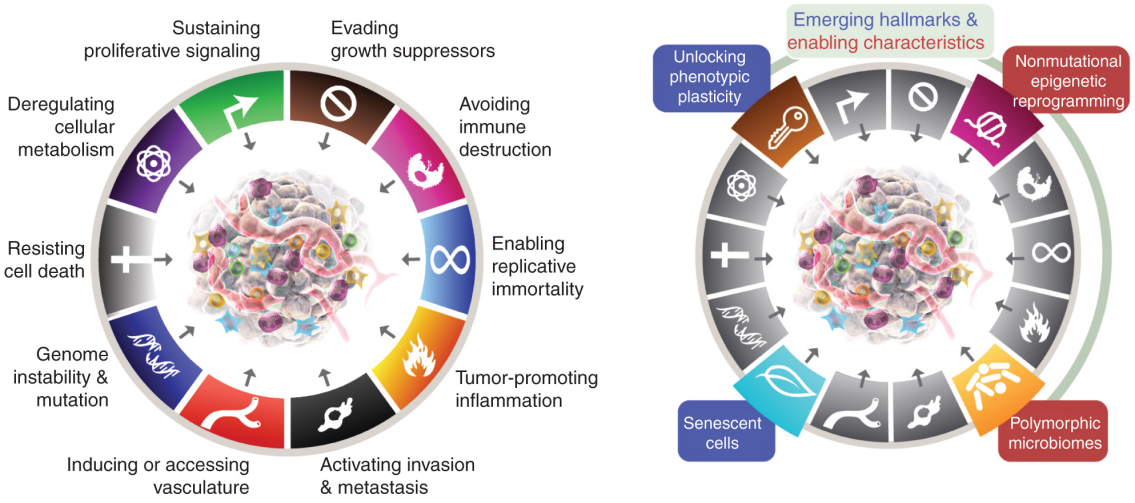

Figure 2.2: **The Hallmarks of Cancer**. These characteristics compose a theoretical framework that defines the commonalities between cancer cells. The left subfigure shows the hallmarks and enabling characteristics proposed in the past (years 2000 and 2011) that have been proven sufficiently characterized and validated to be recognized as functional capabilities acquired by malignant tumour cells. The right subfigure highlights the emerging hallmarks and enabling characteristics proposed recently (year 2022) to integrate the Hallmarks of Cancer. Figure taken from (Hanahan, 2022).

Weinberg, 2000). However, this framework has been continuously updated as new research studies have illuminated neoplastic patterns that were not recognized in the past (Hanahan and Weinberg, 2011; Hanahan, 2022). Henceforth, two new hallmarks have been introduced, namely reprogramming cellular metabolism and avoiding immune destruction, and six other enabling characteristics have been suggested. These include unlocking phenotypic plasticity, tumour-promoting inflammation, genome instability and mutation, non-mutational epigenetic reprogramming, senescent cells and polymorphic microbiomes. A visual summary of the Hallmarks of Cancer is shown in Fig 2.2.

As previously stated in Chapter 1, this dissertation focuses on three aspects of malignant development that can be associated with one or more hallmarks:

i **Single-cell motility**: Activating invasion & metastasis;

ii **Tumour growth**: Sustained proliferative signalling, evading growth suppressors, enabling replicative immortality and deregulating cellular metabolism;

iii **Pattern formation**: Unlocking phenotypic plasticity (i.e., abnormal differentiation);

Therefore, we present a general overview of these hallmarks and their relevance to the biological phenomena to be explored in this dissertation. In particular, we discuss (i) how cancer cells acquire metastatic capacities and define their migration patterns, (ii) the relevance of sustained growth and metabolic reprogramming to promote tumour progression and (iii) how specific spatial patterns arise from blocked differentiation. We highlight that, even though we focus on specific hallmarks

to describe these behaviours, complex cancer behaviour arises from a combination of different hallmarks (e.g., cell migration can have an impact on solid tumour formation).

Metastasis and tissue invasion

Understanding cell motility and its dependence on the physical and chemical properties of the extracellular microenvironment is of particular interest in cancer biology. The invasion of surrounding tissues by tumour cells, a process known as metastasis, represents a critical step in tumour development and leads to a significant decrease in survival rates (Clark and Vignjevic, 2015; Yamaguchi et al., 2005). Metastasis relies on the eptithelial-mesenchymal transition (EMT), a process through which cells lose epithelial features and acquire mesenchymal properties (Ribatti et al., 2020). In other words, epithelial cells that were originally in a stable tissue connected by cell-cell junctions (e.g., E-cadherin adhesions) lose these connections, changing their morphology to become more similar to fibroblasts and acquiring invasive properties (Yang et al., 2020). On the other hand, cell-matrix adhesions are promoted, which enables cells to invade the surrounding tissue. Subsequently, tumour cells enter blood and lymphatic vessels through intravasation and travel to other parts of the body, where they extravasate to colonize other tissues (Morgan-Parkes, 1995). Remarkably, it has been observed that cells favour the regression to epithelial states to form secondary tumours (Yao et al., 2011).

Several previous studies have aimed to assess the pathophysiological processes of primary tumours and explain how the properties of the extracellular microenvironment induce and guide metastasis (Paszek et al., 2005; Seewaldt, 2014; Barney et al., 2016; Polacheck et al., 2013). The ECM is the non-cellular component of tissues that provides support and structure to the cells (Frantz et al., 2010; Kim et al., 2011). The ECM is mainly composed of water, proteins (e.g., collagen, elastin and fibronectin) and polysaccharides. The quantities at which these components are present vary significantly from tissue to tissue and define their characteristic mechanical properties (Butcher et al., 2009; Pizzo et al., 2005). Previous studies have shown that properties such as ECM stiffness influence migration (Guimarães et al., 2020). Moreover, it has been shown that the mechanical properties of the tumour microenvironment differ from those found in normal tissues as a result of the tumour cells' ability to sense and modify their environment (Kim et al., 2011; Yamaguchi et al., 2005; Lu et al., 2012).

Sustained growth

Abnormal proliferation has been recognized as the most fundamental characteristic among all cancer cells for multiple decades (Hanahan and Weinberg, 2011). In fact, several hallmarks that enable sustained proliferation at high rates and inhibit cycle arrest can be associated with abnormal growth in tumour cells (Hanahan and Weinberg, 2000; Feitelson et al., 2015). Tumours originate from small colonies of cells which acquired these enabling characteristics due to mutations (Weston and Harris, 2003). For instance, the deregulation of the mitogen-activated protein kinase (MAPK) cascade have been shown to enable the proliferative potential of cancer cells (Dhillon et al., 2007).

In healthy organisms, MAPK pathways are signal transduction pathways that are activated when growth factors bind to glycoproteins of the receptor tyrosine kinase (RTK) family to promote growth (Braicu et al., 2019). When the effectors of these pathways mutate in cancer cells, it leads to constitutive activation and continuous signal transduction. In other words, cancer cells no longer require external stimuli provided by growth factors to activate growth mechanisms through this pathway. The retinoblastoma (RB) pathway constitutes another example of how unregulated signalling networks promote cancer development (Hanahan and Weinberg, 2011). In particular, the RB protein acts a tumour suppressor that regulates whether cells should advance in the cell cycle. Hence, when it becomes silenced, cancer cells are able to proliferate without control (Cress et al., 2014).

Cancer is considered an evolutionary process, and it is accepted that tumour cells mutate over time and that cells with better adaptive characteristics are selected and continue to proliferate and develop into solid tumours (Casás-Selves and DeGregori, 2011; Vendramin et al., 2021). At initial stages of development, these tumours are avascular and have small dimensions (Jiang et al., 2005; Chaplain, 1996). Thus, cells rely on diffusion to obtain nutrients from the microenvironment. However, once critical tumour dimensions are reached (i.e., a radius of approximately 400-500 µm), diffusion alone is insufficient for nutrients to reach the tumour core (Hirschhaeuser et al., 2010; Casciari et al., 1992a). Consequently, cancer cells must trigger angiogenesis to induce vascularisation that will supply tumours with the nutrients required to survive and continue to grow (Lugano et al., 2020). Accordingly, the ability to promote or access vasculature is relevant to sustain growth, even though it is not explored in this dissertation.

Additionally, metabolism reprogramming has been increasingly recognized as a mechanism to sustain growth. Cell metabolism is the series of biochemical reactions that enable cells to produce the energy required for survival and maintenance. Under some circumstances, metabolism may be altered to fit the cells' energy requirements, for instance, in proliferating cells with high energy demands (Jones and Thompson, 2009; Lunt and Vander Heiden, 2011). Since tumour cells are highly proliferative, metabolism plays a crucial role in cancer progression since it supports the aberrant proliferation and survival dynamics of tumours. Studies have shown that cancer cells can reprogram their metabolism and favour metabolic pathways that enable cell division at fast rates but are less efficient, such as aerobic glycolysis (Hanahan and Weinberg, 2011; Warburg, 1956b; Palm, 2021). In fact, altered metabolism is the foundation for [18F]-fluorodeoxyglucose-positron emission tomography (FDG-PET), a well-established scanning technique employed to identify and diagnose tumours (Danhier et al., 2017; Martinez-Outschoorn et al., 2017). Specifically, FDG-PET quantifies how much FDG, a glucose analogue, is consumed by cells under the assumption that most cancer cells consume more glucose (Duhaylongsod et al., 1995). Even though metabolic reprogramming has been studied more in recent years, there are still many unanswered questions regarding why tumour cells evolved to metabolize glucose through energetically inefficient pathways (Liberti and Locasale, 2016).

Abnormal differentiation

Cellular differentiation is the process through which unspecialised cells mature and acquire characteristics specific to a given tissue type (Jögi et al., 2012). Unlike progenitor stem cells, healthy adult stem cells are already differentiated and have limited potential to generate different cell types (Gupta et al., 2019). Under specific conditions, cells have been shown to evade this terminal differentiation state and acquire phenotypic plasticity, i.e., the ability to change their phenotype. For instance, phenotypic plasticity has been explored extensively in the literature as a relevant component of development, tissue repair and regeneration (Jessen et al., 2015; Fusco and Minelli, 2010). However, this biological event has also been observed in cancer development. In particular, previous studies have shown that cancer cells can dedifferentiate, reversing to progenitor-like stages, they may remain in a poorly-differentiated stated, or they can transdifferentiate, acquiring characteristics from a tissue different from that they should have originated (Hanahan, 2022).

Phenotypic plasticity is generally less accepted as a cancer hallmark than abnormal growth and metastasis because it occurs in specific cancer types (Hanahan, 2022; Yuan et al., 2019). For instance, neuroblastoma is the most lethal and most common extracranial solid tumour in infants (Maris, 2010; Wright, 1910; Shimada et al., 1999b) and it arises from blocked neuronal differentiation during development. Neuroblastoma tumour cells derive from neural crest cells that differentiate into sympathetic neurons during embryogenesis in response to the presence of neuronal growth factors (Johnsen et al., 2019; Kholodenko et al., 2018). However, genetic alterations and environmental cues halt the differentiation process (Mora and Gerald, 2004). The stage at which this blockage occurs defines the neuroblastoma tumour subtypes, which can be classified into three principal categories: undifferentiated, poorly-differentiated, and differentiating (Shimada et al., 1999b). Undifferentiated neuroblastic tumours tend to be more malignant that those that became more differentiated and acquired characteristics similar to those of healthy tissues (Ponzoni et al., 2022). Besides, the morphology of neuroblastic cells differs based on their differentiation grade.

Progenitor cells are initially undifferentiated and present a round morphology, yet during neuronal development these cells create and extend neuronal processes to establish the structure of a neuron (Pahlman and Hoehner, 1996). In poorly-differentiated neuroblastoma, this leads to the formation of neurons with under-developed neuronal processes (Wippold and Perry, 2006). In addition, the actual spatial distribution of these poorly-differentiated cells is an interesting aspect of neuroblastoma, given that they are organized as rosettes (Wright, 1910). Rosettes are small groups of cells organised in circular configurations around a common core that, in neuroblastoma, is filed with fibres (Wippold and Perry, 2006). It is currently accepted that the rosettes observed in neuroblastoma arise from the mechanical interactions between the cells' under-developed neuronal processes, which adhere and become entangled, resulting in a circular arrangement of the cell bodies around this common core (Moss, 1983). However, this theory has not been confirmed since it is not possible to study neuroblastoma during development *in vivo*, and the ability to recreate these structures *in vitro* is limited with the current technology (Duarte Campos et al., 2019).

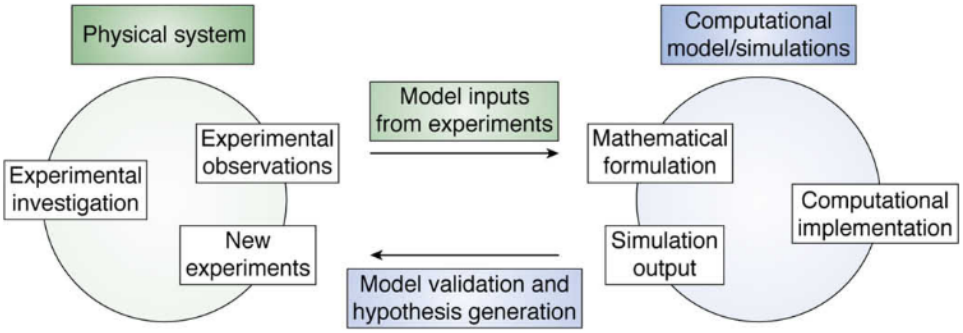

Figure 2.3: **Representations of the relationship between experimental and computational research.** Experimental data can be integrated into computational models that aim to propose mechanisms that explain how cellular behaviour arises. In turn, the hypotheses generated by the computational approach can be employed to plan future experiments with a more specific research question. Adapted from (Banwarth-Kuhn and Sindi, 2020).

2.3 Computational oncology

2.3.1 Modelling cancer systems

Designing a computational model of a biological system usually consists of posing a question, stating the assumptions and variables that should be included in the model and translating these conditions into equations and code (Byrne, 2010). Additionally, simulations should be run, and model results must be compared with the original research question to evaluate the model's validity. Accordingly, it is common to establish collaborations between experimentalists and modellers to guarantee that biological assumptions are translated to mathematical terms correctly, and that model results are appropriate. Fig 2.3 shows a representation of the common discourse between experimental and computational researchers to implement models that integrate real-world data, which, in turn, can generate hypotheses that will guide the planning of future experiments (Banwarth-Kuhn and Sindi, 2020).

Mathematical models are convenient tools to generate and test hypothesis, yet they are unable to thouroghly and flawlessly depict nature, as expressed by the prevalent aphorism *"All models are wrong, but some are useful"* (Box, 1976). Consequently, mathematical modelling should be regarded as an iterative process (Byrne, 2010). A model may generate hypotheses that appear to be biologically valid. However, it is critical that these predictions are not taken as truth and that further experiments are designed and performed to tests these hypotheses. Furthermore, models can be implemented in distinct manners, from traditional ordinary differential equations (ODEs) and partial differential equations (PDEs) models to more recent approaches like ABMs, based on their suitability to break down and understand the research problem (Bekisz and Geris, 2020).

2.3.2 Continuum vs discrete models

Biological systems can be modelled through distinct mathematical approaches, including continuum, discrete and hybrid frameworks (Edelman et al., 2010; Rejniak and Anderson, 2011). The spatial resolution at which cancer cells are reproduced differs between these implementations, with models ranging from microscopic, individual cell-based to more macroscopic, cell population-based representations (Altrock et al., 2015; Preziosi and Tosin, 2009). Continuum approaches, such as ODEs and PDEs, usually consider tumours cell populations at a large scale (Schaller and Meyer-Hermann, 2006). In other words, tumours are generally represented as a single cell population whose size changes over time or as group of different cell populations (Araujo, 2004). Therefore, the ability to account for spatial heterogeneity is limited and does not consider cell-level detail. Nonetheless, the behaviour of individual cells and stochastic events are not taken into account, and thus continuum approaches and cannot model heterogeneous behaviour at the cell-level (Schaller and Meyer-Hermann, 2006). Conversely, discrete modelling techniques, also known as ABMs, can achiev a higher level of detail since they simulate tumours as a group of individual cells acting as agents that follow a set of rules and interact with each other and the microenvironment (Metzcar et al., 2019; Van Liedekerke et al., 2015).

ABMs describe cells at the individual level and are thus particularly able to reproduce the heterogeneous behaviour of biological systems (Metzcar et al., 2019) and the interactions between cells and the microenvironment (Mathias et al., 2022; Macnamara, 2021; Gonçalves and Garcia-Aznar, 2021). For example, cellular automata (CA) models consider that each cell is represented by a single voxel, whereas cellular Potts models (CPMs) simulate a single cell as more than one voxel. Thus, an approximate cell morphology can be estimated, and the mechanical interactions with the surrounding neighbours can be computed (Macnamara, 2021). Off-lattice approaches do not consider that the domain is divided into a grid, enabling cells to move freely through space based on the forces exerted on them (Metzcar et al., 2019). centre-based models (CBMs) are an example of the off-lattice approach, and they consider that cells are characterized by their central position and a simplified geometry, such as a sphere. More sophisticated techniques that describe cell shape and allow for cell deformation (e.g., vertex-based and discrete element models) have also been developed, yet they are computationally expensive to simulate large tumours (Rejniak and Anderson, 2011; González-Valverde and García-Aznar, 2017). However, efficient parallelization strategies can be used to take advantage of computational resources and improve the speed of ABM-based simulations (Montagud et al., 2021; Fachada et al., 2017).

Hybrid modelling is a technique that can be used to couple detailed discrete descriptions of cellular systems at the cell level with continuum models of the surrounding microenvironment (Rejniak and Anderson, 2011). Specifically, hybrid modelling is used to study tumour growth as the result of the response of individual cells to the concentration of substances such as nutrients, metabolic waste and therapeutic agents that diffuse and are consumed/produced in the system (Altrock et al., 2015; Gerlee and Anderson, 2007). The spatial distribution of the chemical species present in the microenvironment is generally modelled as a set of reaction-diffusion equations

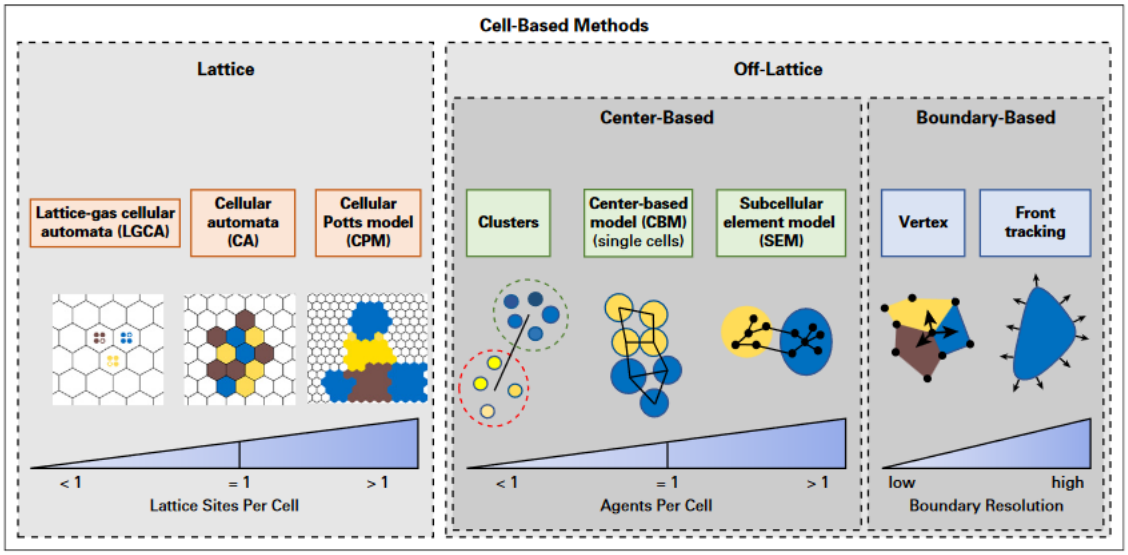

Figure 2.4: **Cell-based modelling frameworks.** Discrete models can generally be categorized into two main categories, i.e., on-lattice and off-lattice models, based on whether the cells are bound to occupy fixed positions defined by a lattice or if they can freely move through the domain. Taken from (Metzcar et al., 2019).

written as PDEs where discrete agents act as sources and sinks (Byrne, 2010).

2.3.3 Centre-based models (CBMs)

Cell morphology and biological behaviour

In CBMs, cells are assumed to be sufficiently represented by their centre of mass and a simplified geometry, such as a sphere (Macnamara, 2021; Metzcar et al., 2019). Although previous works modelled cell shape by assuming that cells behave as ellipsoids and modelling anisotropy and polarity (Palsson and Othmer, 2000; Rey and García-Aznar, 2013; Menzel and Ohta, 2012), the simplest CBMs do not resolve cell shape and, instead, consider that cells behave as isotropic viscoelastic spherical particles which can undergo small deformations (Camley and Rappel, 2017). Furthermore, CBMs simulate cells as autonomous agents that follow rules that define their biological and physical behaviour, such as cell cycling, cell death and mechanics, as shown in Fig 2.4. In this work, it is considered that CBMs are hybrid models that include continuum representations of chemical substances present in the extracellular microenvironment, as explained in Section 2.3.2. Examples of software platforms that enable the creation of hybrid CBMs models include *Biocellion* (Kang et al., 2014), *BioDynaMo* (Breitwieser et al., 2022), *CellSys* (Hoehme and Drasdo, 2010), *Chaste* (Mirams et al., 2013), *Gell* (Du et al., 2022), *iDynoMiCs* (Lardon et al., 2011) *MecaGen* (Delile et al., 2017) and *PhysiCell* (Ghaffarizadeh et al., 2018). Nevertheless, discrete approaches that consider only cell behaviour and mechanics have also been used in computational biology and are available through frameworks such as *CBMOS* (Mathias et al., 2022), *FLAME* (Richmond et al., 2010) and *yalla* (Germann et al., 2019).

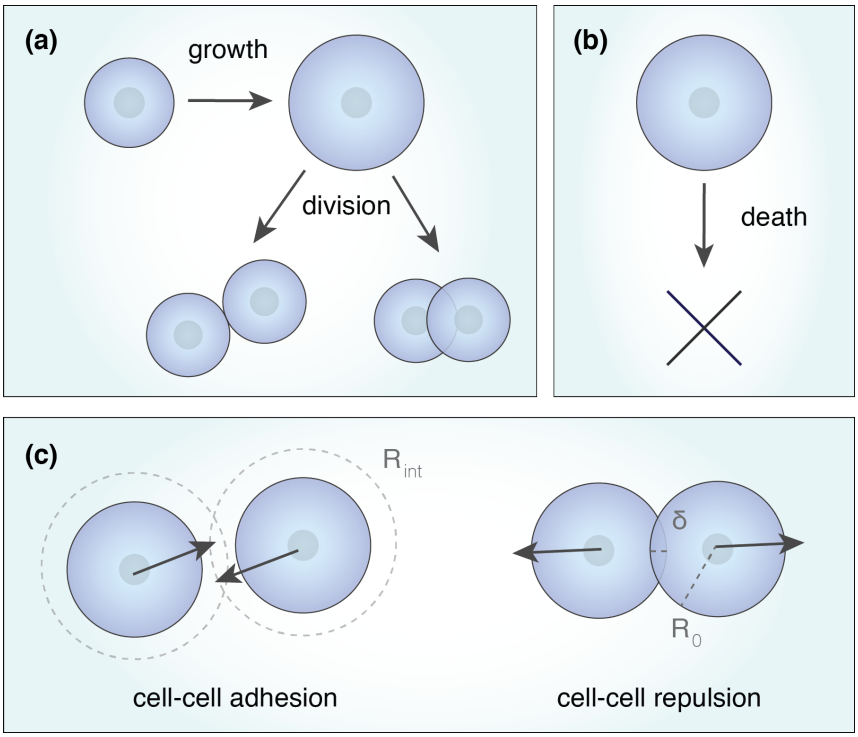

Figure 2.5: **Biological and physical rules in centre-based models.** Centre-based models assume that cells are represented by their middle point and a spherical geometry and follow rules that define their life cycle, death and mechanics. (a) Representation of cell growth and subsequent cell division. Cells may automatically originate two spherical daughter cells that repel each other. Alternatively, a dumbbell shape can be modelled to simulate the cell morphology during division. (b) Cell death can be simulated through agent removal from the simulation domain, which may include a phase where size decreases. (c) Cell-cell mechanics are computed based on pairwise interactions that depend on the distance between the cells' centre positions. Adapted from (Van Liedekerke et al., 2018).

Previous works based on CBMs considered cell growth during division through an increase in cellular volume (Drasdo, 2007; Ghaffarizadeh et al., 2018; González-Valverde and García-Aznar, 2017), as represented in Fig 2.5(a). During cell division, cells duplicate their DNA and synthesise macromolecules required to produce two viable cells (McIntosh et al., 2012), doubling their size approximately. Subsequently, an actin ring is formed at the cell's division plane, which contracts and generates two individual cells. These stages of development are generally modelled by introducing two new spherical agents in close proximity at the end of the cell cycle, which should be repelled by each other until equilibrium is reached (Macklin et al., 2012; Mathias et al., 2020). Alternatively, a more realistic representation of the deformations observed in cell shape during division can be simulated by approximating cell shape as a dumbbell that, subsequently, separates into two cells (Van Liedekerke et al., 2018; Drasdo, 2007).

Cells are represented as autonomous agents with an intrinsic cell cycle regulated by an internal clock that tracks and updates their state over time. In addition, the rates at which cell division occurs can be modelled as deterministic or stochastic processes (Ghaffarizadeh et al., 2018). When

considering a deterministic system, it is established that these events occur at fixed time intervals. In other words, when a cell doubling time is settled, all cells will undergo cell division once this time point has passed in their internal clock. Conversely, stochastic modelling is based on setting rates that dictate the probability of cell division (P_{div}) to occur at a given time interval Δt and is represented by Eq 2.1:

$$P_{div} = 1 - e^{r_{div}\Delta t} \approx r_{div}\Delta t \tag{2.1}$$

The cell cycle can be modelled at different complexity levels, with models ranging from a single transition rule that defines cell division to multiple states that describe the different cell cycle phases (Macklin et al., 2012; Pleyer and Fleck, 2023; Kempf et al., 2010). Moreover, additional rules can be defined to simulate the effect of the surrounding microenvironment on cell proliferation. For instance, cells may enter cycle arrest due to contact inhibition, i.e., when the forces exerted on them by other cells inhibit proliferation (Baker and Simpson, 2010; Pavel et al., 2018). Likewise, due to the hybrid nature of these frameworks, the cell cycle may be linked to the substances represented as continuum concentration fields (de Montigny et al., 2021; Macklin et al., 2012; Hyun and Macklin, 2013). Accordingly, cells located in regions where nutrient density is low may become quiescent due to the limited access to growth-enabling substances.

Cell death is simulated by agent removal from the simulation domain, through similar approaches as those used to model the cell cycle. In other words, deterministic or stochastic models can be used, and the microenvironment may also modulate this process. When considering cell death, both apoptosis, i.e., programmed cell death, and necrosis, which is induced by external factors, can be reproduced (Macklin et al., 2013; D'Arcy, 2019). Cell death can also be employed to simulate therapeutic agents through chemical substances or physical agents that can attach to cells and kill them (Ghaffarizadeh et al., 2018; Ozik et al., 2018).

Physical interactions and motility

In CBMs, the cells' centre of mass can be updated based on an equation of motion (Van Liedekerke et al., 2018). Generally, three factors are taken into account in this equation: the friction imposed by the microenvironment, the mechanical interaction forces between cells (i.e., adhesion and repulsion forces) and the active migration forces exerted by the cells (Camley and Rappel, 2017; Macnamara, 2021). Moreover, because the extracellular microenvironment has a low Reynolds number, it is generally assumed that inertia can be neglected and friction dominates this system (Purcell, 1977; Odell et al., 1981). Accordingly, the equation of motion for a spherical cell i can be written as Eq 2.2:

$$\mathbf{0} \approx \mathbf{F}^i_{drag} + \mathbf{F}^i_{loc} + \sum_{j \in \mathcal{N}(i)} \mathbf{F}^{ij}_{cci} \tag{2.2}$$

where $\mathbf{F}^i_{drag}$ represents dissipative forces caused by friction (e.g., fluid drag and matrix adhesion),

$\mathbf{F}^i_{loc}$ represents the active locomotive forces generated by cells $\mathbf{F}^{ij}_{cci}$ represents the cell-cell interactions between the cell i and a neighbouring cell, j. It is a common and reasonable assumption to model drag forces as $\mathbf{F}^i_{drag} = -\frac{1}{\mu}\mathbf{v}^i$, meaning that these forces are proportional to the cells' velocity, where μ defines the mechanical properties of the ECM and the resistance to cell movement (Camley and Rappel, 2017; Zaman et al., 2005). Nonetheless, other approaches include, for instance, explicitly modelling the frictional forces imposed by not only the ECM but also other cells (Van Liedekerke et al., 2015).

Cells exert forces on the surrounding ECM to migrate by creating protrusions in a given direction, forming adhesions with the matrix and contracting the cytoskeleton to move in a preferential direction (Merino-Casallo et al., 2022a). Henceforth, these active migration forces are usually represented in mathematical models by their magnitude, which correlates to cell speed, and direction (Rey and García-Aznar, 2013). Migration forces can be modelled to represent random or directed movement. When cells do not have a preferred direction movement, cell motion can be described as a random walk (Codling et al., 2008). Conversely, in the presence of chemical factors known as chemoattractants, i.e., which stimulate cell migration towards these substances, the direction of preferred movement aligns with the chemical gradients (Ghaffarizadeh et al., 2018). Moreover, models of cell migration can assume that cells move at constant speed and that only the direction of movement changes. Nonetheless, other approaches suggest that cell speed may follow a distribution, for example it may fluctuate around a mean value, thus capturing the heterogeneous nature of cell movement better (Campos et al., 2010).

Lastly, the physical interactions between two neighbouring cells can be depicted through pairwise potential functions which consider both cell-cell adhesion and repulsion as a function of the overlap between the two cells (Mathias et al., 2022; Rey and García-Aznar, 2013). Neighbouring cells are known to create adhesion bonds due to cadherin and caterin molecules when they are nearby (Ramis-Conde et al., 2008; Macnamara, 2021). Conversely, repulsion forces are also exerted when cells are compressed together since their internal cytoskeleton cannot deform extensively (González-Valverde et al., 2016). Consequently, it is well-established in CBMs that three regions define how two cells interact based on their overlap, commonly written as δ, a resting length R_0, usually considered to be equal to the cells' diameter, and an established interaction radius, R_{int}, as shown in Fig 2.5C (Mathias et al., 2020; Rey and García-Aznar, 2013). Specifically, when two cells overlap, i.e., the distance between their positions is smaller than the resting length, repulsive forces are dominant to assure that this overlap is eliminated. Conversely, adhesion forces play a more dominant role at distances larger than R_0 but smaller than R_{int}, making the cells adhere and stabilize at the resting length. For distances larger than the radius of interaction, the interaction forces become negligible. Different models may be employed to model cell-cell interactions, including linear springs (Drasdo and Loeffler, 2001), generalized linear springs (Mirams et al., 2013), Hertzian contacts (Drasdo and Höhme, 2005), Johnson–Kendal–Roberts potentials (Basan et al., 2011), Morse potentials (Rey and García-Aznar, 2013), piecewise quadratic forces (Ghaffarizadeh et al., 2018) and cubic forces (Delile et al., 2017).

Extracellular microenvironment

CBMs can be extended to account for representations of not only the cells but also the extracellular microenvironment. Like many ABM-based approaches, CBMs are usually coupled with PDEs to simulate the substances in the microenvironment. These reaction-diffusion equations can take into account substance diffusion and decay as well as consumption and uptake by bulk sources/sinks and cells (Ghaffarizadeh et al., 2018). Changes in the concentration of a substance i can be modelled in accordance with a reaction-diffusion equation:

$$\frac{\partial \rho_i}{\partial t} = D_i \nabla^2 \rho_i - \lambda_i \rho_i + S(\rho^* - \rho) - U\rho_i + \sum_{cells,k} \delta(\mathbf{x} - \mathbf{x}_k) W_k [S_k(\rho_k^* - \rho) - U_{k,i}\rho_i] \text{ in } \Omega \tag{2.3}$$

with Dirichlet boundary conditions in $\partial\Omega$. Here, ρ_i represents the substance concentration, D_i is the diffusion coefficient, λ_i is the decay rate, S is the bulk supply rate, ρ^* represents the saturation density at which secretion stops and U is the bulk uptake. In addition, for a cell k in the domain, $\delta(\mathbf{x})$ is the Dirac function, $\mathbf{x}_k$ is the cell position, W_k represents the cell volume, S_k is the cell secretion rate and $U_{k,i}$ is the cell consumption rate. For instance, this approach has been used to simulate the effect of hypoxia (Rocha et al., 2021; Macklin et al., 2012), nutrient-depletion (Shan et al., 2018; Jagiella et al., 2016) and therapeutic agents (Ghaffarizadeh et al., 2018; Jafari Nivlouei et al., 2022) on tumour development.

As previously stated, the local concentration can be evaluated by the agents in a simulation and considered to make decisions on cell behaviour (e.g., arrest proliferation, induce cell death and change the direction of migration). Furthermore, ABM-based approaches enable the simulation of biological behaviour at different scales, including through the definition of subcellular metabolic and signalling pathways (Rejniak and Anderson, 2011; Jafari Nivlouei et al., 2022). This approach provides a detailed microscopic description of the processes that regulate cell behaviour and provides information that can be useful when it is relevant to study the effect of specific molecules (Ford Versypt, 2021). Intracellular models can be implemented through differential equations (Somogyi et al., 2015), Boolean models (Letort et al., 2019) and neural networks (Gerlee, 2013), for instance.

Lastly, CBMs can account for representations of the ECM and its mechanical properties. In their equation of motion (see Eq 2.2), CBMs already simulate the effect of drag forces imposed by the ECM. Nonetheless, it is possible to incorporate more realistic descriptions of this component of the extracellular microenvironment (Macnamara, 2021; Senthilkumar et al., 2022). The ECM is composed by a meshwork of fibres that define its mechanical properties (Pathak and Kumar, 2011). Starting with a straightforward approach, one way to take into account spatial heterogeneity is to simulate the ECM substance whose concentration can be sampled by cellular agents to update their equation of motion based on the local viscosity (Gonçalves and Garcia-Aznar, 2021). Alternatively, the CBM model can be coupled to a tunable continuum solid material to account for the tissue passive mechanics (González-Valverde and García-Aznar, 2018, 2017). Finally, ECM

fibres can be modelled as agents (spheres or cylinders) that interact with cells, which is the most realistic yet computationally expensive approach (Letort et al., 2019; Macnamara et al., 2020).

2.4 Summary

This chapter provided an overview of the biological processes that will be covered in this dissertation and the computational frameworks used to model them. The general objective of this dissertation is to simulate cancer systems using computational models. Accordingly, we discussed the mechanisms that lead to tumour development and presented a theoretical framework that unifies different malignant characteristics observed in distinct cancer types. In addition, we explored some of these characteristics in more detail because they will be studied in the following chapters. Specifically, Chapter 3 deals with migration and metastasis, Chapter 4 explores tumour growth and metabolic reprogramming and, lastly, Chapter 5 aims to reproduce the emergence of specific spatial patterns that arise due to abnormal differentiation.

Furthermore, this chapter contextualized the current trends in computational oncology, from model formulation to specific modelling frameworks. In this dissertation, we aimed to simulate cancer systems from a bottom-up perspective, starting with individual cell behaviour and progressing to collective trends. Therefore, we chose to model our systems through CBMs, which were discussed in detail in this chapter. We based this choice on the need to simulate each cell as an individual agent and the advantage of being able to study cell-cell and cell-matrix interactions through force-based equations. We started this dissertation with a general CBM-based framework and extended it to account for more complex phenomena. Accordingly, in Chapters 3-5, we discuss new and specific biological behaviours, how they were added to the model and what we can learn from the results obtained.

CHAPTER 3

Single-cell motility

Contents

3.1 Introduction 26
3.2 Model optimization and data integration: an overview 29
3.2.1 Microfluidic devices and migration assays 30
3.2.2 Model calibration routines and data-driven models 31
3.3 Modelling 3D cell motility in collagen matrices of different densities 32
3.3.1 Study outline 32
3.3.2 Methodology 33
3.3.3 Results 37
3.3.4 Conclusions 38
3.4 An improved data-driven model of chemically-regulated motility 41
3.4.1 Study outline 41
3.4.2 Methodology 42
3.4.3 Results 49
3.4.4 Conclusions 56
3.5 Summary 59

This chapter is based on:

Inês Godinho Gonçalves and José Manuel García-Aznar. **"Extracellular matrix density regulates the formation of tumour spheroids through cell migration."** PLoS computational biology 17.2 (2021): e1008764.

Inês Godinho Gonçalves, Nieves Movilla, Carlos Borau and José Manuel García-Aznar. **"A novel integrated experimental and computational approach to unravel fibroblast motility in response to chemical gradients in 3D environments."** Integrative Biology (2023)

3.1 Introduction

Some of the most fundamental processes in life rely on the complex process of cell migration (Vicente-Manzanares and Horwitz, 2011; Ladoux and Mège, 2017). For instance, unicellular organisms rely on migration to move away from hazardous conditions and into nutrient-rich surroundings. Additionally, cell migration is involved in multiple processes that ensure the normal growth and maintenance of multicellular organisms, including morphogenetic events, tissue repair and the immune response (te Boekhorst et al., 2016). Nonetheless, unregulated cell motility can also lead to metastasis (Krakhmal et al., 2015), which significantly increases cancer mortality, and play a role in chronic inflammation (Liu et al., 2021). Cell migration can be performed individually or collectively (Friedl and Wolf, 2010; Buttenschön and Edelstein-Keshet, 2020). Although collective migration tends to be prevalent in physiological phenomena such as tissue regeneration and wound healing (Mayor and Etienne-Manneville, 2016), it requires an extended understanding of the chemical and mechanical interaction between cells that maintain tissue organization (Trepat et al., 2012; Canel et al., 2013; Yang et al., 2019). Therefore, this chapter focuses on characterizing single-cell motility.

Individual migration phenotypes are usually classified as amoeboid or mesenchymal (Bear and Haugh, 2014; Talkenberger et al., 2017). Fig 3.1A shows a comparison between these two migration modes. On the one hand, amoeboid migration is characterized by rapid cell movement and constant changes in cell shape to move through the paths of least resistance in the microenvironment (Lämmermann and Sixt, 2009). The rapid movement of the cells is accompanied and facilitated by weak adhesion interactions with the substrate that allow cells to squeeze through pores and small structures (Lämmermann et al., 2008). Consequently, highly motile cells such as those present in the inflammatory response (e.g., neutrophils and lymphocytes) generally acquire amoeboid migration patterns. On the other hand, mesenchymal migration relies on cell-matrix adhesions through which cells exert forces on the surrounding environment to move (Merino-Casallo et al., 2022a). Accordingly, it is a slower mechanism, usually observed in cells like fibroblasts that rely on cell adhesions to sense the mechanical and chemical properties of the surrounding ECM (Bear and Haugh, 2014). Moreover, these cells may exhibit the ability to degrade the ECM and generate new paths to pass through (Yamada and Sixt, 2019).

Cancer cells are regularly associated with mesenchymal migration because they undergo the EMT, through which cell-cell and cell-matrix adhesions are lost, disrupting the structure of epithelial tissues and leading to the emergence of mesenchymal-like migration patterns (Kalluri and Weinberg, 2009; Yang et al., 2020). However, amoeboid features have also been observed in cancer cells (Krakhmal et al., 2015), for instance, when protease inhibitors are present and cells become unable to degrade the surrounding matrix (Friedl and Wolf, 2003; Wolf et al., 2003; Orgaz et al., 2014). Furthermore, recent studies have shown that amoeboid and mesenchymal motility patterns are not mutually exclusive (Graziani et al., 2022). According to these studies, cancer cells present the phenotype plasticity required to shift between these two states (Talkenberger et al., 2017; Paňková et al., 2010). Therefore, more studies are needed to understand this behaviour and

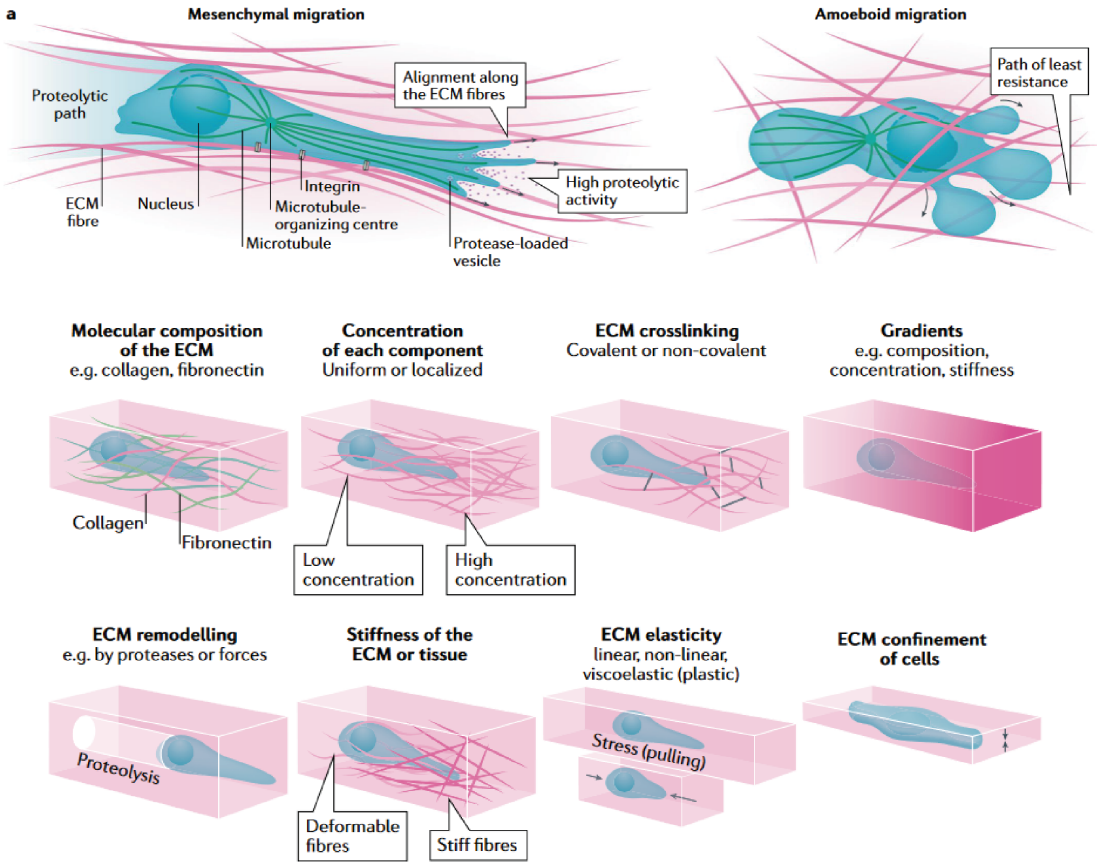

Figure 3.1: **Cell motility mechanisms in 3D fibrous matrices.** (A) Comparison between the mesenchymal and amoeboid phenotypes, the two major migration modes in individual cells. Mesenchymal migration is characterized by "path-generating" patterns because cells degrade and change their environment to move. Conversely, cells that adopt a amoeboid phenotype tend to follow "path-finding" patterns, where they squeeze through the matrix pores and try to move along the path of least resistance. (B) Main ECM features that influence and guide cell motility in 3D. Both mechanical (e.g., matrix stiffness, elasticity and confinement) and chemical (e.g., molecular composition, chemical gradients) properties can have an impact on cell migration. In addition, cells can modify the structure of the ECM, altering its properties. Adapted from (Yamada and Sixt, 2019).

advance the knowledge of how the external factors may regulate it.

Environmental cues have been shown to modulate cell motility as cells are able to sense and respond to the properties of their surroundings. For example, cells can follow gradients of soluble or surface-attached chemoattractants (e.g., chemokines and growth factors) through chemotaxis and haptotaxis, respectively. Furthermore, cells respond to the mechanical properties of the extracellular microenvironment (Discher et al., 2005; Janmey and McCulloch, 2007). Matrix stiffness, which characterizes the matrix's resistance to deformation in response to forces, has been extensively studied in 2D (Ehrbar et al., 2011; Paszek et al., 2005; Lo et al., 2000; Pathak and Kumar, 2011). Nonetheless, these findings may not apply to 3D environments, which are more similar to complex physiological conditions (Yamada and Sixt, 2019). The complexity of 3D cultures makes it increasingly difficult to disassociate the effects induced by matrix stiffness from those produced by the matrix architecture. In 3D, stiffness arises from structural changes that also affect the matrix architecture (Pathak and Kumar, 2011). In other words, matrices with a higher fibre density may be stiffer, but they also present smaller pore sizes. Therefore, cells seeded in these matrices face high confinement levels and become unable to squeeze through the matrix to continue moving (Charras and Sahai, 2014).

In the last two decades, advanced cell culture tools such as microfluidic devices have been developed to study cell motility in 3D configurations (Mehta et al., 2022; van Duinen et al., 2015), as opposed to more conventional 2D approaches (Kramer et al., 2013). Microfluidic devices are engineering tools that enable the study and manipulation of fluids at the submillimetre length scale and, at the same time, the acquisition of high-quality image data (Sackmann et al., 2014; Moreno-Arotzena et al., 2014). Fig 3.2 shows an example of a microfluidic device used to study cell motility. Despite these advances, there are still disadvantages to experimental research, such as the increased costs and long time scales (Bekisz and Geris, 2020). Accordingly, to facilitate the study of these biological questions, computational modelling has been used to create models and simulation frameworks that can replicate experimental settings while overcoming the aforementioned disadvantages (Merino-Casallo et al., 2022a; Rangarajan and Zaman, 2008; Shatkin et al., 2020). Furthermore, the high-quality imaging data generated in microfluidic assays can be considered in the design and implementation of these computational approaches to enhance their biological relevance.

Taking this into account, this chapter aims to explore how computational models can be used to reproduce experimental data from migration assays and make predictions about the mechanisms that guide cell motility. To achieve this, Section 3.2 provides an overview of the experimental configurations used to perform migration assays in 3D and discusses model calibration and optimization routines. Subsequently, two applications are discussed. Section 3.3 presents an application which focused on implementing a computational framework that accounted for the presence of the ECM and its effect on single-cell motility. This work was based on a previously published study and qualitative model calibration routines were performed. The second application, discussed in Section 3.4, intended to extend the computational model developed in Section 3.3 and integrate it into an automatic model calibration pipeline to predict which model input parameters produce

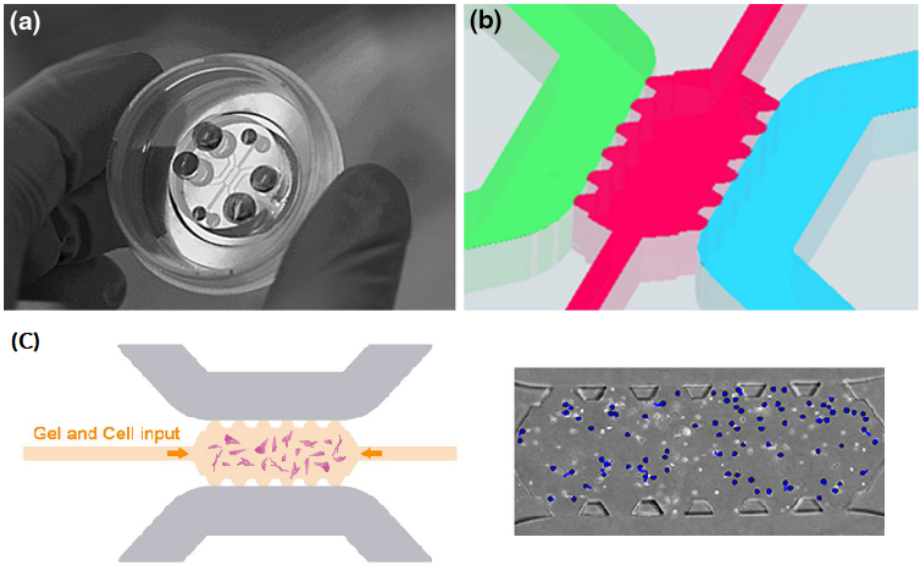

Figure 3.2: **Microfluidic devices.** (a) Microdluidic device in a 35 mm glass-bottom petri dish. (b) Geometry of the microfluidic device. The hydrogel is kept confined within the central channel (pink). In direct contact to the gel, two main media channels (green and blue) ensure hydration and diffusion through the hydrogel. (c) Migration assays can be performed by seeding cells in the hydrogel chamber and tracking individual cells over time. Adapted from Moreno-Arotzena et al. (2015) and Plou et al. (2018).

the best similarity between experimental and computational data. The data used in this study was obtained in original experiments that measured the chemotactic response of fibroblasts when exposed to the presence of growth factors. Thus, the model was properly extended to reproduce the effect of chemical substances on cell migration.

3.2 Model optimization and data integration: an overview

Mechanistic models are characterized by a large number of parameters that require validation (Hasenauer et al., 2015). Furthermore, ABMs can be particularly cumbersome to characterize given that they are determined by the interactions ocurring between multiple agents that often follow stochastic rules (Gonzalez-de Aledo et al., 2018). As such, parameter space exploration is crucial to characterize how models behave. In addition, it is important to take biological data into account to define these parameter spaces, to assure that model behaviour is as realistic as possible (Banga, 2008). Experimental data may also be used to optimize models and identify which model mechanisms were the most relevant to reproduce the observed biological behaviour (Stéphanou et al., 2020; Lima et al., 2021).

In this section, we describe how sophisticated experimental setups based on the use of microfluidic devices and time-lapse imaging can be used to characterize cell motility in 3D environments. Furthermore, we introduce some methods, such as parameter space exploration and model

optimization frameworks, focusing on advanced automatic routines based on Bayesian statistics.

3.2.1 Microfluidic devices and migration assays

In conventional 2D cell cultures, cells are grown over flat substrates, which is not realistic when compared with physiological conditions where cells can be found embedded in 3D environments (van Duinen et al., 2015). Taking into account that the nature and composition of the ECM are quite difficult to replicate *in vitro*, experimental studies conducted with 3D scaffolds have been helpful to improve the understanding of how the ECM regulates 3D cell migration (Nyga et al., 2011; Moreno-Arotzena et al., 2015; Movilla et al., 2018, 2019). In particular, hydrogels mainly composed of type I collagen, the most abundant substance of the natural ECM components (Frantz et al., 2010), are commonly used in experimental settings. Among other benefits, collagen matrices offer versatility and enable the possibility to produce matrices with different mechanical properties based on their composition and preparation procedures (Sapudom and Pompe, 2018). For example, collagen concentrations can be altered or changes in pH can be implemented to modify the different matrix properties (Roeder et al., 2002).

Polymeric matrices and scaffolds can be integrated in microfluidic devices to add a new complexity level to these simplified representations of the ECM through the creation of complex and dynamic environments (Li et al., 2012). Specifically, the high level of control over the microfluidic system's properties allows researchers to design experiments under different conditions by, for example, changing the density of the collagen matrix inside the device or introducing chemical substances in the devices (Plou et al., 2018; Pérez-Rodríguez et al., 2018; Moreno-Arotzena et al., 2015). These features make microfluidic devices a convenient tool to conduct migration assays with time-lapse microscopy approaches (Mehta et al., 2022). In these studies, the position of individual cells is tracked over time, resulting in a time series dataset that can be employed to quantify migration metrics such as the cells' persistence time and directionality (Pijuan et al., 2019; Pérez-Rodríguez et al., 2018).

Apart from cancer cells, several migration studies have been performed with fibroblasts since they are one of the most well-recognized models of mesenchymal migration (Doyle et al., 2021). Fibroblasts are the most abundant cells of the connective tissue and are known to synthesize and organize proteins such as collagen that form the ECM (Petrie and Yamada, 2015; Tschumperlin, 2013). Accordingly, fibroblasts play a relevant role in wound repair (Li and Wang, 2011) and tissue remodelling (Jacob et al., 2012), both of which require cells to migrate through the ECM to reach the areas where they will act. Once fibroblasts have reached the injured tissue, they can promote wound repair by reorganizing the ECM. Moreover, fibroblasts are associated with cancer cells during tumour development and they play a role on matrix remodelling during metastasis to promote invasiveness (Kalluri and Zeisberg, 2006; Kalluri, 2016).

Given the relevance of fibroblasts as a model of mesenchymal migration, it is important to study the mechanisms that define fibroblast motility and the factors that may regulate them (Bear and Haugh, 2014). One of the first aspects that modulate fibroblast activity during development and tissue repair is the presence of chemical substances that attract these cells, guiding them

through chemical gradients in a process known as chemotaxis (Martins et al., 2012). Growth factors such as PDGF-BB have been shown to act as chemoattractants in these processes (Heldin and Westermark, 1999; Deuel et al., 1991). Following an injury, some of the cells responsible for the inflammatory response release these substances to direct fibroblasts, which recognize these signals and orient themselves towards the direction of higher growth factor concentration (P., 2013). In addition, cancer cells have been shown to secrete PDGF-BB to recruit and activate fibroblasts (Bronzert et al., 1987).

Microfluidic devices can be used to study chemically-regulated cell motility by inducing chemical gradients in the microfluidic chambers (Moreno-Arotzena et al., 2015; Li et al., 2012). Taking the device shown in Fig 3.2C as an example, it can be seen that the it is composed of a central chamber containing an hydrogel matrix and two adjacent channels where cell media and other substances can be introduced. Therefore, a chemotactic gradient may be induced by introducing a substance (e.g., a chemoattractant such as PDGF-BB) in one of the chambers (Moreno-Arotzena et al., 2014; Cóndor et al., 2018). It is expected that this substance will diffuse through the central chamber, reaching the cells that are seeded in it. Subsequently, the migration patterns of these cells may be evaluated and compared to experiments where these chemical factors were not present to assess the chemotactic effect of the substance used in the study (Moreno-Arotzena et al., 2015).

3.2.2 Model calibration routines and data-driven models

Computational models can be seen as simulators that produce output data from a set of input parameters. Some input parameters may be known from the literature, or at least expected to lie inside a known interval. For example, cell doubling times may be available in the literature and they can be measured experimentally. Nonetheless, other parameters need to be inferred, which can be a difficult task (Fernández Slezak et al., 2010; Jørgensen et al., 2022). In addition, different model parameters can present distinct effects on the model outputs depending on their sensitivity (Fernández Slezak et al., 2010). In other words, small changes in the input values of a sensitive parameter will produce significant changes on model results. On the contrary, less sensitive parameters will have smaller effects on the simulations and allow for more flexible input parameter values. Accordingly, parameter spaces must be explored to quantify parameter uncertainty (Renardy et al., 2018).

At the most basic level, parameter exploration routines can be performed manually or programmatically (Read et al., 2018). Manual calibration implies that a user should define a parameter set, change the model inputs, run the simulation and process the results in a procedural and iterative manner. These tasks are often repetitive and, once defined, they do not actually require user actions, though. Therefore, scripts can be written to automatically perform them in a standardized and efficient manner (Newland et al., 2018; Read et al., 2018). Grid search, for example, is one of the most straightforward automatic calibration routines (Bergstra and Bengio, 2012). This search workflow starts with the definition of a search space, which defines the range of the parameter values to be studied. Subsequently, this algorithm systematically samples every point within the grid space, providing a comprehensive overview of the model's response. Alternatively, random

search algorithms can be employed to randomly sample a limited number of points within this space to reduce computational costs (Bergstra and Bengio, 2012).

Grid and random search methods require exhaustive parameter exploration. Consequently, they scale poorly when dealing with problems with several parameters and computationally expensive models. In contrast, new algorithms such as Bayesian Optimization (BO) predict the best parameter values for a given task without exhaustively searching the parameter space (Lei et al., 2021). BO is a methodology based on Bayes' theorem, commonly used to optimize functions that can be treated as a black-box, i.e., when the algorithm only needs to know the function inputs and outputs (Merino-Casallo et al., 2018; Mascheroni et al., 2021; Shahriari et al., 2016). At a basic level, this methodology takes into account some target data (e.g., data from *in vitro* experiments) and knowledge from previous model iterations to adequately choose the next set of parameters to be tested. Accordingly, BO significantly reduces the number of iterations needed to fit the model to experimental data (Shahriari et al., 2016).

BO can be regarded as a general optimization algorithm. In optimization problems, an *objective function* is defined to measure the quality of the solution (Banga, 2008). Subsequently, model parameters are varied according to the established model constraints and the objective function is evaluated. Assuming that BO is applied to optimize a computational model with experimental data, the objective of this algorithm is to maximize the similarity between the empirical and simulated datasets (Lei et al., 2021). Therefore, this problem can be defined by Eq 3.1 to find a set of optimal parameters (x) that maximize the function $f(x)$ in a given parameter space $\mathcal{X}$.

$$x^* = \underset{x \in \mathcal{X}}{\operatorname{argmax}} f(x) \tag{3.1}$$

The similarity between two datasets can be measured through metrics such as the Bhattacharyya coefficient (BC), which quantifies the overlap between two discrete distributions (Comaniciu et al., 2000). Conversely, distance metrics (e.g., the Kolmogorov–Smirnov statistic) can be used to compare the experimental and computational datasets if the global objective is changed to be a minimization problem instead, so that the differences between the measured and generated data are minimal (Jørgensen et al., 2022; Read et al., 2018; Lei et al., 2021).

3.3 Modelling 3D cell motility in collagen matrices of different densities

3.3.1 Study outline

The study presented in this section serves as the foundation for the computational models presented in this dissertation. Therefore, here we focused on simulating and characterizing what we consider to be the most basic complexity level to be addressed in this work, single-cell motility. However, we implemented a computational model that can be extended to incorporate the biological behaviours to be presented and discussed in Chapters 4 and 5. Furthermore, we strived

to assure that our model was biologically relevant and valid by replicating experimental results previously published by Plou et al. (2018). The premise of this previous work was to use microfluidic devices to seed individual cells of a non-small cell lung cancer (NSCLC) cell line in collagen matrices of varying concentrations and assess how collagen density modulated cell behaviour. Specifically, Plou et al. (2018) performed migration assays that enabled the study of single-cell motility by tracking individual cell trajectories for 24 hours and computing metrics such as the distance travelled by cells and their velocities.

Overall, this previous experimental study showed that cell motility was hindered in matrices of high collagen concentrations. In fact, as the collagen density increased, cells encountered more difficulties moving through the ECM due to steric hindrance. Accordingly, cells became increasingly more confined, leading to shorter travel distances and lower cell velocity values. Our goal in this study was to qualitatively replicate these general trends and results. Accordingly, we based our model design process on reproducing the experimental setup described by Plou et al. (2018) and using the published data to calibrate our model. We built our model as an extension to *PhysiCell* (Ghaffarizadeh et al., 2018), an open-source modelling framework that simulates cells through a CBM coupled to a continuum description of the microenvironment. We highlight that, although the results of previous studies are at the core of our model, we sought to build a flexible framework that can be extended to other results and scenarios.

3.3.2 Methodology

Modelling framework

The first step of the implementation process consisted of choosing an adequate modelling framework in concordance with our main objective. Since we considered that mechanical interactions should be regarded at the cellular level, we established that a discrete model would suit the problem better than a continuum approach. Discrete models are known to have an increased level of complexity, usually associated with higher computational costs (Jeon et al., 2010). Hence, researchers developed hybrid continuum-discrete models to reduce computational costs by simulating the extracellular environment as a continuum while still accounting for cell-level information. Recently, *PhysiCell* (Ghaffarizadeh et al., 2018) was presented as an open-source hybrid 3D cell simulator. For the continuum part of the model, an associated piece of software, BioFVM (Ghaffarizadeh et al., 2016), is used to solve reaction-diffusion equations. In particular, BioFVM can account for substance diffusion and decay for several user-defined substances, as well as the existence of both bulk and cell-centred sources and sinks (Ghaffarizadeh et al., 2016).

Based on *PhysiCell*'s advanced state of development, along with its flexibility to include new user-defined modules, we chose to develop our model as an extension to this framework. To do so, we relied on built-in functions that accounted for tracking the cell volume, cell proliferation and cell death, and we extended the features that model cell motility and cell mechanics. Moreover, we introduced a representation of the ECM, which the *PhysiCell* system did not include. Thus, collagen was incorporated into the model as a non-diffusing substance defined by its density.

Table 3.1: **Reference parameter values for the single-cell motility model.**

Symbol	Parameter	Value	Unit	References
Δt_{mech}	Mechanics time step	0.1	min	Ghaffarizadeh et al. (2018)
$\rho^0_{Collagen}$	Collagen Concentration	[2.5, 4.0, 6.0]	mg/mL	Plou et al. (2018)
μ	Drag Coefficient	[7.96, 18.42, 39.15]	Pa · s	Valero et al. (2018)
R	Cell Radius	8.4	μm	Ghaffarizadeh et al. (2018)
$t_{persist}$	Cell persistence time	20.0	min	Ghaffarizadeh et al. (2018)

Subsequently, this value was used to compute cell-matrix interactions, as further explained below. For this first approach, we assumed that the collagen density was uniform throughout the domain and that cells did not remodel the matrix, so the collagen concentration did not change over time. More details on the parameters used to implement our model are presented in Table 3.1.

Introducing the mechanical properties of collagen matrices

PhysiCell updates cell positions by computing the effect of cell-cell forces acting on each cell, as represented in Fig 3.3. Accordingly, considering a cell i interacting with some neighbouring cells, $\mathcal{N}(i)$, we can consider the equilibrium equation taking the over-damped assumption:

$$\mathbf{0} \approx \sum_{j \in \mathcal{N}(i)} (\mathbf{F}^{ij}_{cca} + \mathbf{F}^{ij}_{ccr}) + \mathbf{F}^{i}_{drag} + \mathbf{F}^{i}_{loc} \tag{3.2}$$

where $\mathbf{F}^{ij}_{cca}$ and $\mathbf{F}^{ij}_{ccr}$ are cell-cell adhesive and repulsive forces, $\mathbf{F}^{i}_{drag}$ represents dissipative forces (such as fluid drag and matrix adhesion) and $\mathbf{F}^{i}_{loc}$ is the net locomotive force. In this work, cells were simulated as single agents (i.e., one cell per simulation), given that the results presented by (Plou et al., 2018) indicated that cells moved as individuals and did not interact with each other. Thus, the cell-cell interaction terms ($\mathbf{F}^{ij}_{cca}$ and $\mathbf{F}^{ij}_{ccr}$) were disregarded. Furthermore, we assume that cell-matrix forces can be described through Eq 3.3:

$$\mathbf{F}^{i}_{drag} = -\mu \mathbf{v}_i \tag{3.3}$$

where $\mathbf{v}_i$ is the cell velocity and μ represents the drag coefficient. The mechanical properties of the collagen matrices used by Plou et al. (2018) were characterized in a previous publication (Valero et al., 2018) through rheological analysis.

Valero et al. (2018) characterized the viscoelastic properties of collagen matrices using a stress-controlled rheometer and showed that the dynamic viscosity, which quantifies the material's resistance to flow, increased with collagen density. Hence, we expected cells to face increased difficulty when migrating in denser, and consequently more viscous, matrices. Taking this expectation into account, we considered that the drag coefficient, μ, could be estimated using this data. Specifically, for each collagen density, we defined μ as the mean value for the measured dynamic viscosity values, as presented in Table 3.1. The viscosity values were considered to be uniform

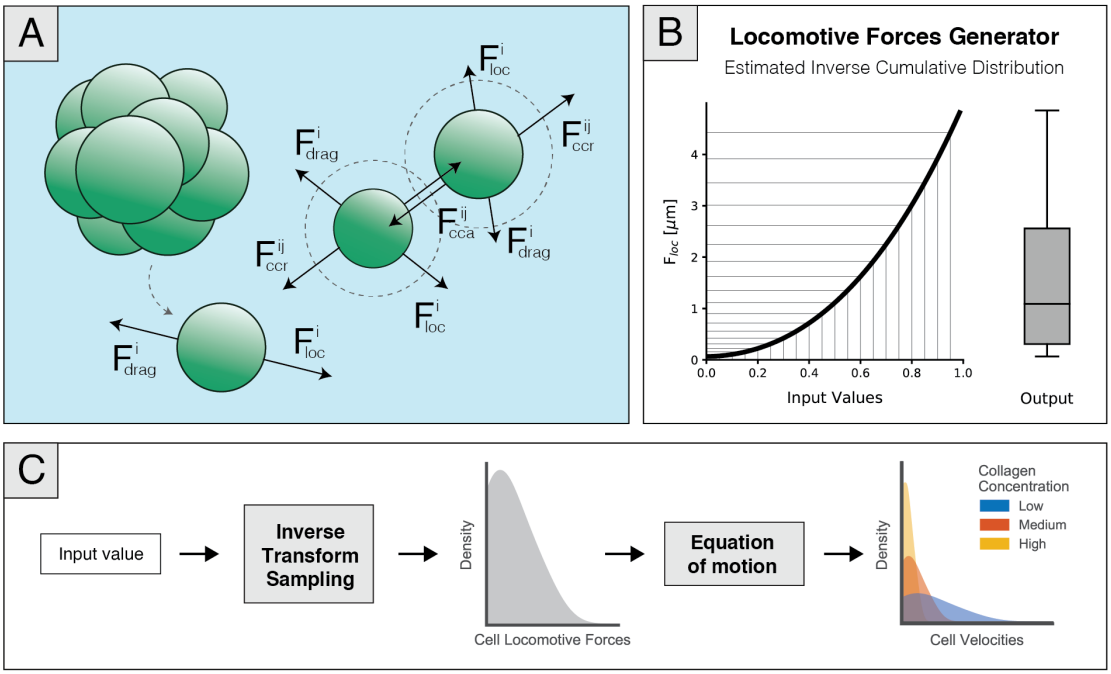

Figure 3.3: **Forces acting on the single-cell motility model.** (A) Force diagram showing the cell-cell and cell-matrix interactions present in *PhysiCell*. Although *PhysiCell* accounts for cell-cell adhesion and repulsion of cells inside an interaction radius, this simplified model of single-cell motility only considers cell-generated locomotive forces and drag forces imposed by the ECM. (B) Locomotive forces generator function modelled through an estimation of the inverse cumulative distribution of experimental cell velocities. When fed uniformly distributed random values between 0 and 1 (x-axis), the represented function produced a new set of values that followed the desired force distribution (y-axis), as shown by the representative boxplot showcased as the output of this function. (C) Implemented workflow to model cell-generated locomotive forces, considering no cell-cell interactions. At an average of 20 simulated minutes, cells were allowed to change their velocity both in magnitude and direction. Accordingly, at those time points, we generated a new force value through the inverse sampling method. The output of this function was subsequently incorporated into the equation of motion producing three different velocity distributions, each corresponding to a matrix density value.

throughout the domain since we assumed that, as a first approach, the collagen density was also uniform.

Defining cell-generated forces

The results obtained by Plou et al. (2018) showed that there was a diversity of behaviours in cell motility, as seen by the non-uniform distribution of cell velocities. Taking into account that cell-cell interactions were not considered in this work, the equation of motion (Eq 3.2) was simplified to $\mathbf{v}_i = \frac{1}{\mu}\mathbf{F}^i_{loc}$. Therefore, the velocity at which cells moved depended solely on the viscosity of the collagen matrix, μ, and the locomotive forces generated by the cells. Given that we assumed that this viscosity was uniform for each collagen concentration and given by a constant μ, it was apparent that the heterogeneity in behaviour came from the locomotive forces. Thus, we defined that the distribution of forces should resemble the velocity distribution characterized experimentally.

We considered that cells were able to change their velocity as a result of a change in the magnitude and direction of the active locomotive forces. This ocurred stochastically and multiple times during the simulations, according to the cells' persistence time, as given by Eq 3.4:

$$P(v_{update}) = \frac{\Delta t_{mech}}{t_{persist}} \tag{3.4}$$

where, $P(v_{update})$ is the probability of changing cell velocity, Δt_{mech} is the time step at which mechanical interactions were computed, and $t_{persist}$ is the cell's persistence time. The values assigned to these model parameters can be found in Table 3.1. Regarding the magnitude update, we needed to create a function that generated random values following the force distribution that, in turn, would produce velocity distributions similar to those measured in the experiments. We chose to solve this problem through the inverse transform sampling method, which is generally used to obtain random values from any probability distribution. In particular, the inverse cumulative distribution is computed and, by feeding this function uniformly random numbers between 0 and 1, it generates a range of values that follows the distribution we aimed to model, as represented in Fig 3.3B.

Henceforth, we created a generator function by estimating an inverse cumulative density function from the experimental cumulative density distribution, which ultimately resulted in Eq 3.5:

$$\mathrm{F}_{loc}(x) = 1.56x^3 + 3.27x^2 + 0.07x + 0.06 \tag{3.5}$$

It must be noted that, as stated previously, our objective was to capture the general tendencies seen experimentally and not to single-handily replicate the experimental data on a small scale. Hence, we chose to estimate a general curve based on the three available empirical cumulative cell velocity distributions presented in Plou et al. (2018) (for the different collagen concentrations), so that it could provide good results for all collagen densities, rather than fit the model to each curve. Subsequently, we iteratively changed its coefficients based on the obtained velocity distributions, so that the results resembled those seen experimentally. An extensive study on how these parameters

may influence our results can be found in Appendix B, Section B.1. Furthermore, this analysis also includes an overview of how the model may respond to other density distribution functions.

Finally, to compute the direction of these forces, it was assumed that the cells adopted a completely random walk. Therefore, every time the velocity of a cell i was updated, a unit vector that defines a random direction, $\mathbf{e}^i$, was computed. Afterwards, this value was multiplied by the magnitude value given by Eq 3.5, thus defining $\mathbf{F}^i_{loc}$:

$$\mathbf{F}^i_{loc} = \mathrm{F}_{loc}\mathbf{e}^i \tag{3.6}$$

Simulation workflow

Our simulations aimed to reproduce the experimental setup employed by (Plou et al., 2018) to track cell trajectories and cell speeds for a period of 24 hours. We considered a single cell for each simulation, taking into account that we did not consider cell-cell interactions, and we tracked its trajectory and recorded spatial data at the same time points as those considered in the experiments (i.e., every 20 minutes). Furthermore, we consider that cells cannot duplicate to meet the requirements of this study, namely, avoiding cell-cell interactions. Moreover, given that the expected lifetime of the NSCLC cell line surpasses 24 hours, cells were programmed not to enter cell death during the simulations.

Similar to what was done experimentally, we computed the mean cell velocity by taking the average value of cell velocities, calculated as the distance travelled between time points, divided by the recording time. Additionally, we computed effective cell velocities as the distance between the initial and final positions of the cell, divided by the time of the simulation. However, experimental quantification was previously conducted using 2D imaging techniques, through confocal microscopy, and these results were presented in the form of 2D plots. Accordingly, we followed the same data processing procedures to facilitate the comparison between the experimental and computational results. This simulation workflow, summarized in Fig 3.4, was used to run 80 replicates for each of the matrix density values, resulting in a similar number of data points to those obtained experimentally. Moreover, we manually calibrated the model by comparing its outputs with the qualitative results published by Plou et al. (2018).

3.3.3 Results

Increasing values of collagen density hinder individual cell migration

As previously stated, this work aimed to define cell-generated forces and to evaluate the model's ability to qualitatively describe the effect of matrix density on individual cell migration. In fact, our results show that cell trajectories become more contained as the collagen concentration increases, as represented in Fig 3.5. It is important to state that regarding individual cell trajectories, the experimental results show that cells appear to spread out in a horizontal direction. We stipulate that this behaviour is promoted by the alignment of the collagen fibres, which is a result of the procedures used to fill out the microfluidic chips and collagen polymerization and has been shown

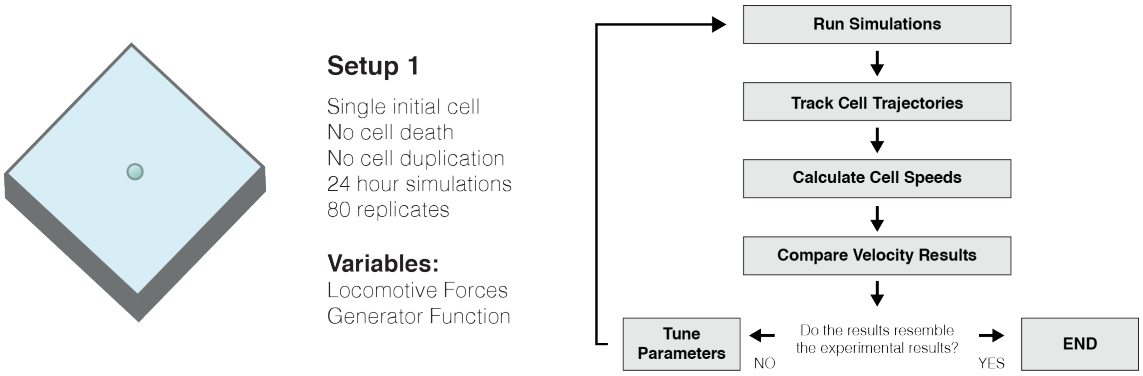

Figure 3.4: **Simulation and calibration workflow for the single-cell motility model.** Single cells were initialized and their trajectories were tracked and recorded. Mean and effective velocities were computed from the recorded spatial data and manual calibration was performed to find the model input parameters that produced the best results in comparison with the available qualitative empirical data.

to regulate cell migration by promoting more directed migration patterns (Del Amo et al., 2018). In contrast, simulated cells adopt more random motility patterns. However, this outcome is expected since we did not define a preferential direction of migration, as our focus was studying cell movement in general and not replicating individual cell trajectories. Nonetheless, considering that it has been shown that fibre alignment can affect cell motility (Del Amo et al., 2018; Fraley et al., 2015), the model could be easily extended to consider directionality through Eq 3.6 by defining a preferential direction, rather than a random vector. In terms of distance travelled, the computational and simulated results are similar. We note that although these results are presented through 2D plots, cells migrate in 3D space in both the simulated and experimental conditions.

Furthermore, our results show that both mean and effective cell velocities adopted smaller values as matrices present higher density values, as shown in Fig 3.5 and Table 3.2. Therefore, we concluded that our simulation results replicated the effect of high ECM density values on individual cell migration. Fig 3.5 shows that the model generally captures this non-uniform distribution of cell velocities quite well. As expected, the distribution does not apply to each of the density values specifically, since we used a general approximation, rather than a specific distribution for each matrix density value. However, we show that good results can be achieved without the need for an increased level of complexity.

3.3.4 Conclusions

In this study, we presented a centre-based model extension to simulate individual and collective cell behaviour in which the regulatory effect of matrix density on cell migration and tumour formation was introduced, using previously published experimental data (Plou et al., 2018), to calibrate the model. To achieve this, we extended *PhysiCell* to account for the presence of the ECM and we characterized its mechanical properties with previously published rheology data (Valero et al.,

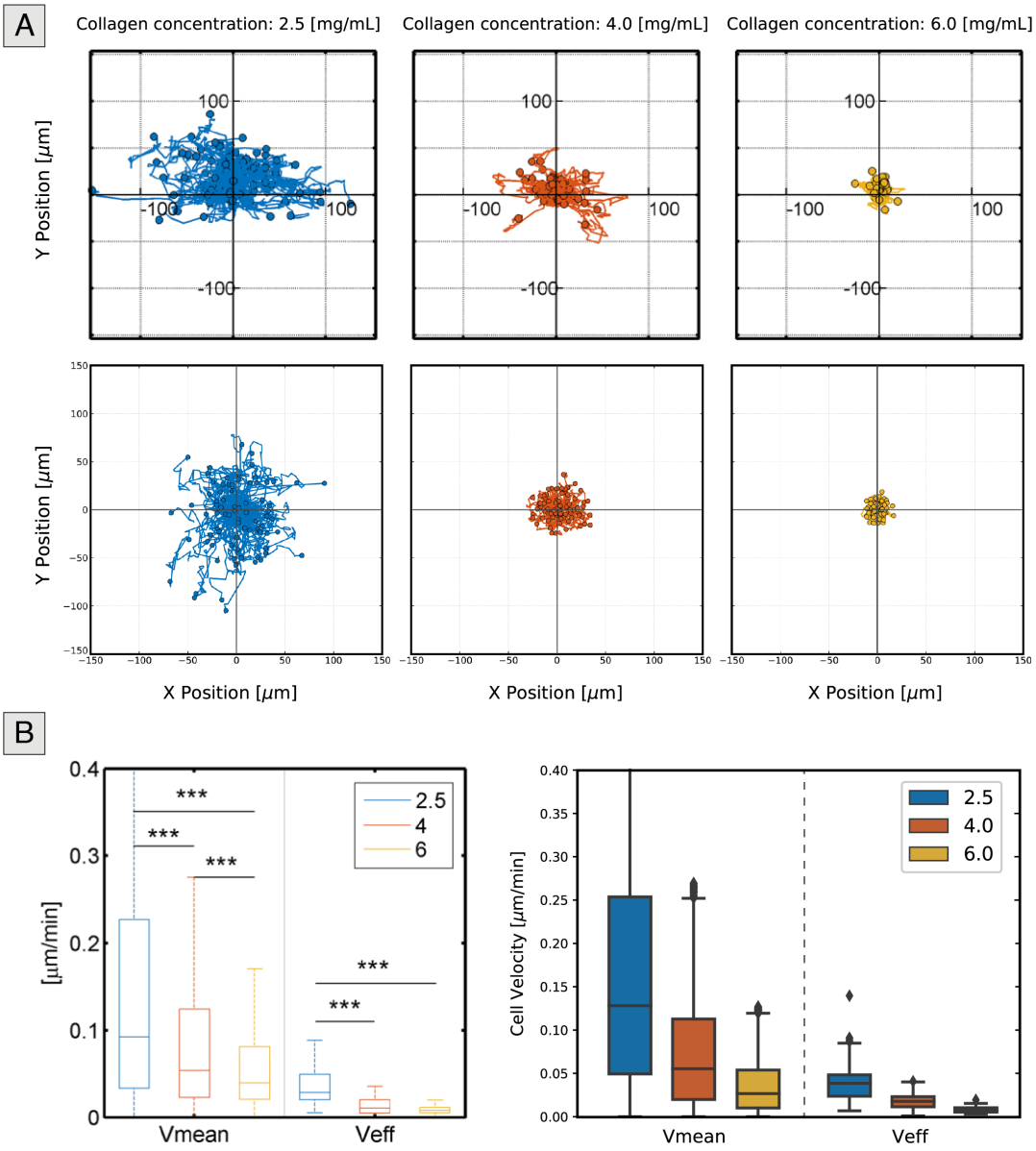

Figure 3.5: **Experimental and simulated results for the individual migration setup.** (A) Representation of relative cell trajectories for the experimental (top) and computational (bottom) results. As the density of the collagen matrix increases, cells become more confined, resulting in reduced cell movement. (B) Mean and effective cell velocities for cells seeded in matrices of varying collagen density. Both the experimental (left) and computational (right) results indicate that as the density increases, cells travel shorter distances due to the restrictions imposed by the matrix and both the mean and effective speeds decrease.

Table 3.2: **Statistical data of the mean and effective cell velocities for different collagen concentrations for both experimental settings and computational simulations.**

V_{mean} [μm/min]	N		median		mean		se	
Density [mg/mL]	Exp	Comp	Exp	Comp	Exp	Comp	Exp	Comp
2.5	5181	5840	0.0922	0.1281	0.1696	0.1702	0.0020	0.0019
4.0	3807	5840	0.0537	0.0553	0.1040	0.0746	0.0023	0.0009
6.0	3837	5840	0.0392	0.0266	0.0630	0.0355	0.0023	0.0004
V_{eff} [μm/min]	N		median		mean		se	
Density [mg/mL]	Exp	Comp	Exp	Comp	Exp	Comp	Exp	Comp
2.5	82	80	0.0285	0.0390	0.0365	0.0421	0.0020	0.0026
4.0	57	80	0.0105	0.0182	0.0136	0.0185	0.0024	0.0010
6.0	60	80	0.0079	0.0082	0.0086	0.0086	0.0024	0.0004

2018). This model extension was published on GitHub[1] and is publicly-available for *PhysiCell* users to introduce this representation of the physical microenvironment in their models. Furthermore, we were able to qualitatively describe how an increase in matrix density lead to smaller cell velocity values and how this, in turn, suppressed the invasion of single cells.

We also introduced a representation of the distribution of cell-generated locomotive forces, which are fitted to resemble the experimental results. Therefore, the force components were temporally dynamic and distinct for each cell in order to replicate the heterogeneity in cell behaviour generally observed in nature (Metzner et al., 2021). Using this model, we can capture the heterogeneity seen in cell behaviour in our data and other previously published experiments Ehrbar et al. (2011); Quaranta et al. (2009). In a biological sense, this result indicates that cells are more likely to migrate at lower speeds. However, in a few instances, cells may migrate at much higher velocities than the general trend seen for that set of conditions. This result could also be indicative of a low motile fraction, but this is currently not implemented in the model; i.e., all cells are programmed to migrate with similar behaviour.

Regarding the limitations of our model, one of the shortcomings that we identified was its simplified depiction of collagen matrices. On the one hand, the model disregards the fibrous nature of collagen matrices and the effects that it may have on motion directionality, which was not studied here. On the other hand, we assume that the collagen matrices are homogeneous materials, which is not completely accurate for these experimental settings, in which there are local regions with higher and lower numbers of fibres Haeger et al. (2014); Valero et al. (2018). As a first approach, we are currently choosing to characterize drag forces through the value of the mean dynamic viscosity of the matrix for each collagen concentration. Nonetheless, we do know that viscosity values vary according to a normal distribution Valero et al. (2018). Accordingly, we could have defined the ECM not by constant viscosity values, but through this spatial distribution, which would provide us with a better description of matrix heterogeneity. However, instead, we

[1] https://github.com/m2be-igg/PhysiCell-ECM

considered the ECM as a substance (i.e., collagen density) with constant mechanical properties, and we indirectly evaluated the heterogeneity through the velocities measured in the experiments. Hence, the impact of this simplification in our results is quite low.

Finally, due to the centre-based nature of the model, it was not possible to gain a realistic approximation of how cells may deform. However, for this particular study, we consider the decreased computational cost and the ability to simulate a high number of cells to be preferable to a detailed representation of the microenvironment and cell geometry. Furthermore, since this model was fitted specifically to a set of experimental data related to a specific cell line and conditions, we stipulate that the model may not accurately describe experimental settings with different cell lines and matrix density values. Nonetheless, we believe that the model could be easily updated by fitting parameters to other experimental results, while the main mechanisms should not need significant changes.

3.4 An improved data-driven model of chemically-regulated motility

3.4.1 Study outline

In this work, we aimed to characterize the role of chemotactic gradients induced by PDGF-BB on human fibroblasts motility using an integrated experimental-computational approach. We based our implementation on previously published works that characterize the migratory response of cells seeded in microfluidic devices to chemical gradients (Moreno-Arotzena et al., 2014; Del Amo et al., 2017). As seen in Fig 3.6, the microfluidic devices were composed of two types of components: a central region filled with a collagen-based hydrogel, and two adjacent channels that acted as reservoirs for culture medium and other substances, such as PDGF-BB The central region of the devices contained one or three chambers, according to the studied conditions. Based on the geometry of these devices, we created chemotactic gradients by introducing PDGF-BB in one of the adjacent channels of the microfluidic devices. However, instead of seeding the fibroblasts inside collagen matrices embedded in the devices, as done in previous works, we created a cell monolayer on one of the walls of the cell media channels in contact with the collagen hydrogel. Thus, we were able to clearly visualize the migration patterns of the fibroblasts, starting at the monolayer and advancing through the collagen matrix.

In addition, we used a discrete-based computational model to replicate the migration trends observed *in vitro* and identify the model parameters that allowed us to reproduce the experimental data. In particular, we replicated the experimental conditions using a *PhysiCell* (Ghaffarizadeh et al., 2018) model, which we extended to account for the presence of chemical substances in the microenvironment and their effect on cell motility. Using this knowledge and relating it to the biological meaning of the model parameters, we aimed to understand which mechanisms produced the chemotactic response of isolated fibroblasts to PDGF-BB gradients. In particular, we implemented a Python-based automatic optimization pipeline to couple *PhysiCell* and a publicly-available BO library (Nogueira, 2014), which allowed us to replicate the experimental results and

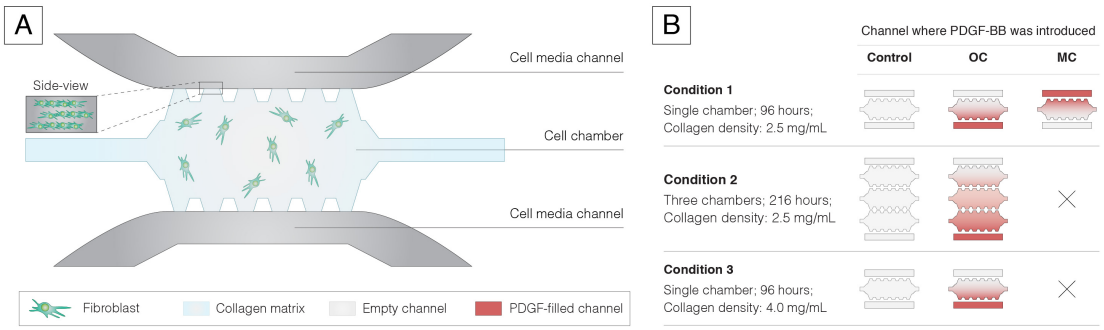

Figure 3.6: **Experimental-computational study design to characterize fibroblast motility in response to chemotactic gradients.** (A) Geometry of the microfluidic devices used in our study. The central microfluidic chamber contains a collagen matrix gel, which fibroblasts entered as they migrated from the cell monolayer in which they were seeded, as shown in the detailed side-view representation of the cell monolayer. (B) Three main experimental conditions under study and their implementation. Chemical gradients are represented through colour gradients which diffuse from the channel where PDGF-BB was placed. We defined that the monolayer would always be placed on the top channel, where cells were seeded.

predict which factors have the most significant impact on approximating the computational and experimental results.

To achieve this, a collaboration was established between the PhD candidate and members of the Mutiscale in Mechanical and Biological Engineering (M2BE) group, who designed and developed a set of microfluidic-based experiments to quantify the response of migrating fibroblasts to chemical stimulation. Furthermore, image processing analysis was performed by these researchers to measure the distance between the current position of the fibroblasts and their initial position in the fibroblast monolayer, thus quantifying their advance in the microfluidic device. Therefore, in this work, we will briefly describe the experimental research process employed to use this data as it is relevant to understand the computational model subsequently implemented by the PhD candidate. However, further technical details about the experimental setup and image processing analysis used in these experiments can be found in the work published by Movilla et al. (2023).

3.4.2 Methodology

Study design

This study aimed to explore three main research questions: (i) describing the effect of chemotactic gradients on cell motility; (ii) evaluating how the response of fibroblasts changed over time; and (iii) characterizing the chemotactic effect in matrices of distinct collagen concentrations. A graphical representation of the experiments performed to address each research question is shown 3.6B. The specific objectives of each condition are stated below:

i **Condition 1** aimed to characterize the response of migrating fibroblasts to the presence of a chemical gradient. Furthermore, we have studied how the spatial location of the gradient

may affect cells differently by considering both experiments where a chemical gradient was induced from the monolayer channel, and experiments with a chemical gradient induced from the channel that is opposite to the monolayer. Cells were seeded in collagen matrices with a concentration of 2.5 mg/mL and the experiments were performed over 4 days.

ii **Condition 2** focused on extending the previous configuration for longer periods, specifically 9 days, to define the evolution of the chemotactic effect over time. Three chambers were used to allow cells to migrate for larger distances, and the collagen density of the matrices was kept at 2.5 mg/mL. Here, we only created a chemotactic gradient by introducing PDGF-BB in the opposite channel to the cell monolayer, and we used a control condition.

iii Lastly, **Condition 3** aimed to characterize how distinct collagen densities modulated the chemotactic effect. Hence, we used single-chamber microfluidic devices embedded with collagen matrices with a concentration of 4 mg/mL. We considered a chemotactic gradient created by introducing PDGF-BB in the opposite channel to the monolayer, as well as a control experiment. Experiments were run over 4 days.

In addition, we performed a short study to quantify intercellular distance and whether it was modulated by the chemical factors to confirm if cells were moving as individuals or collectively. The results for this study are presented in Appendix B, Section B.1.5.

Experimental setup

The experiments that generated the data used in this study were conducted in microfluidic devices fabricated in polydimethylsiloxane (PDMS) following the methodology implemented by Shin et al. (2012). The geometry of the microfluidic devices consisted of a central chamber in which we introduced the hydrogel and two adjacent media channels that allowed for the introduction of cell culture medium and other substances (e.g., PDGF-BB). The spatial location of the media channels enabled the production of chemical gradients through the addition of a growth factor in one of the channels (Moreno-Arotzena et al., 2014; Del Amo et al., 2017). The study previously published by Moreno-Arotzena et al. (2014) presents further details about the geometry of these microfluidic devices.

A type I collagen gel solution was used to prepare collagen hydrogels of different concentrations, namely 2.5 and 4.0 mg/mL, following the methodology proposed by Shin et al. (2012). Furthermore, normal human dermal fibroblasts (NHDF) were transducted with a green fluorescent protein (GFP) to enable cell tracking and were subsequently added to the side channel of the microfluidic device to a final dilution of 2.5 x 10^5 cells/mL to form the fibroblast monolayer. After 24 hours of incubation since hydrogel seeding, 5.0 ng/mL of PDGF-BB was added to one of the adjacent microfluidic channels. Thus, a chemical gradient was established by a diffusive process across the hydrogel when PDGF-BB was added to the devices. The adjacent channels were filled every two days with the appropriate culture medium (i.e., growth medium with PDGF-BB for

the chemotaxis conditions and growth medium for the control conditions). The experiments were reproduced one to four independent times, with 4-10 technical replicates in each experiment.

Phase contrast and fluorescent (GFP) images were captured every 24 hours for 4-9 consecutive days, depending on the experimental condition. The focal plane was selected to be in the middle along the z-axis of the device, focusing on cells fully embedded within the 3D network. Each image was then automatically evaluated with a custom-made MATLAB script to count the number of cells in each channel and quantify the advancement of the individual fibroblasts in the perpendicular direction to the monolayer. It was assumed that the position of each fibroblast was given by its centre, despite the elongated morphology of these cells. Statistical significance between the distance data for cells grown in environments with and without PDGF-BB was evaluated through one-way ANOVA tests, followed by a post-hoc Tukey HSD test when necessary.

Model implementation

As done in Section 3.3, our model was implemented in *PhysiCell* and cell motility was modelled as a combination of cell-generated locomotive forces and interactions with the environment, as previously described in Eq 3.2. In addition, we extended our model to consider cell-wall repulsion ($\mathbf{F}^i_{cwr}$) between the cells and the walls of the microfluidic device. Thus, cell positions were updated by computing the effect of all the forces acting on a cell i, as defined by Eq 3.7.

$$\mathbf{v}_i = \frac{1}{\mu}\left(\sum_{j\in\mathcal{N}(i)}(\mathbf{F}^{ij}_{cca}+\mathbf{F}^{ij}_{ccr})+\mathbf{F}^i_{loc}+\mathbf{F}^i_{cwr}\right) \tag{3.7}$$

A force diagram representing these forces is shown in 3.7. In this work, we neglected cell-wall adhesion forces to better reproduce the *in vitro* behaviours observed in our microfluidic assays. In these observations, cells tended to favour migration from the monolayer to invade the 3D collagen matrix. Since we adopted a model simplification that placed cells near a wall to replicate the monolayer, we established that cell-wall adhesion should be neglected so that cells would prefer invading the matrix. These interactions were modelled through potential functions, as fully described in previous works (Ghaffarizadeh et al., 2018), and their corresponding coefficient values are presented in Table 1. Here, we consider that, under these settings, fibroblasts migrate from the monolayer through the collagen matrices as individual cells, and cell-cell connections are primarily transient. Thus, as cells do not create permanent cell-cell adhesion structures, their motility differs from the patterns usually seen during in vivo wound healing, characterized by the significant formation and maintenance of cell-cell adhesion structures (Sunyer et al., 2016; Ozcelikkale et al., 2017; Tambe et al., 2011). Accordingly, we focused primarily on the role of the cell-generated forces.

Additionally, we extended *PhysiCell*'s standard cell motility functions to provide finer control over cell movement. Based on our experimental observations, we concluded that, although fibroblasts were seeded in 3D conditions, cells had restricted vertical motion because of the small height of the microfluidic device, yet this was not reflected in the 2D component. Thus, we defined that

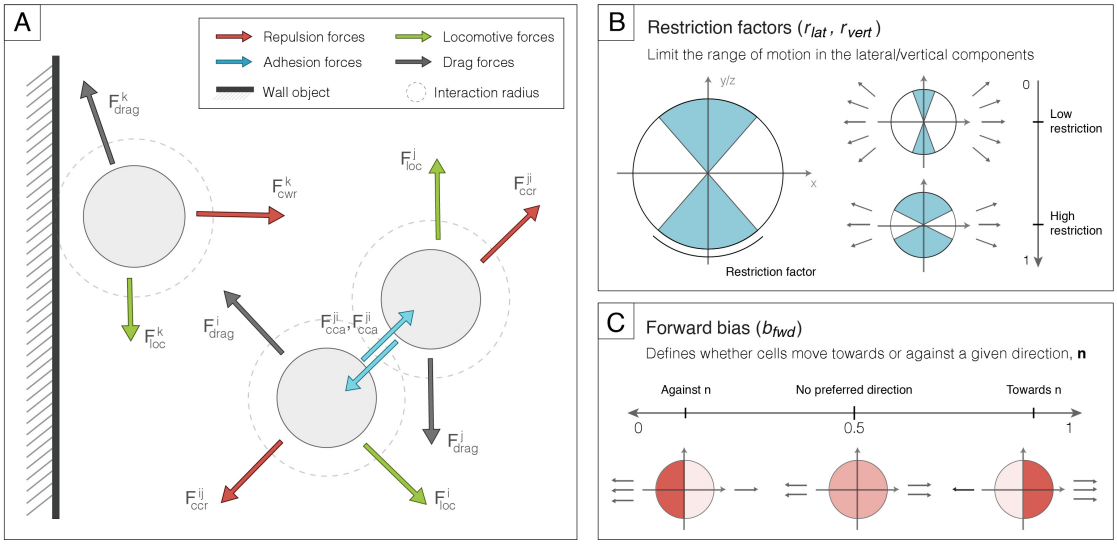

Figure 3.7: Representation of the forces considered in our computational model. Representation of the model's mechanics and cell motility functions. (A) Simplified representation of the force diagram for the interactions between the cells and the microenvironment. The model considers cell-cell repulsion and adhesion ($\mathbf{F}^{ij}_{ccr}$ and $\mathbf{F}^{ij}_{cca}$, respectively), repulsion forces between the cells and the walls that define the boundaries of the computational domain ($\mathbf{F}^{i}_{crw}$), cell-generated locomotive forces ($\mathbf{F}^{i}_{loc}j$). Furthermore, the model takes into account the interactions between the cells and the ECM through the drag coefficient given by previously obtained experimental data that characterize the mechanical properties of collagen matrices (μ). (B, C) Representation of how the direction of cell-generated forces is computed every time the cell velocity is updated, based on a set of parameters defined at the start of the simulation (n, b_{fwd}, r_{lat}, r_{vert}). These parameters define how motility may be restricted in its vertical and lateral components (represented by the coloured areas) and whether cells are more probable to move towards or away from a defined direction (represented by the opacity of the coloured areas in the lateral component). The figure shows an example of a unit vector n that can be chosen based on the represented parameters.

Table 3.3: **Reference parameter values for the chemotaxis model.** Values in brackets represent ranges, showcasing that the parameter values were not fixed and varied based on the configuration of the simulation to reflect the presence and spatial location of chemotactic gradients.

Symbol	Parameter	Value	Unit	References
Δt_{mech}	Mechanics time step	0.1	min	Ghaffarizadeh et al. (2018)
Δt_{cell}	Cell processes time step	6.0	min	Ghaffarizadeh et al. (2018)
$\rho^0_{Collagen}$	Collagen concentration	[2.5, 4.0]	mg/mL	Experimental
μ	Drag Coefficient	[7.96, 18.42]	Pa · s	Valero et al. (2018)
R	Cell radius	7.0	μm	Thoumine et al. (1999)
R_A	Cell adhesion distance	$1.25R$	μm	Ghaffarizadeh et al. (2018)
t_{double}	Cell doubling time	50.0	h	Gupta et al. (2005), estimated
t_{death}	Cell lifetime	720.0	h	Rubin (2002)
$t_{persist}$	Cell persistence time	[10.0-60.0]	min	Calibrated
c_{cca}	Cell adhesion coefficient	[1.0-10.0]	-	Ghaffarizadeh et al. (2018), calibrated
c_{ccr}	Cell repulsion coefficient	[10.0-100.0]	-	Ghaffarizadeh et al. (2018), calibrated
σ	Locomotive forces factor	[0.0-6.0]	-	Calibrated
b_{fwd}	Forward bias	[0.0-1.0]	-	Calibrated
r_{lat}	Lateral restriction factor	[0.0-1.0]	-	Calibrated
r_{vert}	Vertical restriction factor	[0.0-1.0]	-	Calibrated

our model should enable restriction over these two components of cell motility independently. We achieved this by introducing two new model parameters, r_{lat} and r_{vert}, to be defined in more detail in the text. In addition, we aimed to model the effect of chemotaxis while still preserving some of the randomness seen in cell movement. In order words, we defined that, in our model, fibroblasts seeded in the presence of PDGF-BB should have a greater probability to move along the resulting chemotactic gradient. Yet, they must retain some freedom to move in the opposite direction. Thus, we defined a new parameter, b_{fwd}, which modulated the probability of moving towards a given gradient. Taking this into consideration, we established that, every time cell-generated forces were updated for a cell i, their components were defined by Eq 3.8.

$$\mathbf{F}^i_{loc} = \mathrm{F}_{loc} \cdot \sigma \cdot \mathbf{e}^i(\mathbf{n}, b_{fwd}, r_{lat}, r_{vert}) \tag{3.8}$$

Here, $\mathbf{e}^i$ is a random unit vector that represents the direction of the forces and is defined according to user-defined parameters ($\mathbf{n}$, b_{fwd} r_{lat}, r_{vert}). As previously stated, r_{lat} and r_{vert} regulated the range of lateral and vertical movement, respectively. For example, if r_{vert} was defined as 0, the cells were able to move freely in 3D, yet if this parameter had a value of 1, the cellular movement would be completely restricted to 2D. A similar approach applies to r_{lat}, which determined whether cell motility strictly followed a bias vector, $\mathbf{n}$, or if cells could stray away laterally from this direction. The probability of a cell moving towards or away from n was given by b_{fwd}. For a user-defined value of 0, the cell always moved away from the n, and for a value of 1 the cell always moved towardsn. In the current version of the model, we assumed n to be a normal vector to the monolayer, which pointed towards the opposite wall. The effect of these parameters is visually represented in Fig 3.7 (subfigures B and C).

Moreover, F^i_{loc} describes the magnitude of cell-generated forces. To capture the heterogeneous behaviour of cell movement, we modelled force magnitudes through a Rayleigh distribution, a positively-skewed distribution, defined by its mean value (σ), following previously published works (Mark et al., 2018). Accordingly, the magnitude of cell-generated forces was not fixed, but rather varied over time. Hence, we were able to reproduce some of the variability in cell movement that may arise due to cell-matrix interactions, which we did not explicitly account for in our model. We considered that the cell persistence time was regulated by the presence of PDGF-BB and fitted its value for every experimental condition. The range that defined the parameter search space for this model input is presented in Table 3.3.

Regarding cell proliferation and death, we considered a simple cell cycling model that was defined by the cell proliferation rate, and we modelled apoptosis by defining a cell death rate. The selected values for these model parameters were based on results found in the literature (Rubin, 2002; Gupta et al., 2005) and the observed behaviours in our experiments, and they can be found in Table 3.3. As we aimed to use our computational model in combination with data from in vitro experiments, the model required some refinements to accurately replicate the experimental design. Having this into account, the computational domain was defined to resemble the central region of the microfluidic chips. Besides, cells were initialized to reproduce a simplified representation of a monolayer in the top wall of the computational domain, forming a hexagonal grid in which cells did not touch their neighbours.

Model optimization

In this study, we aimed to maximize the similarity between the computational and experimental datasets for each of the conditions under study. Accordingly, we defined a fitting function that took into account both datasets to compute their similarity as given by the BC, which is a metric that can be used to compute the similarity between two discrete distributions (Comaniciu et al., 2000). Particularly, here we considered the normalized *in vitro* and *in silico* histograms for the distance travelled by cells at each day. Having defined these data structures, the similarity between each bin of the histogram ($hist^i$) was computed and the BC was calculated based on Eq 3.9. Given that our results consist of data representative of several days, we computed the BC for each day and, subsequently, we calculated the average of these values.

$$BC = \sum_{i=1}^{N} \sqrt{hist^i_{experimental} \cdot hist^i_{simulated}} \tag{3.9}$$

We implemented BO into our computational framework by building a Python-based pipeline that linked the compiled C++ *PhysiCell* code to in-house Python scripts that analyse the experimental and computational data to compute the similarity between the two datasets, as defined by the BC. Additionally, we used a publicly available Python package that implements BO algorithms (Nogueira, 2014) and integrated it into our framework to find the parameters that maximize the similarity between experimental and computational data, as shown in Fig 3.8. As an initial

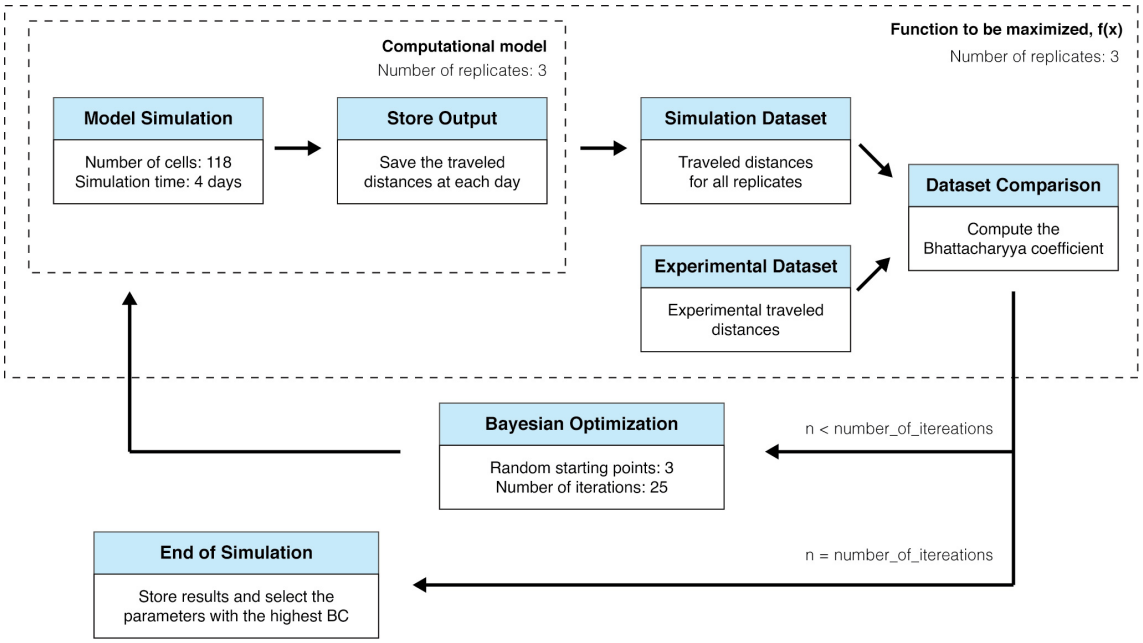

Figure 3.8: **Simulation and optimization workflow for the chemotaxis model.** The optimization framework was composed by three levels: the model (*PhysiCell*), the data-analysis component (in-house Python scripts) and the optimization algorithm (developed by Nogueira (2014)). Input parameters at the start of the simulation were defined by the BO algorithm based on a set of user-defined ranges, presented in Table 3.3. It was defined that the model would start by exploring 3 random points in the parameter space, and subsequently select the next values to be evaluated. At each iteration, the algorithm selected the parameter values that provided the best similarity results for the previous iterations, and estimated a new set of parameters to converge to a solution. This process was repeated during 25 iterations. At each iteration, the second level was run 3 times and the average BC value was computed. In turn, every time the second level was run, the *PhysiCell* model was called 3 times, to get a total number of cells similar to that obtained for in vitro experiments.

approach, we used the experimental data for the control conditions to calibrate the model parameters that regulate cell motility', previously presented in Table 3.3, without taking into account the effect of PDGF-BB. Henceforth, we defined an optimization routine that started by sampling three random points of the parameter space. This was done to avoid converging to a local solution of the maximization problem without exploring the parameter space, overlooking other possible solutions.

For the following iterations, the optimization algorithm chose new parameter values based on the knowledge gathered from other iterations, i.e., which values produced the maximal similarity values. Subsequently, we assumed that the presence of PDGF-BB would primarily lead to changes in the persistence time, the magnitude of forces, and the lateral restriction and forward motility bias factors. Hence, we fixed the remaining parameters and ran the optimization algorithm focusing solely on these values. In addition, we probed the parameter space at the start with the optimization pipeline with the best set of parameters obtained for the control conditions, as we assumed that these values had biological significance and should be sampled. To reduce the computational cost

of running a large number of simulations, we narrowed the size of the parameter search space at each iteration through the sequential domain reduction technique (Stander and Craig, 2002). This approach was offered by the Python library in use (Nogueira, 2014). Hence, at each new iteration, the region of interest of the parameter space was updated based on the results of the previous iteration to be centred around the parameters that maximized the similarity. Besides, the size of the region of interest was reduced. Consequently, the optimization function was less exploratory but converged faster to an optimal value.

Simulation workflow

As shown in Fig 3.8, each *PhysiCell* model simulation was comprised of 118 cells, and 3 replicates were considered to achieve the same number of data points as in the experimental results, approximately. Furthermore, at each iteration of the optimization algorithm, we considered three independent sets of computational results. For each dataset, the BC was computed in regards to the experimental data, and we considered the average similarity for these values to compute a mean BC value, that was passed to the optimization algorithm as our target metric. The BO algorithm consisted of a total of 25 iterations for each of the experimental conditions. At the end of the optimization pipeline, we selected the parameter combinations that maximized our optimization function. In cases where distinct parameter sets produced similar optimal values, we analysed the results and selected the best parameters based on their biological significance and their coherence with the chemical gradients produced under each experimental condition.

3.4.3 Results

PDGF-BB gradients induce and guide human fibroblast motility

In Condition 1, we focused on characterizing the chemotactic effect of PDGF-BB on fibroblasts seeded in single-chamber devices with 2.5 mg/mL collagen matrices in the direction perpendicular to the monolayer. The results of this study are shown in Fig 3.9. Comparing the control experiment with those where PDGF-BB was present (i.e., monolayer channel (MC) opposite channel (OC) conditions), we observed that fibroblasts presented a more homogeneous spatial distribution in the presence of this chemoattractant substance, as shown in Fig 3.9A. Moreover, we observed that this behaviour was further enhanced when PDGF-BB was placed in the OC. The enhanced motility of fibroblasts in the presence of PDGF-BB started being apparent after 24 hours, and it became more evident over time. Fig 3.9B presents the distribution of the travelled distances for each day of the experimental conditions. This quantitative representation allowed us to verify the previously described behaviours. After 48 hours, the median travelled distances were 96.8, 143.6 and 201.2 µm, for the control, MC and OC conditions, respectively. Therefore, the median values for the travelled distances were increased by 48.3% when PDGF-BB was added in the MC, and there was an increase of 107.8% when it was placed in the OC. Similarly, after 96 hours the median travelled distances were 141.6, 220.4 and 310.0 µm, corresponding to an increase of 55.6% and 118.9% for the aforementioned conditions. Furthermore, the performed statistical analysis indicated that the

differences between the distance data were statistically significant for all of the studied conditions (p-value < 0.001).

These results suggest that the presence of PDGF-BB enhanced the motility of fibroblasts, allowing them to advance farther into the microfluidic chip. It is also apparent that the direction of the chemical gradient was a relevant factor to promote cell migration. When the chemotactic factor was placed in the MC, cells showed enhanced motility in comparison to the control conditions. However, this effect was more pronounced when PDGF-BB was placed in the OC. We conducted additional experiments to test if the migration patterns of fibroblast would change in response to the use of PDGF-BB at a higher concentration. Additionally, we performed another study in which we ceased the administration of PDGF-BB at 48 hours. Overall, we observed that the distances travelled by fibroblasts were not affected by these conditions. These results can be found in Appendix B, Section B.1.3.

Model predictions suggest that active cell forces define the response to chemotactic gradients

Using the data presented in the previous section, we implemented and calibrated a computational model to predict the underlying mechanisms of the observed motility patterns. We used a discrete-based model to study fibroblast motility as a combination of purely random and gradient-based migration. In addition, we adjusted the model input parameters to capture the characteristic behaviours of each experimental condition, and we defined the computational domain to replicate an approximated geometry of the microfluidic devices used in the experiments. We used our experimental data alongside automatic optimization techniques to identify the model parameters that were the most relevant to replicate the observed in vitro results. During the optimization procedure, we considered the magnitude of the active cell locomotive forces (σ), the lateral and vertical movement restriction factors (r_{lat} and r_{vert}, respectively), the probability of moving towards or away from the OC (b_{fwd}) and the motility persistence time ($t_{persist}$). Using this approach, we tuned our computational model to produce results similar to the experimental data. Fig 3.9C shows the overlap between the experimental and computational distributions for the distances travelled by fibroblasts over time, under control conditions. We conclude that datasets are similar, as indicated by the similarity coefficient (BC), also presented in this figure. However, we recognize that, at day 4, the simulations produced distance results that were more normally distributed than the experimental data. Hence, we propose that the model may not be able to account for the different dynamics that ECM cur at the leading edge of the invasion, where there are fewer cells usually named as leaders, and the trailing edge, which presents a larger number of cells characterized as followers (Mayor and Etienne-Manneville, 2016). This phenomenon has been documented extensively in collective cell migration but, since our model considers individual cell motility, it was not possible to replicate it in this work. The BC values for the conditions described in the previous section are represented in Table 3.4, alongside the parameter values that provided the best similarity values between experimental and computational data.

Comparing the three conditions under study, we concluded that the forward bias (b_{fwd}) should be reduced to simulate the results obtained when PDGF-BB was added to the MC, whereas it

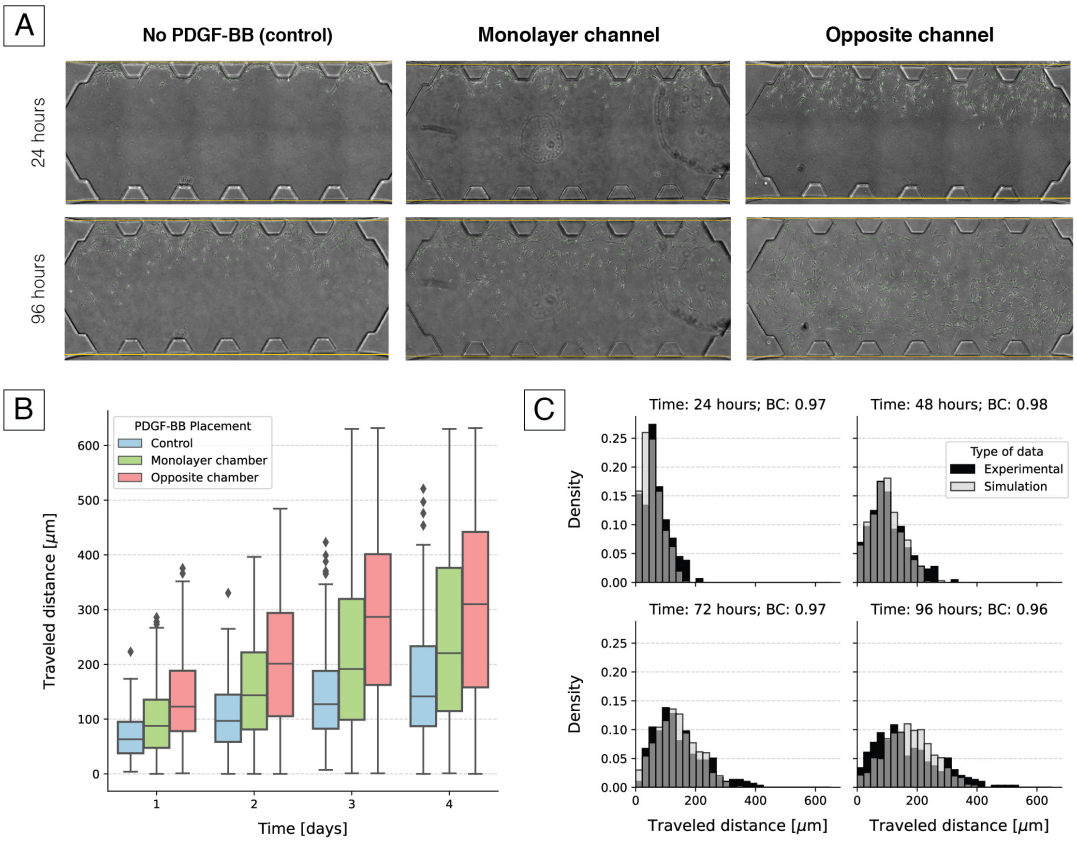

Figure 3.9: **Experimental and computational results for the travelled distances in 2.5 mg/mL matrices over 4 days.** (A) Spatial distribution of migrating fibroblasts seeded in microfluidic devices over 4 days, with and without the introduction of PDGF-BB at a concentration of 5 ng/mL. (B) Quantification of the distances travelled by fibroblasts, with and without chemical stimulation. Distances were computed based on the perpendicular distance between the current position of each cell and the position of the monolayer. The differences between all groups on each day of experiments were evaluated to be statistically significant (p-value < 0.001). (C) Overlap between the experimental and computational results for the distribution of the distances travelled by cells over 4 days, under control conditions. The represented histograms serve as a graphical representation of the similarity between the two datasets and showcase the model's ability to replicate the experimental data. Similarity was computed through the BC, which evaluates the overlap between two discrete histograms and quantifies it as a value between 0 (representing no overlap between the two groups) and 1 (representing two identical groups). We ran 3 replicates for our simulations and computed the BC between simulation results and experimental data for each group.

Table 3.4: **Model optimization results for the experiments where fibroblasts were grown in 2.5 mg/mL collagen matrices over 4 days.** We tuned our computational model through an automatic optimization procedure based on BO to replicate each of the experimental conditions studied for fibroblasts seeded in 2.5 mg/mL collagen matrices. We quantified the similarity between the experimental and computational data using the BC, which measures the overlap between two datasets and defines it as a value between 0 (no overlap) and 1(maximum similarity). The table shows the most relevant model input parameters that produced the best similarity values, alongside their values. The addition of PDGF-BB could be replicated through an increase in the magnitude of the cells' active locomotive forces (σ) and the cells' persistence time ($t_{persist}$), which resulted in faster and more directed movement, respectively. In turn, the forward bias value (b_{fwd}) was modified to reflect the direction of the chemotactic gradient. For the MC experiments, the bias factor was decreased to reflect the attraction of cells towards the monolayer, where the chemical concentration was highest. On the contrary, the addition of PDGF-BB in the OC was simulated by increasing the bias factor.

Condition	Parameter Values					BC Value
	σ	$t_{persist}$	b_{fwd}	r_{lat}	r_{vert}	
Control	1.4	17.2	0.56			0.972 ± 0.004
PDGF-BB in the monolayer channel	2.4	21.0	0.55	0.76	0.36	0.975 ± 0.001
PDGF-BB in the opposite channel	2.0	24.4	0.58			0.937 ± 0.004

should be increased to replicate the OC results. In other words, to replicate the results of the MC experiments, the best simulation results were obtained when cells were more likely to migrate backwards in comparison to the control conditions. On the contrary, to achieve the behaviour observed when PDGF-BB was introduced in the ECM , the probability of migrating towards this channel had to be increased. This is coherent with the expected PDGF-BB gradient for each of these cases. In regard to the cell-generated forces, the model results obtained for σ seem to indicate that active cell forces should be stronger in the presence of PDGF-BB than in control conditions. Moreover, the magnitude of these forces had to be further enhanced to replicate the MC experiments. Lastly, we concluded that our model suggests that the presence of PDGF-BB had a role in regulating cell directionality, since the cell persistence times had to be increased to replicate the experimental data. Therefore, we highlight the role of chemotaxis, which induces cell movement and directionality, in contrast to chemokinesis, which only stimulates cell movement in random patterns (Kamiyama et al., 1998). Nonetheless, these observations should be interpreted as model predictions that require experimental validation, using techniques tailored to quantify the variables identified by the model at the cell level.

The chemotactic effect of PDGF-BB is persistent over time

To assess the effect of PDGF-BB on fibroblast motility over extended periods, in Condition 2 we extended our previous experimental configuration to include three chambers instead of a single chamber. Experiments were run for 9 days, and we considered a collagen density of 2.5 mg/mL. Besides, we only performed control and OC experiments, as these proved to provide the most

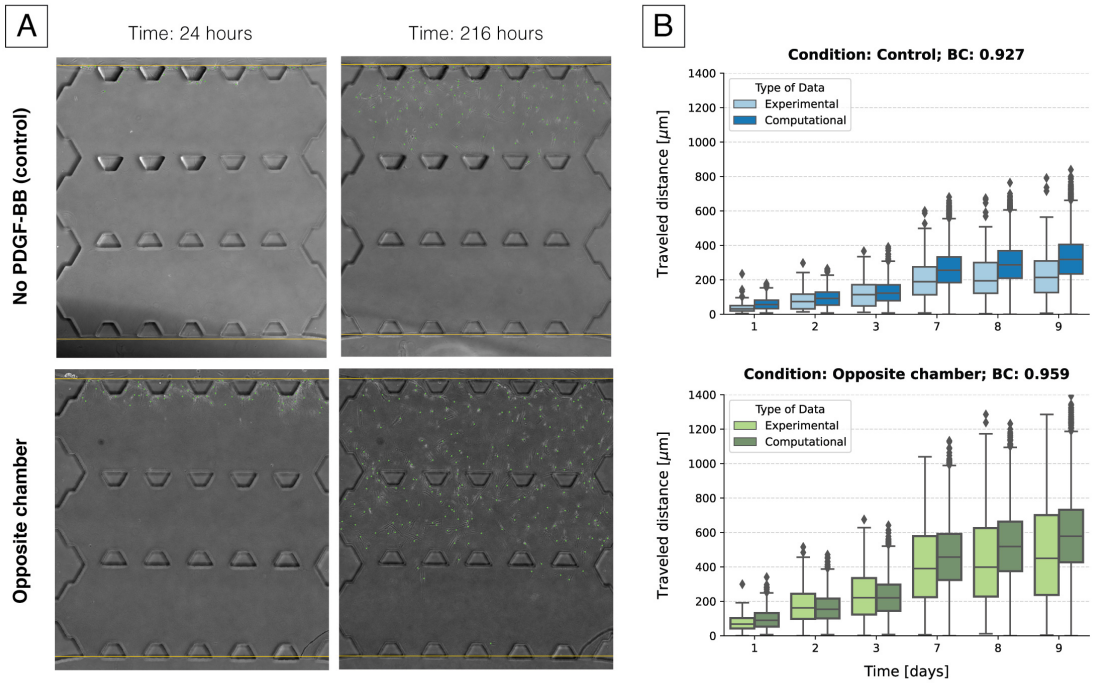

Figure 3.10: **Experimental and computational results for the travelled distances in 2.5 mg/mL matrices over 9 days.** (A) Spatial distribution of migrating fibroblasts. (B) Quantification of the experimental and computational distances travelled by fibroblasts in collagen matrices. Comparing the experimental and computational data, the similarity values indicate that the model was able to capture the general trends of the migrating fibroblasts (BC > 0.93). Nevertheless, the direct comparison of both datasets shows that the model produced larger distances values than those observed experimentally during the last days of simulations, specifically under control conditions.

significant results in the previous sections. Fig 3.10A shows that the presence of a PDGF-BB gradient induced enhanced fibroblast motility. We observed that after 216 hours the fibroblasts seeded in matrices with chemical stimulation were able to migrate twice as much as those seeded in control conditions. This was validated by the quantitative results for this metric, presented in Fig 3.10B. The median values for the distance travelled by fibroblasts after 48 hours were 65.6 and 161.0 μm, for the control and the PDGF-BB assays, respectively. Hence, we concluded that the presence of PDGF-BB lead to an increase of 145.4%. After 216 hours, the median value increased from 205.7 μm in the control conditions to 449.2 μm in the presence of the chemical factor. In turn, this represented an increase of 118.3%, showcasing once more the enhanced motility of fibroblasts seeded in the presence of PDGF-BB. Furthermore, the performed statistical analysis indicated that the differences between the distance data in the absence and presence of PDGF-BB were statistically significant (p-value < 0.001).

We also studied the effect of chemical stimulation over 9 days through our computational model. Henceforth, we used the parameter values that provided the best results for the experiments presented in the previous section. Particularly, we considered the values obtained for the control condition and the assays in which PDGF-BB was introduced in the opposite channel to the

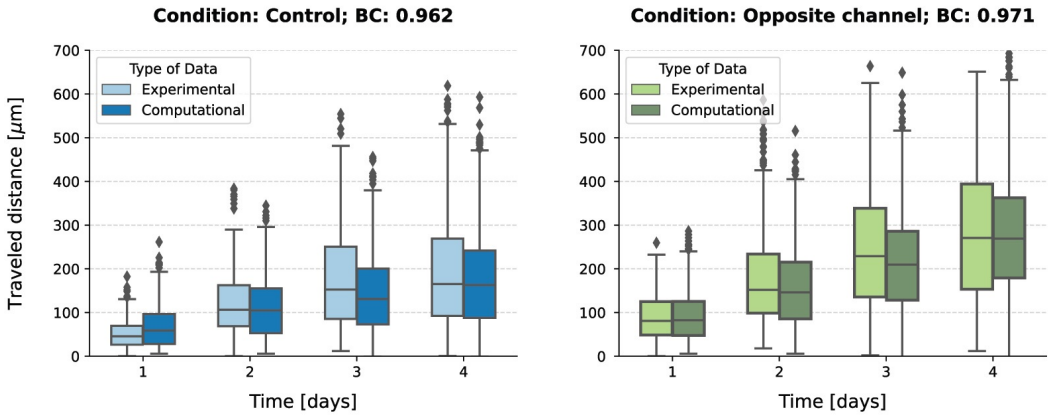

Figure 3.11: **Experimental and computational results for the travelled distances in 4.0 mg/mL matrices over 4 days.** Quantification of the experimental and computational distances travelled by fibroblasts. Comparing the experimental and computational data, we conclude that our model was able to accurately replicate the emergent behaviours (BC > 0.96).

monolayer. We ran the simulations for increased periods to resemble the new experiments, considering a total of 9 simulated days. We compared the experimental and computational datasets and obtained a BC of 0.93 for the control conditions, and a BC of 0.96 for the condition with chemical stimulation. These results suggest that, despite a small decrease in the BC values, the same input parameters values were able to replicate the experimental data for both the experiments carried out for 4 days and those that lasted for 9 days. Consequently, it is apparent that, at the computational level, there were no significant differences in cell behaviour despite the extended simulation time. Comparing the two datasets represented in Fig 3.10B for the control conditions, we can conclude that the model was able to replicate the experimental data during the first days of simulation. However, by day 7, we observed that the simulated values were larger than the values obtained experimentally. Thus, we propose that the motility mechanisms defined by the initial model parameters did not fully capture cell behaviour in the last days of experiments. The same pattern could be observed to a lesser extent for the condition that considered chemical stimulation.

Collagen density hinders the chemotactic effect of PDGF-BB

In Condition 3, we aimed to assess how differences in the concentration of collagen modulated the cell response to the presence of PDGF-BB gradients. Therefore, we embedded collagen matrices with a density of 4.0 mg/mL in single-chamber microfluidic devices. As in Condition 2, two experiments were performed (control and OC). The results of this study are shown in Fig 3.10. Based on this data, we concluded that fibroblast motility was enhanced when PDGF-BB was introduced in the system, as we also observed in the previous sections. Furthermore, statistical analysis indicated that the differences between the distance data in the absence and presence of PDGF-BB were still statistically significant (p-value < 0.001).

After 48 hours, the median values were 106.3 and 148.9 μm for the control condition and the experiments with PDGF-BB, respectively. Accordingly, the travelled distance values increased by 40.2% due to the addition of the chemical factor. Similarly, after 96 hours there was an increase of 61.6%, as the median value increased from 165.8 μm for the control conditions, to 268.1 μm for the OC experiments. Nonetheless, we should consider that, as we have previously demonstrated, the median value for the travelled distances after 96 hours increased by 118% for fibroblasts seeded in collagen matrices with a concentration of 2.5 mg/mL. Hence, even though motility was enhanced by the addition of PDGF-BB in 4.0 mg/mL matrices, the resulting chemotactic effect was not as strong as that observed for cells seeded in matrices with lower collagen densities. To justify these results, we suggest that the diffusion of PDGF-BB was faster in matrices with low collagen densities (i.e., matrices with larger pore size), which facilitated the chemotactic effect. In addition, cells may have faced some difficulty when entering the collagen matrices, probably due to the increased density values. A more in-depth comparison between the results obtained for distinct collagen densities can be found in Appendix B, Section B.1.4.

As done in the previous sections, we simulated these experiments with our computational model and used the experimental data to identify the most relevant input parameters to replicate the observed migration trends. Moreover, we compared the values we obtained for this condition to those we have obtained for the experiments considering a collagen density of 2.5 mg/mL. We represented the presence of the collagen matrix in our model through its physical properties, namely its viscosity, which had a direct effect on the motility of each cell. We ran our model optimization pipeline using the presented experimental data as our target data, and we studied which model input parameters values produced the best similarity results. Table 3 presents the parameters that were the most relevant to replicate the experimental data, alongside their values and the associated BC value. Once again, we observed that the parameters that define the active cell motility forces (i.e., σ, b_{fwd} and $t_{persist}$) were identified by the computational framework as the main regulators of the chemotactic response. In addition, the model results suggested that the vertical restriction factor should increase in comparison to previous simulations, which we assume to be representative of the higher confinement levels that cells face as the collagen density increases.

In regard to the parameters that modulate the cell-generated locomotive forces, by comparing control and chemical conditions, we observed similar patterns to those previously described when PDGF-BB was added to microfluidic devices with 2.5 mg/mL collagen matrices. On the one hand, to replicate the chemotactic effect produced by PDGF-BB, the best set of results was obtained when the magnitude of the cell-generated forces was increased. Similarly, these results also relied on higher values for the forward bias, which dictated that cells were more likely to move towards the channel opposite to the monolayer. Unlike the previous results showed, though, we did not find a significant difference between the persistence times obtained for the control and the PDGF-BB conditions. However, we observed that the persistence time for these simulations (49.8 minutes) was particularly higher than those obtained for lower collagen concentrations (17.2-24.4 minutes). Hence, it is suggested that fibroblasts seeded in collagen matrices of higher concentration changed the magnitude and direction of their locomotive forces fewer times than those

Table 3.5: **Model optimization results for the experiments where fibroblasts were grown in 4.0 mg/mL collagen matrices over 4 days.** The table shows the most relevant model input parameters that produced the best similarity values, alongside their values. The similarity between the experimental and computational datasets is given by the BC, which is defined as a value between 0 and 1, with 1 corresponding to maximum similarity. As suggested by the results shown in the previous sections, the addition of PDGF-BB seems to increase the magnitude of cell-generated forces and cell directionality. In addition, it is apparent based on these results that the increased density of the collagen matrices promoted more restricted migration patterns, as given by the high vertical restriction factor values, and more directed movement, indicated by the high persistence times.

Condition	Parameter Values					BC Value
	σ	$t_{persist}$	b_{fwd}	r_{lat}	r_{vert}	
Control	2.8	49.8	0.56	0.89	0.36	0.972 ± 0.004
PDGF-BB in the opposite channel	3.2		0.61			0.937 ± 0.004

seeded in low concentrations. Lastly, we highlight that the magnitude of the active cell forces was higher in these simulations than those obtained in the previous sections. In other words, the computational results suggested that the cells applied stronger forces to move through the matrix. As previously stated, these results should be interpreted as model predictions about the underlying mechanisms of fibroblast chemotaxis at the cell level, yet they must be validated.

3.4.4 Conclusions

Here, we presented an improved data-informed that allowed us to have more control over cell motility and simulate the effect of chemotactic gradients. In addition, we were able to calibrate our model with the use of novel experimental quantitative data. Both the *PhysiCell* model extensions to simulate chemotaxis and the optimization pipeline scripts are publicly-available on GitHub[2]. Overall, in this work we observed that the presence of PDGF-BB promoted cell motility in fibroblasts as shown in previous works (Heldin and Westermark, 1999). Moreover, our results showed that fibroblast motility differed based on the direction of the induced chemotactic gradient. Comparing the datasets for the MC and OC experiments, we remark that motility was enhanced in both cases, but the invasion of the collagen matrix by fibroblasts was more pronounced when PDGF-BB was added to the OC. We propose that motility was enhanced but limited in the MC condition because cells were already close to the PDGF-BB source. When PDGF-BB was added to the OC, we postulate that matrix invasion was enhanced due to chemotactic signals sensed by cells, which promoted the invasion of the collagen matrix to reach the growth factor source. We further observed that the induced chemical effect produced by the addition of PDGF-BB in the opposite channel was sustained over larger periods of time, as we have observed the same trends for experiments that lasted for 9 days.

[2]https://github.com/m2be-igg/pdgf-induced-motility

The combined use of our experimental data and a descriptive computational model allowed us to infer some of the mechanisms that may have originated the migration patterns observed in vitro. We concluded that the directionality and magnitude of the cell-generated locomotive forces, as well as the persistence time of the migrating fibroblasts, were key aspects to simulate the chemotactic effect induced by PDGF-BB. We further observed that the effects imposed on these parameters were sustained over time. We also highlight that our model enabled us to gain new insights about the behaviours observed when there were no chemical gradients in the devices. Generally, our results showed that cells were initially seeded in a collective structure but moved through the collagen matrix as individuals, as seen in previous works (Miron-Mendoza et al., 2012). Moreover, we observed that cells seeded under control conditions in particular still showed a preference to migrate towards the wall opposite to the cell monolayer, even though there was not a chemotactic gradient guiding them. In our computational model, this was represented by the increased forward bias values. We infer that this could have been motivated, in the experiments, by the large number of fibroblasts seeded in the monolayer. Cells may have shifted their motility to invade the hydrogel and migrate towards cell-free regions, which provided more space for individual growth and motility (Oraiopoulou et al., 2018). Consequently, it is possible that fibroblasts seeded at lower cell concentrations will not migrate as much as there will be more cell-free regions close to the monolayer. Nevertheless, we observed that, as cells advanced through the microfluidic chip and reached lower cell density regions, forward motility appeared to have been conserved and preferred over random movement. Moreover, this behaviour could be representative of a preference by cells to invade the 3D microenvironment of the collagen matrices, rather than remaining attached to the 2D monolayer (Bayless et al., 2009).

Additionally, we considered the effect of the physical properties of the collagen matrices on the cells' motility patterns. In regard to the control conditions, we did not detect relevant differences in the gel invasion between cells seeded in matrices with a collagen density of 2.5 mg/mL and those with a density of 4 mg/mL. This was consistent with previous works which showed that the collagen density of the matrices does not influence the spatial distribution of individual fibroblasts over time (Del Amo et al., 2017). Our model results suggest that, even though the increased viscosity values may have reduced cell velocity as the cells faced more difficulties to move through the matrix (Petrie and Yamada, 2015), the enhanced directionality imposed by the matrix fibres produced similar cell spatial distributions (Dickinson et al., 1994). It was apparent that fibroblasts adopted more directed and more restricted migration patterns, which we assume to be due to the confinement levels produced by smaller pore sizes, and increased the magnitude of the cell-generated forces (Dickinson et al., 1994; Miron-Mendoza et al., 2012). In other words, whereas cells migrated longer distances in low density matrices, their motility patterns were random. Therefore, the resulting displacement values (i.e., the difference between the cell's final and initial position) were similar to those seen at high collagen density values.

Our model results suggest that the chemical stimulation of PDGF-BB still ocurred in spite of the increased matrix density caused by higher collagen concentrations. However, we remark that, even though the addition of PDGF-BB still enhanced fibroblast motility through an increase in the

travelled distance values, this effect was not as pronounced as that seen in fibroblasts seeded in collagen matrices of lower density. Taking into account that the increased collagen density leads to differences in the structure of the collagen matrices, and, particularly, produces smaller pore sizes (Kamiyama et al., 1998), we hypothesize that the diffusion of PDGF-BB may have been affected by these differences. In fact, it is expected that the diffusion coefficient of PDGF-BB decreases in matrices with reduced pore sizes (Moreno-Arotzena et al., 2014). Therefore, as the diffusion of PDGF-BB must have been impaired and slowed down by the physical properties of the matrix, fibroblasts probably presented a delayed and weaker response to the presence of chemical gradients.

Even though our model produced good results that were very similar to the distance distributions measured experimentally (BC > 0.93), it presents some limitations. We primarily modelled the presence of chemotactic gradients and their effects through the initial values of the model's input parameters. Accordingly, we assumed that the model response to these factors was constant over time, which did not fully capture the reality of our experiments. The diffusion of PDGF-BB is a dynamic process. Thus, it is expected that the chemical gradients may evolve over time and space. For instance, cells spatially closer to the PDGF-BB source should be more sensitive to the chemotactic substance than cells that are farther from the source, and thus sense smaller PDGF-BB concentrations. Similarly, the diffusion of these chemical factors in collagen matrices may be heterogeneous, PDGF-BB molecules may degrade or become attached to the proteins of the matrix (Moreno-Arotzena et al., 2014; Griffith and Swartz, 2006). Therefore, distinct regions of the microfluidic chip may present higher PDGF-BB concentrations, which would have a local effect on fibroblast motility. This also applies to the model's inability to properly reproduce the experimental results obtained in Condition 2, where fibroblasts were seeded in microfluidic devices with 3 cell chambers. In this case, it would have been especially crucial to be able to simulate the delayed chemotactic response of fibroblasts, as it took more time for PDGF-BB to diffuse through the entire device and reach the cells. However, once this chemical factor reached the cells, their migration patterns were similar to those seen for Condition 1.

Lastly, we should highlight that our model predictions must be further validated through the use of techniques more suited to study cell movement at smaller time and spatial scales. For example, we confirmed our predictions on the directionality of cell movement by tracking the position of fibroblasts at a time scale of minutes and not days. These additional assays are described in Appendix B, Section B.1.6, showing that fibroblast seeded in the same conditions as our main study present migration persistence times in the order of minutes, in agreement with the model results. However, the predictions about the vertical restriction factor and force magnitude should be validated. The former could be achieved by actually quantifying cell movement in the vertical (z) direction, while the latter entails the use of higher-end technologies. Nonetheless, we believe that our framework is a valuable tool to make these kinds of predictions to identify the main regulators of an observed experimental behaviour and plan future experiments accordingly.

3.5 Summary

In this chapter, we implemented the first building block of this dissertation through the design of computational models that simulate single-cell motility. Starting with a simple model that simulated cell motility as the combination of cell-generated and drag forces, we could estimate the distribution of cell locomotive forces based on experimental data published in the past and were able to predict that matrix density regulated and hindered single-cell migration. Moreover, we extended this model to include a more realistic representation of cell-generated forces and how they are influenced by external stimuli of the microenvironment. Using an automatic optimization based on Bayesian statistics, we were able to infer which model parameters played a significant role to reproduce experimental data. Specifically, our computational approach was able to identify that cell persistence and directionality were influenced by the presence of chemical factors and that cells exerted stronger forces in response to chemical stimuli and high matrix density. Henceforth, these results lead to the design of new experiments that confirmed the model predictions in regard to the cell persistence times, highlighting the relevance of combined experimental-computational approaches to unravel biological behaviour.

It is also interesting to compare the outputs obtained for each application covered in this chapter. When examining the models' responses to increases in collagen density, it is clear that the two applications displayed distinct biological behaviours. In the first scenario, it was assumed that the ECM had no effect on the cell-generated force generator. In the second case, we added a degree of freedom by defining a variable that controlled the amplitude of locomotive forces, and we discovered that our model optimization process favoured an increase in this value when collagen density rose as well. Given that the first example relies on qualitative examination of the experimental data, we are unable to draw a firm conclusion as to why these disparities were identified. Possibly, the assumption first example oversimplified, and better results could have been achieved using a dynamic magnitude value. However, these results were satisfactory. Thus, given that this adaptation was not essential in the first case, this could be representative of biological differences between the two cell lines used (NSCLC and NHDF). Fibroblasts tend to present a mesenchymal migration mode that is heavily influenced by the surrounding environment and that relies on creating cell-matrix adhesions through which cells can sense the properties of the ECM and modulate their response accordingly. In this case, it makes sense to consider the force magnitude as a function of matrix density. Conversely, amoeboid-like cells move through the matrix without adhering to it or changing their force response. Therefore, their motility is hindered when steric hindrance imposed by matrix fibres is a limiting factor to cell movement. We propose that the cells studied in the first example relied on a primarily amoeboid phenotype and, thus, good results were achieved through a general force generator function independent of matrix density. In future experiments, it would certainly be relevant to confirm this prediction.

Overall, this chapter served as a fundamental starting point to this dissertation, and, in the chapters to follow, we will extend the computational model described here to account for more complex phenomena. Furthermore, we established the relevance of parameter estimation routines,

which will be discussed throughout this document, and we explored how advanced microfluidic-based experimental configurations allow us to obtain data to calibrate our models, a topic that will also be addressed in detail in Chapter 4. It must also be stated that the optimization pipeline developed in this chapter is model agnostic and can be modified to optimize other models, including higher complexity models as those to be studied in the chapters to follow.

CHAPTER 4

Tumour growth

Contents

4.1 Introduction 62
4.2 Traditional models of tumour growth 63
4.3 Modelling tumour spheroid growth in collagen matrices 66
4.3.1 Study outline 66
4.3.2 Methodology 67
4.3.3 Results 71
4.3.4 Conclusions 73
4.4 A review on computational models of cancer glucose metabolism 75
4.4.1 Glucose metabolism in cancer cells 75
4.4.2 Modelling cellular systems and glucose metabolism 79
4.4.3 Future perspectives 83
4.5 Summary 84

This chapter is based on:

Inês Godinho Gonçalves and José Manuel García-Aznar. **"Extracellular matrix density regulates the formation of tumour spheroids through cell migration."** PLoS computational biology 17.2 (2021): e1008764.

Inês Godinho Gonçalves and José Manuel García-Aznar. **" Hybrid computational models of multicellular tumour growth considering glucose metabolism."** Computational and Structural Biotechnology Journal 21 (2023): 1262-1271.

4.1 Introduction

Tumours arise from cancer cells with abnormal proliferative potential which divide and form multicellular structures. A particular area of study that has yet to be understood in cancer development is related to metastatic colonization, i.e., how metastatic cells adapt to survive in a new tissue (Massagué and Obenauf, 2016). Although the large majority of circulating individual tumour cells perish before they are able to produce a secondary tumour, some can invade new tissues and replicate (Massagué and Obenauf, 2016; Nagrath et al., 2007). In the last few years, studies have shown that this ability to survive and proliferate in new tissues depends not only on cells but also on the biophysical microenvironment of the metastatic niche and the cell-matrix interactions (Fidler, 2003). Thus, it is of high importance to understand and characterize how tumour cells respond to different compositions of the microenvironment, to identify the causes that may induce a more invasive phenotype and to develop therapeutic strategies that hinder this behaviour (Massagué and Obenauf, 2016; Carey et al., 2012).

Several studies have studied tumour cells seeded in 3D scaffolds to replicate the initial stages of growth in avascular tumours (Nyga et al., 2011), creating a setting that can also reflect the initial stages of metastatic colonization in new tissue. There are several available techniques used to produce multicellular structures from tumour cells (Lv et al., 2017). multicellular tumour spheroids (MCTSs), for example, can be formed by growing tumour cells in solutions or non-adhesive substrates and can later be introduced in 3D scaffolds (Reynolds and Weiss, 1992; Mueller-Klieser, 1987). However, individual cells may also be directly seeded in 3D matrices or microfluidic devices to understand how cells organize based on the matrix characteristics (Breslin and O'Driscoll, 2013; Zhang and Nagrath, 2013). In comparison to more traditional cell culturing techniques, e.g., cultures grown in 2D or in suspension, MCTSs provide a more realistic histological and functional depiction of solid tumours and their surrounding microenvironment as observed in *in vivo* (Pampaloni et al., 2007).

MCTSs provide a realistic representation of the avascular growth states of tumour growth (Pampaloni et al., 2007) Yet, it is still unclear how the mechanical properties of 3D scaffolds modulate tumour growth. On the one hand, it is well established that matrices of higher density suppress growth by exerting compressive forces on cells, hence producing smaller tumours than those grown in matrices with low densities (Cheng et al., 2009; Loessner et al., 2013; Helmlinger et al., 1997). Consequently, the tumour size tends to decrease as the density of the matrix increases. In general, these experiments have used low-porosity matrices. On the other hand, experimental studies have also shown that matrices with higher density tend to limit cell movement as cells are unable to migrate through the matrix due to steric hindrance (Ehrbar et al., 2011; Plou et al., 2018; Wolf et al., 2013; Cóndor et al., 2019). In turn, this situation promotes individual cell migration in matrices composed of lower collagen concentrations, causing cells to stray from their original cluster, which can subsequently affect tumour size (Plou et al., 2018; Mark et al., 2020; Haeger et al., 2014), as represented in Fig 4.1. Furthermore, the chemical species of the tumour microenvironment have also been shown to modulate cancer growth, specifically glucose, which

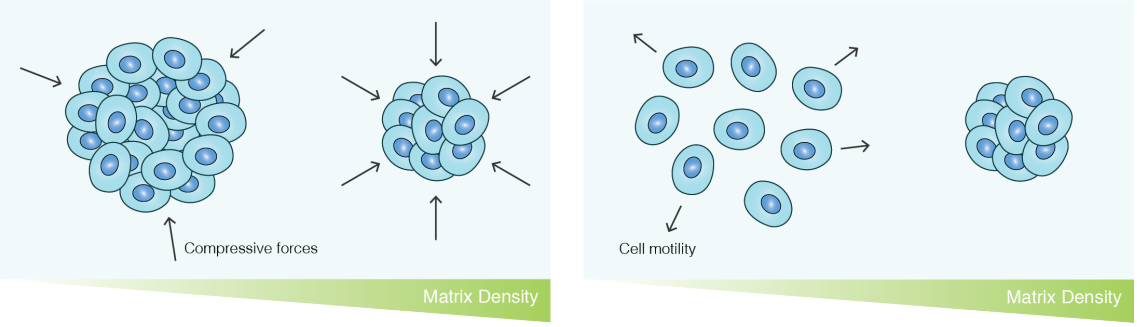

Figure 4.1: **The effect of matrix collagen density on tumour growth.** The left subfigure represents the common assumption that high-density matrices inhibit tumour growth due to compressive forces that lead to cell arrest. The right subfigure shows that, in high-porosity matrices, individual cells are able to move freely when collagen density is low, resulting in individual and sparse cells and no detectable tumours.

is associated with metabolic reprogramming (Casciari et al., 1988, 1992b; Rousset et al., 2022).

This chapter aims to analyse the biological behaviours and external factors that induce tumour growth, specifically in the avascular development stages, and how thee features can be simulated through ABMs. Therefore, we firstly provide a brief overview of traditional models of tumour growth implemented with continuum methods, showcasing how they evolved over time to account for more detailed spatial descriptions of tumour development. However, these approaches are less powerful than CBMs to model cell heterogeneity, which is a topic that will be explored in detail in this chapter. As done in Chapter 3, we investigate the role of mechanical and chemical stimuli present in the tumour microenvironment. In particular Section 4.3 describes the implementation of a model extension to the model described in Chapter 3, Section 3.3, to consider cell-cell mechanics and the cell cycle. Using previously published data of tumour spheroid growth in microfluidic devices, we were able to qualitatively calibrate our model to reproduce the empirical growth curves and we could make predictions about how spheroid formation was regulated by the collagen density through cell migration. Moreover, Section 4.4 aims to showcase how previous computational tumour growth models were employed to investigate nutrient availability and metabolic reprogramming. Although we did not implement glucose metabolism in our computational approach, this review allowed us to understand the current literature on computational models of tumour growth kinetics that consider glucose metabolism. Moreover, we could identify strategies through which we could include these representations in our models.

4.2 Traditional models of tumour growth

For several decades, scientists have relied on mathematical and computational models to understand and reproduce tumour growth and several frameworks have been developed to describe the spatial dynamics of tumours and their microenvironment at different complexity levels (Rejniak and Anderson, 2011; Altrock et al., 2015; Schaller and Meyer-Hermann, 2006; Araujo, 2004; Kansal et al., 2000; McEvoy et al., 2020). Some of the simplest models developed were based on

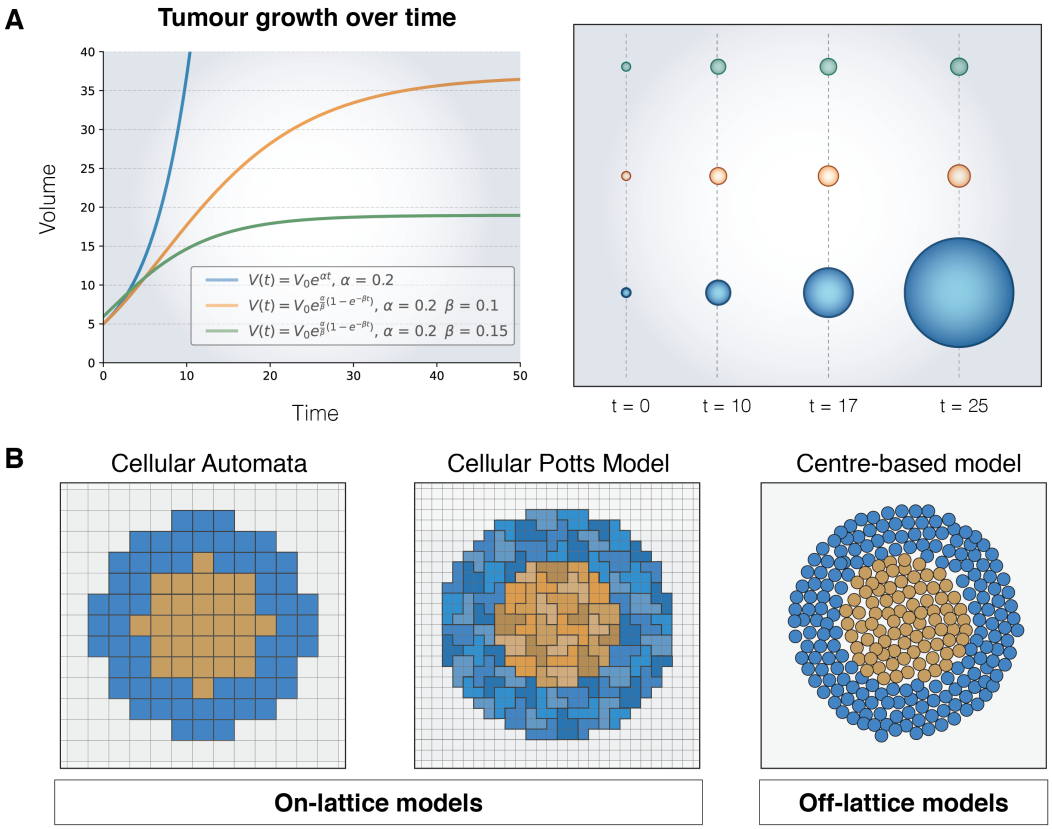

Figure 4.2: **Continuum and discrete models of tumour growth.** (A) Representation of simulated growth curves of avascular models obtained with classical continuum models based on ODEs, namely the exponential (blue) and Gompertzian (green/orange). Although these models provide similar results at early stages of growth, only the Gompertzian model is able to reproduce the saturation in tumour size that is commonly attributed to limited nutrient diffusion once a tumour reaches a critical size. (B) Examples of discrete, or agent-based, models, which, unlike continuum approaches, model tumours a group of individual agents that follow a set of rules that characterize biological phenomena such as proliferation and death. Proliferative cells are shown in cells while yellow cells represent the necrotic core. In the CPM subfigure, different shades of the same colour represent distinct cells since a single cell may occupy more than one voxel.

continuum approaches that aimed to replicate the evolution of the number of cells or the size of a tumour spheroid over time, using ODEs and PDEs (Gerlee, 2013). ODEs can be employed to model macroscopic tumour growth curves under the assumption that tumours are composed of a single, spatially homogeneous cell population (Bull et al., 2020). Several growth laws have been used to this aim, as illustrated in Fig 4.2A. For example, tumour growth can be modelled through an exponential law that describes growth as being proportional to tumour size, represented by its volume, V, as written in Eq 4.1.

$$\frac{dV}{dt} = \alpha V, \text{where } V(t=0) = V_0 \tag{4.1}$$

Here, α is a growth constant that can be fitted to match a given cell population using experimental

data and V_0 is the initial tumour volume. Exponential growth is an adequate model to reproduce the initial stages of growth in a tumour spheroid, when a single cell originates two daughter cells (Murphy et al., 2016). Nonetheless, it fails to capture the subsequent stages where growth becomes arrested due to the limited amount of nutrients and the increased concentrations of metabolic waste that induce cell arrest and death (Kunz-Schughart et al., 1998; Casciari et al., 1992b).

The inhibitory effect caused by the limited diffusion can be modelled by modifying the exponential growth law and considering logistic or Gompertzian growth instead (Murphy et al., 2016; Bull et al., 2020; Benzekry et al., 2014). In the 1960s, Laird (1964) showed that the growth curves of several tumour types could be described by the Gompertzian equation, a generalization of the logistic growth law that reproduces an initial fast growth phase, followed by an exponential decrease in the tumour's growth rate that results in tumour size saturation. This model can be defined by defined by Eq 4.2:

$$\frac{dV}{dt} = \alpha e^{-\beta t} V, \text{where } V(t=0) = V_0 \tag{4.2}$$

where β defines the rate at which growth decays (Benzekry et al., 2014; Norton, 1988). It has been shown to fit experimental data better than exponential and logistic models (Vaghi et al., 2020).Nevertheless, growth laws modelled as ODEs are not able to capture the complex tumour internal spatial organization and structure (Benzekry et al., 2014; Schaller and Meyer-Hermann, 2006). Consequently, more sophisticated models were adapted to capture the internal spatial organization of tumours and describe growth as a result of insufficient nutrient diffusion that leads to cell death and quiescence (Greenspan, 1972; Burton, 1966).

PDEs have been used to simulate the effect of the spatial distribution of diffusing factors that promote or inhibit cell growth and how they modulate the growth rate of tumour cells locally. In the 1970s, Greenspan (1972) developed a model of tumour growth that took into account how a single diffusing factor, here assumed to be glucose, influences tumour size. Assuming radial symmetry and that $V(t) = \frac{4}{3}\pi R^3(t)$ where R defines the tumour radius, the Greenspan model can be defined by Eq 4.3:

$$\frac{dR}{dt} = \frac{1}{R^2}\int_0^R f(c) r^2 dr \tag{4.3}$$

where c is the local concentration of glucose and $f(c)$ is a function that describes how the cells' doubling rate changes in function of glucose concentration. Given that glucose is a growth-promoting factor, $f(c)$ should consider that the local concentration of glucose should increase the cells' doubling rate, until a maximum value is reached, where an increase in glucose concentration no longer has an effect on the doubling rate (Byrne, 2010). The spatial distribution of glucose can be modelled as a reaction-diffusion equation and it is assumed that glucose concentration decreases from the tumour surface to its core. Furthermore, the glucose consumption rate can be adjusted to consider different phenotypes, such as proliferating, quiescent and necrotic (Greenspan, 1972; Burton, 1966). In addition, the Greenspan model can be extended to consider other substances and reproduce differential behaviours accordingly (Bull et al., 2020).

The aforementioned continuum-based models have proven to provide realistic descriptions of tumour growth kinetics at the avascular stages. Nevertheless, these mathematical approaches fail to undisclosed how tumour growth arises from single-cell behaviour. On the other hand, discrete models describe cells at the individual level and are thus more suited to reproduce the heterogeneous behaviour of biological systems (Metzcar et al., 2019) and the interactions between cells and the microenvironment (Mathias et al., 2022; Macnamara, 2021; Gonçalves and Garcia-Aznar, 2021). These models simulate cells as individual agents that follow a set of rules that define their cellular behaviour (e.g., death, proliferation, migration) and how they interact with other cells and the surrounding microenvironment (Wang et al., 2015). ABMs can be grouped into on-lattice and off-lattice models. Fig 4.2B shows an illustration of some of the most commonly used on-lattice and off-lattice models to simulate tumour growth. On-lattice approaches (e.g., CA and CPMs) divide the domain into a grid of cells or voxels, and cells are bound to occupy these specific positions, which makes them computationally efficient (Bull et al., 2020). Conversely, off-lattice approaches such as CBM and vertex models provide more realistic representations of cells, yet they can be considered to be computationally expensive to simulate large tumours (González-Valverde and García-Aznar, 2017; Rejniak and Anderson, 2011). All of the aforementioned approaches have been used to simulate tumour growth as reviewed in more detail in (Bull et al., 2020; Macnamara, 2021; Wang et al., 2015).

4.3 Modelling tumour spheroid growth in collagen matrices

4.3.1 Study outline

The work covered in this chapter describes the implementation of a model extension to the model presented in Chapter 3, Section 3.3, which took into account cell-cell interactions and cell proliferation/death. Therefore, at this stage, a new complexity level was added to our computational approach that now combines the effect of single-cell motility and drag-imposed forces, already characterized in the previous chapter, with cell-cell dynamics that need to be evaluated. This work was also inspired by the experimental studies developed by Plou et al. (2018), where the authors seeded single cells in collagen matrices embeded in a microfluidic setup during 7 days and characterized how collagen density affected cell organization and tumour growth. The authors concluded that the matrices with low concentrations of collagen enabled cells to migrate more freely, the increased motility produced sparser and smaller clusters. In contrast, the restrained motility of cells seeded in matrices of higher collagen density resulted in larger spheroids. Accordingly, it can be stipulated that even though the invasive capability is low, cell surveillance is enhanced by denser matrices, as cells aggregate and form stable structures that evolve into secondary tumours (Sapudom et al., 2015; Aceto et al., 2014).

In this adapted experimental configuration, Plou et al. (2018) allowed cells to proliferate for several days to enable tumour growth. Cell motility metrics were not tracked, and the focus shifted to quantifying how the tumour spheroid size evolved over time. Thus, cluster areas were

Table 4.1: **Reference parameter values for the tumour model.**

Symbol	Parameter	Value	Unit	References
Δt_{mech}	Mechanics time step	0.1	min	Ghaffarizadeh et al. (2018)
Δt_{cell}	Cell processes time step	6.0	min	Ghaffarizadeh et al. (2018)
$\rho^0_{Collagen}$	Collagen Concentration	[2.5, 4.0, 6.0]	mg/mL	Plou et al. (2018)
μ	Drag Coefficient	[7.96, 18.42, 39.15]	Pa · s	Valero et al. (2018)
R	Cell Radius	8.4	μm	Ghaffarizadeh et al. (2018)
R_A	Cell Adhesion Distance	$1.25R$	μm	Ghaffarizadeh et al. (2018)
c_{cca}	Adhesion Coefficient	7.2	-	Ghaffarizadeh et al. (2018), estimated
c_{ccr}	Repulsion Coefficient	380	-	Ghaffarizadeh et al. (2018), estimated
T_{Ki67-}	Cell Quiescence Time	6.5	h	Ghaffarizadeh et al. (2018), estimated
T_{Ki67+}	Cell Proliferation Time	15.5	h	Ghaffarizadeh et al. (2018), estimated
r_D	Cell Death Rate	0.00319	1/h	Ghaffarizadeh et al. (2018)

tracked at time points of 1, 3, 5 and 7 days, and the obtained results were compared between the different matrix density values. The authors also analysed how the cluster eccentricity, measured at day 7, may vary according to the microenvironment. Therefore, in our computational approach, we introduced the new model components to account for the new components that needed to be reproduced (i.e., cell mechanics and proliferation/death) and, subsequently, we reproduced the experimental setup and its associated post-processing procedures.

4.3.2 Methodology

Modelling framework

As previously stated, in this work we intended that our computational framework combined the forces that were previously characterized for the single-cell motility model and extend it to take into account cell-cell adhesion and repulsion and cell cycling. The former could be achieved by updating the equation of motion to detect and compute cell-cell adhesion and repulsion forces ($\mathbf{F}^{ij}_{cca}$ and $\mathbf{F}^{ij}_{ccr}$), respectively), resulting in Eq 4.4.

$$\mathbf{0} \approx \sum_{j \in \mathcal{N}(i)} (\mathbf{F}^{ij}_{cca} + \mathbf{F}^{ij}_{ccr}) + \mathbf{F}^{i}_{drag} + \mathbf{F}^{i}_{loc} \tag{4.4}$$

PhysiCell custom functions were used to detect cells inside a radius of interaction, R_A, at which cell-cell adhesion started to occurr. Cell-cell repulsion was present only when cells were in contact with each other (Ghaffarizadeh et al., 2018). Model parameter values can be found in Table 4.1.

Defining cell-cell interactions

Cell-cell interactions are characterized by potential functions as implemented by Ghaffarizadeh et al. (2018) and are regulated by adhesion and repulsion coefficients, C_{cca} and C_{ccr}, respectively (see Table 4.1). In particular, adhesion and repulsion forces between two cells i and j, separated by a distance of (x_i - x_j), represented as $\mathbf{r}$, are given by Eqs 4.5 and 4.6:

$$\mathbf{F}_{cca}^{ij} = C_{cca} \nabla \phi(\mathbf{r}) \tag{4.5}$$

$$\mathbf{F}_{ccr}^{ij} = C_{ccr} \nabla \psi(\mathbf{r}) \tag{4.6}$$

for which $\nabla\phi$ and $\nabla\psi$, respectively, can be written as:

$$\nabla \phi(\mathbf{r}) = \begin{cases} \left(1 - \frac{|\mathbf{r}|}{R_A}\right)^2 \frac{\mathbf{r}}{|\mathbf{r}|} & \text{if } |\mathbf{r}| \leq R_A \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{4.7}$$

$$\nabla \psi(\mathbf{r}) = \begin{cases} -\left(1 - \frac{|\mathbf{r}|}{R}\right)^2 \frac{\mathbf{r}}{|\mathbf{r}|} & \text{if } |\mathbf{r}| \leq R \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{4.8}$$

where R_A represents the maximum distance between two cells at which adhesion forces are present, and R is the maximum distance for which repulsion forces are present. R is, simultaneously, the radius of a single cell. Below these values, interaction forces between cells are assumed to be null.

Defining cell events

Unlike the previously described setup, cells were allowed to proliferate and undergo apoptosis. Specifically, it was considered that cells could either be in a quiescent or a proliferative state, which would correspond to Ki67- and Ki67+, respectively Warth et al. (2014). Taking into account the cell state for a cell i, represented as $\mathcal{S}_i$, the probability of the cell in a Ki67- phase entering a Ki67+ phase in a time interval $[t, t+\Delta t]$ can be given by:

$$\text{Prob}(\mathcal{S}_i(t+\Delta t) = \text{Ki67+} | \mathcal{S}_i(t) = \text{Ki67-}) = 1 - e^{-\frac{1}{T_{\text{Ki67-}}}\Delta t} \approx \frac{\Delta t}{T_{\text{Ki67-}}} \tag{4.9}$$

where $T_{\text{ki67-}}$ represents the time spent in the Ki67- state. Similarly, the probability of a proliferative cell to enter a quiescent state is:

$$\text{Prob}(\mathcal{S}_i(t+\Delta t) = \text{Ki67-} | \mathcal{S}_i(t) = \text{Ki67+}) = 1 - e^{-\frac{1}{T_{\text{Ki67+}}}\Delta t} \approx \frac{\Delta t}{T_{\text{Ki67+}}} \tag{4.10}$$

Regarding cell death for both states, the probability of a cell entering apoptosis, represented by a state D, is given by the death rate, r_D, as described in Eq 4.11.

$$\text{Prob}(\mathcal{S}_i(t+\Delta t) = D | \mathcal{S}_i(t) = \text{Ki67-/+}) = 1 - e^{-r_D \Delta t} \approx r_D \Delta t \tag{4.11}$$

The time steps (Δt) used in our simulations are coherent to those originally defined by *PhysiCell* (Ghaffarizadeh et al., 2018). Accordingly, the time step used for cell processes such as described above, namely cell death and cell cycle, is in the order of minutes (6 min), whereas the time steps used for the diffusion and mechanical analyses are, respectively, 0.01 and 0.1 min. *PhysiCell* also includes the possibility of modelling proliferation rules as a dynamic process dependent of

Table 4.2: **Parameter values for the post-processing of tumour growth modelling results.**

Parameter	Values	Unit	References
z-Height of Interest	[-48, 48]	μm	Estimated
DBSCAN Radius	18	μm	Estimated
DBSCAN Minimum Number of Cells	3 (2 at day one)	-	Estimated

substance concentration. However, we assumed that the experimental settings corresponded to a well-oxygenated environment, assuming conventional cell culture oxygen values as well as the high permeability of PDMS-based microfluidic chips (Thomas et al., 2011). Furthermore, given that the cell concentration was not high, we considered that oxygen's effect on spheroid formation could be disregarded. Nonetheless, an extended study of the oxygen levels in our simulations is provided in Appendix B, Section B.2.2, where it is shown that hypoxia values were not reached.

Cluster classification and metrics

However, we adapted this procedure to our data and decided to automate this process since we needed to iteratively run simulations and obtain cluster metrics to assess how cell-cell interactions could be calibrated to obtain better results. To obtain comparable data representations to those presented by Plou et al. (2018), we used the x and y coordinates of cells present in a height-defined z-region. We only used a section of the z-axis to obtain a similar selection of cells to what would be expected from a confocal microscopy analysis, disregarding cells that would be out of focus. With this spatial data, we used an implementation of the density-based spatial clustering of applications with noise (DBSCAN) algorithm (Ram et al., 2010) to classify each cell into a cluster or identify it as an outlier. In particular, we used the implementation offered by scikit-learn (Pedregosa et al., 2012), a publicly available Python module for machine learning algorithms. DBSCAN is a distance-based algorithm that requires users to define the minimum number of cells in a spheroid and a radius of interest. Consequently, we inferred these values, which are presented in Table 4.2, from experimental observations. An extended sensitivity analysis of the effect of these parameters on our results can be found in Appendix B, Section B.2.1, which also includes a brief study on how area values would vary if other area quantification methods were used (e.g., smallest enclosing circle/ellipse and convex hull).

Having defined the clusters, which were visually assessed to avoid misclassifications, the area of each cluster was computed. Here, we aimed to replicate the procedure used by Plou et al. (2018), in which the cluster area is computed by calculating the area of the circle that best fits each cluster. Accordingly, we started our analysis by calculating the centroid for each cluster, which we defined by the average x and y coordinates of all the cells that formed the cluster. Subsequently, we computed the distance of each cell to that point, which we used to estimate the radius of the cluster by calculating their average value. Using the estimated cluster radius, we computed the area for a circular geometry.

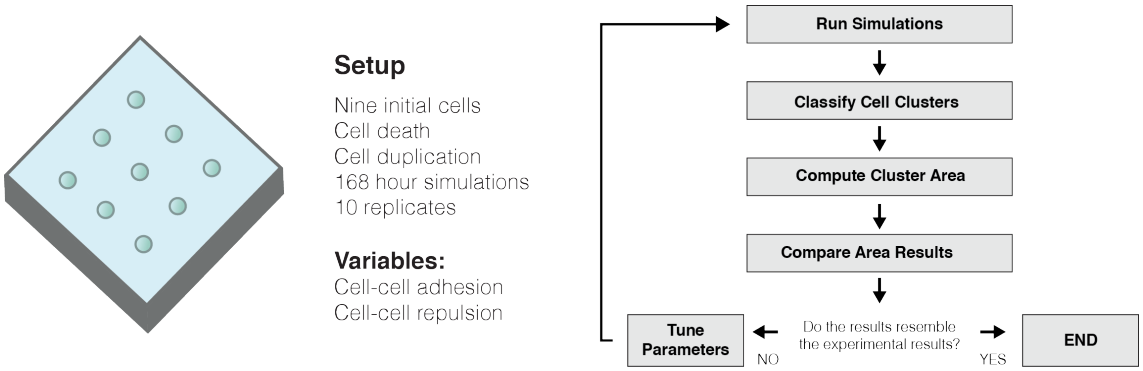

Figure 4.3: **Simulation and calibration workflow for the single-cell motility model.** Single cells were initialized and their trajectories were tracked and recorded. Mean and effective velocities were computed from the recorded spatial data and manual calibration was performed to find the model input parameters that produced the best results in comparison with the available qualitative empirical data.

Last, after seven simulated days, we also computed cluster eccentricity by identifying an equivalent ellipse with the same second-moments as the region defined by the cell centres of each cluster. We calculated ellipse eccentricity as the ratio of the distance between the foci of the ellipse and its major axis length, which we assumed to be representative of the cluster eccentricity. In contrast to our previous strategy, we analysed the entire height of the domain, as we observed that at this point of development, the geometry of the clusters did not significantly change based on the selected subsection. Since the experimental eccentricity values considered only spheroids with an area larger than 1000 m^2 and were obtained through a semi-automatic classification algorithm, the computational results were also manually evaluated to disregard clusters with small area values or that were misclassified by the automatic algorithm.

Simulation workflow

For this setup, simulations were run for seven simulated days and recorded at time points of 24 hours, with a total of 10 replicates for each matrix density value. At time points of 24, 72 and 120 hours, cells were classified into clusters. Experimentally, the results were obtained through a manual classification procedure, using an image-based piece of software. Several microscopy images were taken, representing different slices of the microchip height. Subsequently, the subregion that better captured the area of the clusters was chosen and used as input for a classification algorithm. The authors selected the multicellular clusters' approximate locations in the images of interest, and the algorithm identified the regions of interest and computed the cluster area by fitting a circle to the detected structure. Hence, as presented in Fig 4.3, we aimed to calibrate these cell-cell interaction parameters by analysing the formation of multicellular clusters, and their subsequent growth, as depicted by the experimental quantification of cell area values. Accordingly, we placed nine initial cells in the domain at a distance that would enable individual cluster growth, without promoting the junction of two different clusters, as this was not observed experimentally.

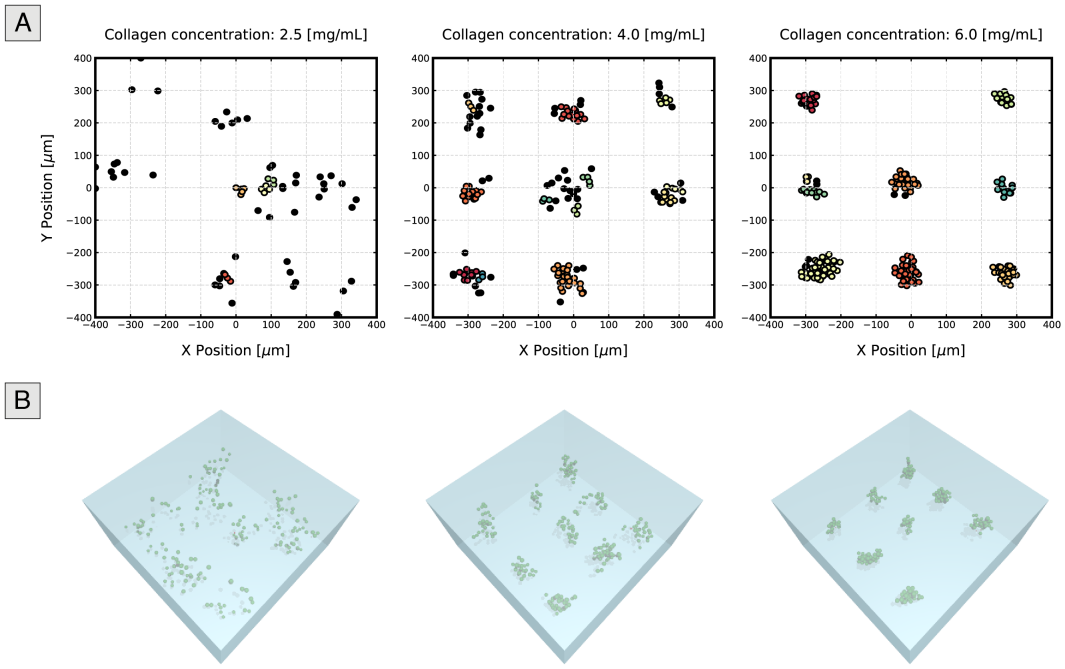

Figure 4.4: **Representation of cell positions after five simulated days of tumour growth.** 2D (top) and 3D (bottom) representations of the coordinates of cells grown in matrices of different collagen concentrations after five days. For the 2D scatter plots, we selected only cells present in a defined height of interest to remove cells that would otherwise be out of focus in microscopy images. A single replicate was chosen for each condition to produce these plots. Different colours represent different clusters, whereas black cells are those considered outliers (i.e., they do not belong to any of the cell groups). The cluster area increases with density, as cells stay closer to their original position. On the other hand, in low collagen density matrices (left), in which individual cell migration is not limited, individual cells are seen to stray away, resulting in a large number of outliers and smaller, sparser tumours.

4.3.3 Results

Low-density matrices produce smaller structures due to cells' ability to migrate

In the original experiments, Plou et al. (2018) observed that tumour cells seeded in high-density collagen matrices produced smaller multicellular structures than those seeded in low-density environments. Similar to the experimental conclusions, the results of our simulations describe how tumour size was affected by matrix density, as seen in Fig 4.4, which presents the cell positions in the 2D plane at the end of the fifth simulated day for each matrix density value. Coloured circles represent cells that belong to a cluster, with each colour corresponding to a different cell group, whereas black cells represent outliers. Based on this figure, we hypothesized that cells seeded in matrices with high collagen density tend to be more confined and form clusters with higher area values than those grown at low densities, in which barely any spheroids were present.

Accordingly, an interesting dichotomy between migration and tumour size could be observed: for conditions where cell migration was allowed, the multicellular clusters experienced less growth,

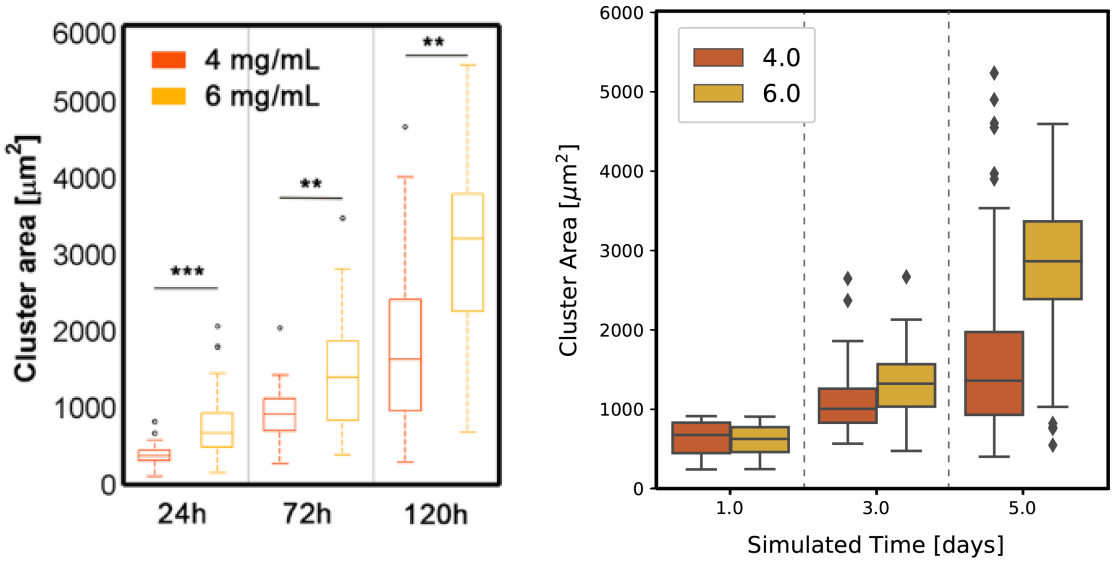

Figure 4.5: **Evolution of cluster area growth over five simulated days.** Distribution of cluster areas at the end of days one, three and five for clusters grown in matrices of medium and high collagen density (4.0 and 6.0 mg/mL, respectively), for the experiment (left) and computational (right) settings. These distributions take into account data from all five replicates. Cells seeded in collagen matrices of low density (2.5 mg/mL) did not show significant multicellular cluster formation and growth and hence were not represented. An increase in tumour size through time can be seen for both the medium and high collagen concentrations, but larger densities induce the formation of clusters of larger areas.

and vice-versa. Previous studies have proposed a possible explanation for the interplay between migration and tumour growth through the "Go-or-Grow" mechanism, which may depend partially on the physical constraints of the ECM (Giese et al., 2003; Hatzikirou et al., 2012). According to this mechanism, migration and proliferation are temporally exclusive: migrating cells are not able to proliferate, resulting in clusters of a smaller size. Nonetheless, it is still unclear whether these two phenomena are not coincident in time (Garay et al., 2013). Interestingly, our results suggest that there was no need for such a mechanism under the studied conditions since we did not implement proliferation suppression in migrating cells. Henceforth, we highlight the difference between tumour growth and cell proliferation and we recognize the effect of individual cell migration on the former, but not necessarily on the latter.

We could confirm that cluster size was enhanced in high-density matrices by classifying and quantifying the area of each cluster. Fig 4.5 depicts the time evolution of cluster areas for matrices of medium and high density at days one, three and five and indicates that our model not only captured cluster area accurately after five days of growth but it could describe the evolution of tumour size through time. Only values for medium and high densities are shown because the number of clusters with a significant area in matrices of low density is not significant.

Cluster eccentricity decreases in response to collagen density

Plou et al. (2018) also computed cluster eccentricity and evaluated these values as a result of matrix density. In these experiments, it was observed that low-density matrices resulted in elongated clusters, whereas high density values favoured the assembly of spheroids with a rounded morphology. Conversely, simulated clusters adopted a relatively round morphology, further promoted in high-density matrices, as shown in Fig 4.6, which presents the results obtained for cluster eccentricity at seven days of simulation. As a general trend, we were able to conclude that eccentricity values increased in low-density matrices due to cell motility. We stipulate that cells started to stray from the original cluster but, ultimately, due to a decrease in cell speed, reconnected to the original structure as cells proliferated. Consequently, more elongated clusters started to appear.

When compared with the experimental values, our values were consistently higher than those computed in the experimental study. We conducted an investigation to understand why this was occurring, which can be found in Appendix B, Section B.2.1, and we attribute this outcome to the intrinsic differences between the two datasets and the used algorithms for cluster classification, as we observe the same trends in each setup. It must be taken into account that the experimental measurements relied on semi-automatic imaging methods, which are pixel-based and, consequently, have lower sensitivity to perturbations in cell coordinates. These small differences can greatly influence the eccentricity values of clusters, especially when considering rounder structures, as we explain in more detail in Section B.2.1. Furthermore, the experimental methods focused more on the contours of the multicellular cluster, whereas we based our computational results on cell coordinates, without taking into account cell geometry. Nonetheless, we consider that the model captured the general dynamics of the experimental setup.

4.3.4 Conclusions

In this study, we used model of single-cell motility that took into account cell-generated and drag forces and extended it to take cell-cell interactions and cell cycling into account. We intended to employ this computational to evaluate the formation of multicellular clusters and identify which of the aforementioned model components were the most relevant to this biological behaviour. Accordingly, we relied on previously published data of experiments where tumour cells were seeded in collagen matrices integrated in microfluidic devices and the formation of multicellular structures from individual cells was studied over time. In general, we showed the model's potential to simulate and describe the formation of tumour spheroids, integrating the biomechanical role of the matrix, which was mediated by its collagen concentration. As a consequence, our approach could qualitatively predict spheroids growth within different collagen densities as a direct result of cell motility. Moreover, we showed that this relationship between cluster size and individual cell migration, often attributed to suppression in cell proliferation of migrating cells, could be achieved without changing the rates of cell proliferation (Giese et al., 2003; Hatzikirou et al., 2012). Accordingly, our results aligned with other studies that do not consider motility and proliferation to

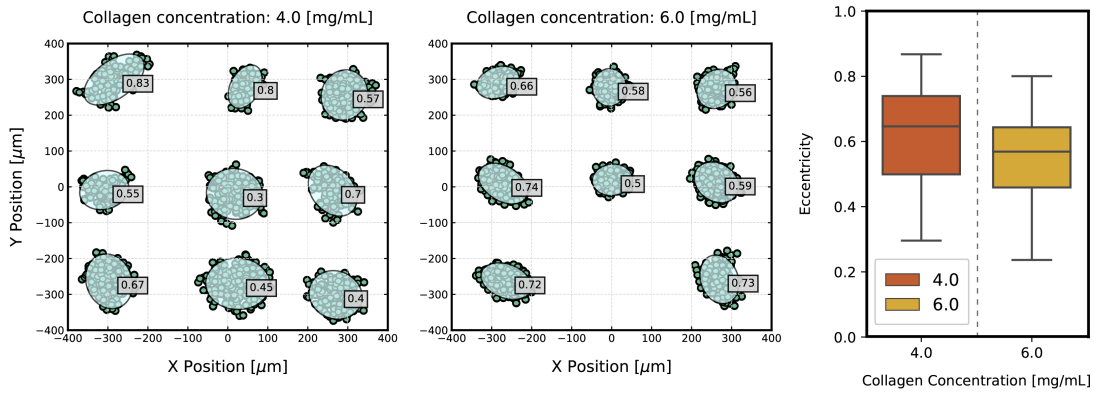

Figure 4.6: **Cluster eccentricity after 7 days of growth.** Representation of the xy cell coordinates for the cells present in the entire height of the domain, overlapped by the clusters' equivalent ellipses and eccentricity values (left). Individual cells and clusters of small areas or areas that were manually evaluated as having been misclassified by the clusterization algorithm are not represented. Clusters of cells grown in collagen matrices of high density are shown to present smaller values of eccentricity (indicating a rounder morphology), while clusters grown in matrices with a lower collagen concentration adopt slightly larger values, as cells migrate away from the cluster, producing more elongated morphologies. The distribution of eccentricity values for both densities (right) further confirms this idea.

be mutually exclusive as well as with the idea that proliferative tumours can also be highly invasive (Garay et al., 2013). This effect is also of relevance in the context of metastatic colonization given that, in low-density matrices, it is less likely for a secondary tumour to develop, as cells are individualized and may die before forming a new tumour.

We recognize that our model has some limitations, as revealed by the study we conducted on cluster eccentricity. Although we were able to simulate the qualitative behaviour of the experimental data, we recognize that model refinements could be implement to provide more realistic results. We postulate that the elongated morphologies and increased eccentricity values could have also been promoted by the alignment of the matrices' collagen fibres. Nevertheless, the model did not capture this property of the collagen matrices, and the simulated cells moved in a random walk, which did not favour cluster elongation in a determined direction, but rather leads to tumour expansion in all directions. This was further promoted by the effect of cell-cell interactions, that may have counteracted individual cell movement as cells moved in different directions. Consequently, the model was not able to predict a direct correlation between increased migration and strand-like morphology, as shown experimentally. A possible strategy to try to simulate the formation of elongated clusters would be to force directional motility instead of a random walk. Nevertheless, we think that, by itself, this is likely not sufficient for producing strand-like clusters, as all cells would migrate in this direction. Hence, the clusters would be expected to move collectively, and not to deform, which, in relative terms, would produce comparable results to those presented here. Consequently, we believe that a leader-follower mechanism (Khalil and Friedl, 2010; Mayor and

Etienne-Manneville, 2016) needs to be implemented so that only some cells migrate in a determined direction, pulling on neighbour cells to force the cluster to grow in that direction.

4.4 A review on computational models of cancer glucose metabolism

Avascular tumour growth has been characterized extensively through experiments and computational models (Araujo, 2004; Casciari et al., 1992a; Altrock et al., 2015). However, there is still a need for models that can provide new insights into cancer metabolism and how it affects tumour progression. Although some mathematical models consider metabolic reprogramming in tumours (Schuster et al., 2015), these usually consider cell populations at a large scale and lack the ability to model heterogeneous behaviour at the cell-level. Therefore, strategies such as hybrid modelling can be used to couple detailed descriptions of cellular systems with models of the surrounding microenvironment and simulate biological behaviour at different scales, including intracellular metabolic pathways (Rejniak and Anderson, 2011).

In this section, we provide a brief review on computational models that consider glucose metabolism. Accordingly, we start by defining the metabolic pathways that are reprogrammed in cancer cells and how they differ from healthy cells. Having established this biological background, we go through previous examples of tumour growth models, from conventional, large-scale implementations to multiscale frameworks that explicitly simulate the targeted metabolic pathways. Lastly, we present some perspectives on the future of multiscale models to simulate glucose metabolism in cancer cells.

4.4.1 Glucose metabolism in cancer cells

Glucose is a nutrient used by both healthy and tumour cells to produce energy in the form of adenosine triphosphate (ATP) (Martinez-Outschoorn et al., 2017; Vander Heiden et al., 2009). Generally, glucose can be catabolized via two main metabolic pathways: glycolysis and oxidative phosphorylation. The former is a less efficient but faster process that can be performed under anaerobic conditions, producing lactate and 2 ATP molecules per glucose molecule. The latter is a more complex, oxygen-dependent pathway that can generate large amounts of energy (approximately 32 ATP molecules for each glucose molecule), producing water and carbon dioxide molecules as a result. In healthy tissues, differentiated cells tend to generate energy through oxidative phosphorylation and resort to glycolysis only under anaerobic conditions. Yet, in the 1920s, studies performed by Warburg (Warburg, 1956b) showed that tumour cells rely on glycolysis, even when oxygen was available, originating a theory commonly known as the "Warburg effect" or "aerobic glycolysis". Currently, it is well accepted that tumour cells reprogram their metabolism and consume glucose at high rates, as glycolysis requires more glucose molecules to produce large amounts of energy (Hanahan and Weinberg, 2011; DeBerardinis and Chandel, 2016). Nevertheless, it is still unclear why cancer cells perform aerobic glycolysis instead of the more effective process of oxidative phosphorylation.

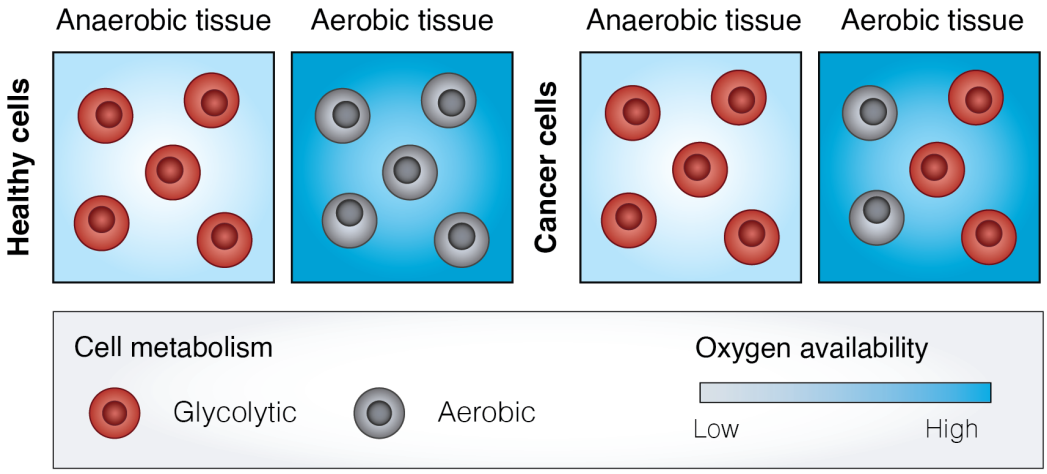

Figure 4.7: **Cell metabolism differences in healthy and cancer cells.** Healthy cells are known to metabolize glucose through glycolysis when they are in anaerobic conditions, i.e., when oxygen is not available. Conversely, healthy cells undergo aerobic respiration when they are in the presence of oxygen. On the other hand, although cancer cells also perform glycolysis in anaerobic conditions, their metabolism shifts in the presence of oxygen. When oxygen is available, some cancer cells perform glycolysis, which is not as energetically efficient, but it enables energy production at faster rates. Different tumour types have been shown to differ in their preference to shift towards glycolysis. Thus, some tumours may have a glycolytic cell population, whereas other tumour types may also present cells that still perform aerobic phosphorylation.

The findings proposed by Warburg suggesting that cancer cells undergo aerobic glycolysis were firstly attributed to defects in their mitochondria that might impair the process of oxidative phosphorylation (Warburg, 1956b,a). Later studies have shown that cancer cells are still able to oxidize glucose, though (Weinhouse, 1955). Furthermore, aerobic glycolysis and oxidative phosphorylation were proven to occur simultaneously at high rates in some tumour types (Martinez-Outschoorn et al., 2017; Pfeiffer et al., 2001), unlike what is normally observed in normal cells, which prioritize one of these metabolic pathways (Hay, 2016). This phenomenon is illustrated in Fig 4.7. Consequently, based on these results, scientists started to postulate that there might be an evolutionary advantage to this metabolic adaptation.

Glucose is commonly associated with energy production, which makes the Warburg effect seem paradoxical in the sense that glycolysis results in significantly less ATP molecules than aerobic respiration (Liberti and Locasale, 2016). Nonetheless, previous studies have shown that aerobic glycolysis enables ATP generation at faster rates than oxidative phosphorylation (Pfeiffer et al., 2001; Bose et al., 2021). Also, cancer cells are able to increase their glucose uptake through the upregulation of glucose transporter 1 (GLUT1) expression. Henceforth, the energetical inefficiency of glycolysis does not compromise cell growth and survival when nutrients and oxygen are abundant, since it is balanced by the ability to produce ATP rapidly (Vander Heiden et al., 2009). In turn, this reveals that, under physiological conditions, the Warburg effect does not interfere with energy production. Furthermore, other studies have shown that aerobic glycolysis

benefits proliferating cells as it enables biomass creation, which is essential to duplicate the cells' internal contents (Ngo et al., 2015), it plays a role in maintaining the redox balance (Gwangwa et al., 2018), and it promotes invasion and metastasis (Alfarouk et al., 2011; Persi et al., 2018).

In addition to the Warburg effect, cancer metabolism is a complex phenomenon and there are other mechanisms that are still being discovered. For instance, it has been shown that cancer cells can modulate the metabolism of fibroblasts, which undergo aerobic glycolysis and produce lactate (Pavlides et al., 2009). Subsequently, the surrounding cancer cells reuptake the metabolites produced by fibroblasts and further catabolize them in the aerobic respiration cycle, which allows them to produce high energy amounts. This process is commonly termed as the "reverse Warburg effect", and it can also occur between distinct tumour cell populations in MCTS, one of which presents a glycolytic phenotype, while the other catabolizes lactate. Therefore, understanding how glucose modulates the evolution of MCTS is also crucial. Moreover, cancer cells metabolize more glutamine than healthy cells to use as a carbon source for macromolecule biosynthesis (Krall and Christofk, 2015; Yoo et al., 2020). Consequently, "glutamine addiction" is also recognized as one of the reprogrammed metabolic pathways in tumour cells (Wise and Thompson, 2010).

MCTSs have been used extensively in research since the 1970s, when Sutherland et al. (1971) first described this experimental model. Specifically, scientists have relied on this model to assess the effect of nutrients on cancer progression and test the effect of new anti-cancer therapies (Hirschhaeuser et al., 2010; Han et al., 2021). At initial stages of growth, MCTSs have small diameters and, thus, the cells rely on diffusion to obtain nutrients from the microenvironment. However, once a critical diameter is reached, diffusion alone is not sufficient for nutrients to reach the spheroid core. In addition, metabolic waste that originates from cell metabolism starts accumulating in the spheroid centre (Hirschhaeuser et al., 2010; Kunz-Schughart et al., 1998). Consequently, large avascular MCTSs present a layered distribution of cells similar to that observed in in vivo solid tumours (Mueller-Klieser, 1987; Murphy et al., 2022) with three concentric regions. At the spheroid core, cells become necrotic since they do not receive enough nutrients to survive. Besides, it is possible to distinguish two cell populations in the spheroid rim: proliferating cells in the periphery, where there are more nutrients and oxygen, and quiescent cells in the internal region close to the necrotic core. The internal structure of a tumour spheroid and the internal distribution of nutrients and metabolic substances is shown in Fig 4.8.

Several previous experimental studies tried to characterize the internal distribution of chemical substances in tumour spheroids (Casciari et al., 1988, 1992b; Rousset et al., 2022). Additionally, multiple of these works relied on mathematical models to predict the distribution of substances such as glucose, oxygen and lactate, as well as their consumption and secretion rates. These integrated frameworks are particularly relevant in metabolic reprogramming research. It is recognized that cancer cells reprogram their metabolism to produce the energy and macromolecules required to keep their high proliferation rates. Specifically, tumour cells favour glycolytic pathways, which are fast yet energetically inefficient, to generate energy at faster rates than normal cells. Nonetheless, scientists have not yet been able to unravel the mechanisms that lead to this evolutionary adaptation.

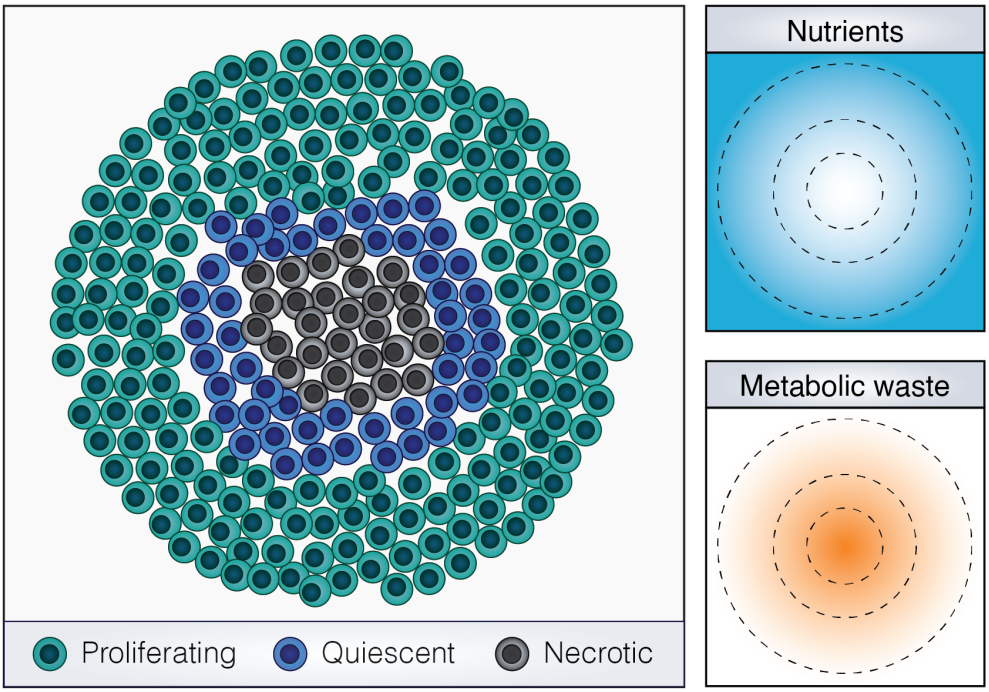

Figure 4.8: **Internal organization of a tumour spheroid.** When tumour spheroids reach a critical size, nutrient diffusion becomes limited and the cells that are in the tumour core start to respond to nutrient shortage by becoming quiescent (represented in blue) or dying (represented in black). Consequently, spheroids present a well-defined internal structure formed by three concentric areas: proliferating, quiescent and necrotic cells. The distribution of nutrients and metabolic waste in tumour spheroids is also characteristic in these structures. Specifically, they are characterized by nutrient shortage as well as an accumulation of metabolic waste at their core.

4.4.2 Modelling cellular systems and glucose metabolism

Intracellular models

Metabolism, like other biological behaviours, can be modelled at different scales. One of the simplest and most commonly employed approaches to incorporate glucose metabolism in hybrid models of tumour growth is to account for the main metabolic processes in the models' continuum component (Shamsi et al., 2018; Astanin and Preziosi, 2009). As previously stated, glucose metabolism relies on two main pathways: glycolysis and oxidative phosphorylation. Glycolysis is characterized by being less energetically efficient than oxidative phosphorylation, since it requires glucose consumption at faster rates and produces fewer ATP molecules. Robertson-Tessi et al. (2015) developed a hybrid model that takes into account oxygen, glucose, ATP and lactate and their effect on tumour cells. and the several previous studies have been designed where cell death was deterministically induced when a critically low concentration of glucose was reached to simulate necrosis. In addition, glucose consumption rates may increase in low-oxygen regions to simulate the metabolic switch between aerobic and anaerobic pathways, while also accounting for the Warburg effect (Astanin and Preziosi, 2009).

Mathematical descriptions of intracellular networks at the subcellular scale have been developed with distinct formalisms being applied to simulate signalling and metabolic pathways (Resendis-Antonio et al., 2015). Metabolic reaction network models aim to create mechanistic representations of the metabolites that take part in a given pathway and how they interact (Edelman et al., 2010). These representations can differ on their assumptions regarding not only whether a system is continuous or discrete in time and space, but also if it is deterministic or stochastic. Furthermore, network models can be generally classified into two main groups: stoichiometric and kinetic models (Yasemi and Jolicoeur, 2021). The former take into account the stoichiometry of the metabolic reactions and their time-independent characteristics, while the latter introduces additional information on metabolite kinetics (Gombert and Nielsen, 2000). Besides, new techniques have been developed to combine kinetic and stoichiometric modelling, coupling the comprehensiveness of constraint-based approaches with the detailed mechanisms of kinetic models (Jamshidi and Palsson, 2010; Júlvez and Oliver, 2020; Sahu et al., 2021).

On the one hand, stoichiometric approaches define metabolic networks as stochiometric matrices based on the number of metabolites, which are characterized as reactants and products, and reactions in a pathway (Edelman et al., 2010). Specifically, the stoichiometric matrices are composed of the metabolites' stoichiometric coefficients and the rows represent the metabolites while the columns define the reactions. Overall, these models define the mass balance over the metabolic network and they are particularly convenient because they can take into account genome-scale metabolic data, without requiring information on the kinetic parameters of the modelled pathways, which can be difficult to measure experimentally (Bordbar et al., 2014; Gatto et al., 2015). Using constraint-based approaches, for example flux balance analysis (FBA), it is possible to find the metabolic pathways that optimize cellular growth and energy production (Edelman et al., 2010). To achieve this, it is assumed that the system has reached a steady-state. Additionally,

some additional constraints, such as the bounds of the flux rates, may be imposed.

On the other hand, kinetic models aim to capture detailed and realistic representations of metabolic system dynamics. Henceforth, metabolite concentrations are modelled over time and they are usually represented by a set of ODEs that take into account specific metabolic reactions and detailed kinetic parameters (Stalidzans et al., 2018). The general expression for the evolution of a metabolite concentration (c) over time (t) is given by Eq 4.12:

$$\frac{dc}{dt} = \sum_{i=1}^{n} k_i r_i \tag{4.12}$$

where n is the number of reactions in which the metabolite takes part, k represents the stoichiometric coefficients and r is the rate of the reaction. The equation rates can be measured experimentally and be mathematically represented by different laws based on the complexity of the interactions between the agents in a given reaction. For example, mass action and Michaelis-Menten kinetics are some of the most used mathematical laws when modelling cell metabolism (Klipp and Liebermeister, 2006). Nonetheless, experimentally measuring the intracellular concentration of metabolites and the corresponding reaction rates can be an arduous task, making kinetic models more difficult to calibrate than stoichiometric approaches.

Recently, researchers have developed several biochemical network models to analyse the Warburg effect and other aspects of metabolic reprogramming (Tripathi et al., 2022; Hashemzadeh et al., 2020; Jia et al., 2019; Lagziel et al., 2019; Bertuzzi et al., 2010; Dai et al., 2016). Nonetheless, only a few models have explicitly taken into account the effect of glucose availability and its metabolism at a larger scale (Jiang et al., 2005; Piotrowska and Angus, 2009; Shan et al., 2018; Cleri, 2019; Schuster et al., 2015; Jagiella et al., 2016; Roy and Finley, 2019). Specifically, there is still a need for models that combine different aspects of cellular behaviour, such as motility, mechanics and cell-cell interactions, into a fully integrated and multiscale model. In the section to follow, we explore some research works that aim to bridge this gap and their implementation.

Multiscale models

One of the most relevant advances that has emerged in recent years due to the increase in computational power is the ability to couple models that describe different spatial and temporal scales (Anderson and Quaranta, 2008; Dada and Mendes, 2011; Walpole et al., 2013). Frameworks that integrate phenomena that occur over distinct scales are usually termed as multiscale models. The idea that cells require some kind of nutrient to survive and proliferate has been incorporated into several models (Gerlee and Anderson, 2007; Dormann and Deutsch, 2002; Lima et al., 2021; Bustamante et al., 2021; Swat et al., 2015) yet only a few have integrated explicit models of glucose metabolism (Rockne et al., 2019). Multiscale hybrid models are particularly relevant in this research field as they are able to integrate intracellular models of glucose metabolism, enabling scientists to understand how changes at the metabolic level will affect cell behaviour. A summary of hybrid models that consider glucose metabolism explicitly can be found in Table 4.3 and Fig 4.9 shows some illustrative results of on-lattice and off-lattice hybrid models.

Table 4.3: Summary of the reviewed multiscale models.

Design	Intracellular	Chemical species	Metabolism	Cell type	Validation	Reference
2D CPM	PDEs	Glucose, oxygen, waste, growth and inhibitory factors	Warburg	EMT6/Ro	Experimental	Jiang et al. (2005)
2D Lattice-gas CA	ODEs	Glucose, oxygen, H+ ions	Not considered	EMT6/Ro	Literature	Piotrowska and Angus (2009)
3D CA	ODEs	Glucose, oxygen, ATP, lactate, waste	Warburg	NSCLC	Experimental	Jagiella et al. (2016)
2D CBM	FBA	Glucose, oxygen, ATP, lactate, glutamine...	Warburg, reverse Warburg and glutamine addiction	Not specified	None	Shan et al. (2018)
2D CPM	ODEs	Glucose, oxygen, ATP, lactate...	Not considered	PaTu8988T	Literature	Roy and Finley (2019)
3D CA	PDEs	Glucose, oxygen, ATP	Not considered	EMT6/Ro	Literature	Cleri (2019)

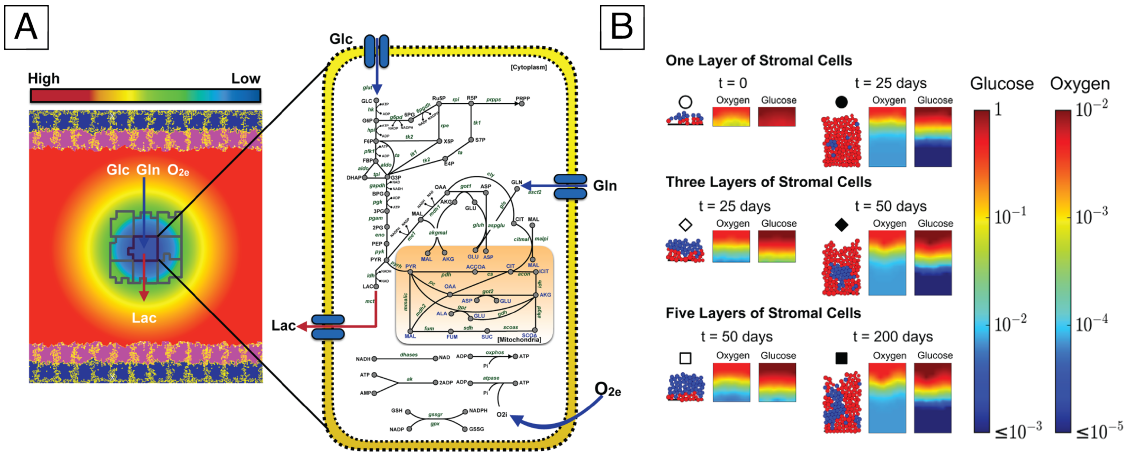

Figure 4.9: **Examples of multiscale hybrid model results.** (A) Illustration of a hybrid model where cells (outlined in grey) are represented by an on-lattice Cellular Potts model and the microenvironment consist of PDE-based descriptions of glucose, oxygen, lactate and glutamine. Single tumour cells consume glucose, glutamine and oxygen and produce lactate based on an intracellular metabolic network represented in the figure which considers both aerobic and anaerobic metabolism. (B) Model results for an off-lattice centre-based model that simulates tumour cells as spherical particles represented by their central point and radius. Reaction-diffusion equations are considered to study the spatial distribution of glucose and oxygen, represented as the colour gradients shown in the figure. Moreover, cell metabolism is introduced through a flux-balance analysis model that simulates the Warburg effect, the reverse Warburg effect and glutamine addiction. Subfigures A and B were taken from (Roy and Finley, 2019) and (Shan et al., 2018) respectively.

Several of these models were formulated with on-lattice implementations. Starting with models developed with CPMs, Jiang et al. (2005) proposed a model that was calibrated with experimental results obtained for EMT6/Ro tumour spheroids, a mammary carcinoma cell line. The authors explicitly modelled the Warburg effect using PDEs and defined that necrosis was induced when the glucose and oxygen concentrations went below a given threshold and lactate levels surpassed a maximum value. Furthermore, at the subcellular scale, the model considered a Boolean regulatory network of the cell cycle, which was modulated by growth and inhibitory factors. Using the CompuCell3D (Swat et al., 2012) modelling framework, Roy and Finley (2019) also implemented a CPM, but incorporated a much more complex intracellular model of cell metabolism, written as a set of ODEs that captured all the reactions in both glycolysis and the aerobic respiration cycle.

Cellular automata models have also been used in this field of research. For example, Cleri (2019) extended a previous agent-based model of cancer growth originally built to account for the effect of cytotoxic agents (Tomezak et al., 2016) to investigate the impact of metabolism on spheroid growth. In his work, the author introduced and implemented a simplified model of glucose metabolism and tested the effect of different nutrient sources, such as constant and sinusoidal glucose supply mechanisms. The model results were qualitatively compared with data from the literature and were found to be relevant. In addition, Piotrowska and Angus (2009) calibrated a lattice-gas cellular automaton model with experimental data available from the literature (Freyer

and Sutherland, 1985). Glucose metabolism was modelled through a set of ODEs as done previously by Venkatasubramanian et al. (2006). Furthermore, Jagiella et al. (2016) implemented an intracellular model of glucose metabolism based on ODEs considering the aerobic and anaerobic pathways effect and investigated which metabolic conditions increased the similarity between the computational and experimental datasets.

Regarding the off-lattice approach, one previous work has employed a CBM to study MCTS formation and the role of glucose metabolism on their growth dynamics. In particular, Shan et al. (2018) used the iDynoMiCS (Lardon et al., 2011) framework to model cell behaviour and integrated a complex intracellular network based on FBA and explicitly reproduced the Warburg effect, the reverse Warburg effect and glutamine addiction. In their work, the authors tested how growth dynamics changed according to the metabolic reprogramming strategy that was adopted by the cells.

4.4.3 Future perspectives

In the last decades, computational modelling of solid tumour growth has evolved to consider biological phenomena that occur over different time and spatial scales. Moreover, cell metabolism has been increasingly recognized as one of the main hallmarks of tumour growth and progression, and it has been accordingly integrated into computational frameworks to be better understood. Overall, these models have been able to simulate experimental data and reveal which type of metabolic response best fits the available results. Furthermore, with the constant increase in data availability, models are evolving to become patient-specific (Karolak et al., 2018; Yang et al., 2021). However, many of the models developed until now are focused on a single scale and there is still a need for fully coupled models that integrate biochemical networks with cellular and extracellular behaviour.

We highlight that, although we focus our review on the avascular stages of MCTS growth, models that take into account neovascularization have also been developed (Patel et al., 2001; Anderson et al., 2009; Smallbone et al., 2007; Antonopoulos et al., 2019; Shamsi et al., 2018). Some of these models also investigate invasion since tumour cells are able to metastasize by entering the circulatory system through vascularization (Robertson-Tessi et al., 2015; McEvoy et al., 2022; Escribano et al., 2019). Nonetheless, these implementations were developed with on-lattice frameworks, which fail to cell mechanics accurately. Moreover, cell-based multiscale models offer an advantage over population-based models since they are able to capture heterogeneity at the individual cell level. This is of particular interest when scaling from tumour spheroid to tumour organoid models. Tumour organoids are 3D self-organized structures grown from patient-derived cancer stem cells (LeSavage et al., 2022; Tuveson and Clevers, 2019). These models enable a higher level of personalization since tumours of the same type (e.g., lung, brain, pancreas) can differ between patients. Henceforth, when developing patient-specific models, it is crucial to be able to calibrate these models with this level of detail. Besides, the internal structure of organoids can be more complex than that of tumour spheroids, thus making it necessary to introduce stochasticity and heterogeneity at the cell level.

Lastly, multiscale models can be further expanded to consider phenomena at the tissue-level. Taking into account that the number of cells in a tissue makes it unfeasible to create a tissue of a model at the individual cell level, this can be achieved by coupling intracellular and cellular models with continuum representations that capture the tissue's mechanics (Fletcher and Osborne, 2022). This strategy is traditionally based on treating the tissue as a continuous material, thus eliminating the heterogeneity that can arise when a tumour is present. Nonetheless, more recent approaches have solved this issue by selecting a region of interest for which a cell-based model is used to evaluate the dynamics of tumour growth. Subsequently, this detailed description can be integrated into a continuum approach (Varella et al., 2022). Therefore, it is possible to reduce the computational power required to study the effect of tumour growth at the tissue-level, and it would certainly be invaluable to be able to use these approaches to recognize how metabolic reprogramming may influence the evolution of a tumour at the tissue-level.

4.5 Summary

In this chapter, we investigated the role of cell-cell mechanics, proliferation and metabolic reprogramming on tumour growth. In addition, we based our analysis on the previously discussed effect of cell motility and the influence of the extracellular microenvironment on cell behaviour. Accordingly, the work developed in this chapter represents the second building block of this dissertation. Building on the knowledge and applications established in Chapter 3, we were able to research tumour growth and characterize how it is modulated by both chemical and mechanical properties of the microenvironment. In particular, we extended the computational model implemented in Chapter 3 to account for proliferation and cell-cell mechanics, moving from individual to collective cell behaviour. On the whole, we concluded that these new mechanisms were relevant to reproduce tumour growth dynamics but, interestingly, we found that cell migration greatly influenced the formation of multicellular structures. Specifically, we observed that when cells were able to move freely due to the mechanical properties of the ECM, tumour growth was not observed, and vice-versa.

At this current stage of development, we did not implement the effect of chemical cues in our computational framework. Nonetheless, we reviewed models of avascular stages of tumour growth and how these works evolved to recognize the relevance of reproducing nutrient availability and metabolic reprogramming. Thus, we concluded that our framework can be extended conveniently to reproduce this intracellular mechanism by introducing glucose, oxygen, lactate, energy and other substances that users may find to be relevant in this process as diffusing and reacting species in the model. Furthermore, cell metabolism can be simulated through simplified consumption/secretion dynamics defined by PDEs or more complex intracellular ODE-based networks can be implemented as subcellular models. We highlight that these biomechanical models of tumour growth may be of particular interest to investigate the interplay between glucose consumption and the mechanical properties of the ECM (Ge et al., 2021; Pickup et al., 2014; Zanotelli et al., 2022; Salvi and DeMali, 2018). Recent studies have highlighted that changes in the mechanical

properties of the ECM modulate cancer metabolism, in particular glucose metabolism (Ge et al., 2021). For instance, it has been shown that cells that are detached from the ECM change their metabolism and decrease their glucose uptake and that migrating cells regulate their glucose uptake in response to the ECM's mechanical properties (Zanotelli et al., 2022; Sullivan et al., 2018). With this in mind, it is highly relevant to develop models of glucose metabolism that allow for the study of cell motility and cell mechanics at the individual level, and how cell-cell and cell-matrix interactions can also play a role in modulating these dynamics.

CHAPTER 5

Pattern formation

Contents

5.1 Introduction 88
5.2 Methodology 90
5.2.1 Study design 91
5.2.2 Neuron representation 93
5.2.3 Cell biological events 93
5.2.4 Cell mechanics and physical interactions 95
5.3 Results 99
5.3.1 Simplified model of Homer-Wright rosette formation 99
5.3.2 Numerical sensitivity analysis studies 100
5.3.3 Tissue-scale simulations 102
5.4 Summary 104

This chapter is based on:

Inês Godinho Gonçalves and Jose Manuel Garcia-Aznar. **"*neurorosettes:* A novel computational modelling framework to explore the formation of Homer-Wright rosettes in neuroblastoma."** Computational Particle Mechanics (under review)

5.1 Introduction

Biological tissues exhibit structured, well-defined architectures which arise from balanced chemical and mechanical interactions between cells and the surrounding microenvironment (Masserdotti, 2006; Hagios et al., 1998). Thus, deregulated tissue organisation may result from cell-level disturbances, such as genetic defects or altered phenotypes (Otte et al., 2021). For instance, tumours disrupt the architecture of normal tissues since they impact the mechanical properties of the microenvironment and the interactions between cells (Almagro et al., 2022; Nelson and Bissell, 2006; Noble et al., 2021). Indeed, pathologists use histologic records to diagnose health problems, such as cancer and heart diseases, because it is possible to distinguish between healthy and diseased tissues based on their microarchitecture (Acar et al., 2012). For instance, rosettes are histologic architectural patterns represented in Fig 5.1 found in a variety of tissues, most notably the nervous system (Wippold and Perry, 2006; Dell'Aquila et al., 2020).

Rosettes are defined as small groups of cells organised in circular configurations around a common core, as illustrated in Fig 5.1. There are various subtypes of rosettes, including HW, true ependymal, and neurocytic rosettes (Wippold and Perry, 2006; Matsumura et al., 2014). HW rosettes have a well-defined circular geometry, as shown in Figure 5.1a. Moreover, their central core is composed of neuronal processes, commonly designated as neuropil, which are meshwork of fibres derived from the developing neuronal processes of the sympathetic nervous system. True ependymal rosettes (see Figure 5.1b) have a well-defined circular geometry as well, but, unlike HW rosettes, they enclose an empty lumen. Lastly, neurocytic rosettes (see Figure 5.1c) have a central core filled with neuropil, but they are usually more elongated and present a more irregular contour than HW rosettes. In addition, reports have also revealed different architectures, such as rosettes that enclose a lumen with cytoplasmic processes (Wippold and Perry, 2006), pseudorosettes that surround a blood vessel (Chow and Brittingham, 1987) and rosettes with basophilic cells inside the central core (Das et al., 2014). Nevertheless, in this research work, we primarily focused on the factors that lead to the emergence of HW rosettes.

HW rosettes are a rosette subtype that arises in neuroblastoma, the most lethal and common extracranial tumour in infants (Maris, 2010; Wright, 1910; Shimada et al., 1999b). The presence of HW rosettes in histological records is usually taken into account during the diagnosis of this type of tumour and it is characterized by the existence of circular arrangements of cells that surround a common core containing fibres. Although previous studies have suggested that HW rosettes originate from poorly-differentiated neurons and result from cell-cell interactions (Wippold and Perry, 2006; Moss, 1983), these observations have not been further confirmed and validated. Taking into account that these structures originate during neuronal development and are triggered by specific differentiation stimuli, it is not a trivial task to investigate and characterise their formation experimentally (Duarte Campos et al., 2019).

Computational modelling has been extensively employed to simulate biological tissues and tumour environments in order to answer questions that would otherwise be challenging to evaluate (Bull and Byrne, 2022; Enderling and Rejniak, 2013; Zhang et al., 2007; Krakauer et al., 2011; Li

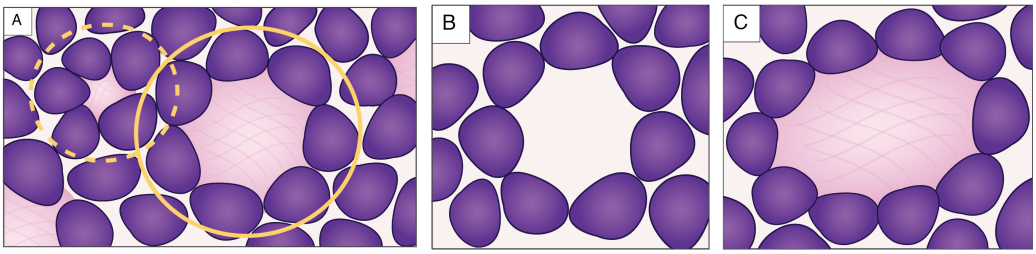

Figure 5.1: **Simplified representations of different types of rosettes.** (a) shows a tissue with HW rosettes, which are circular arrangements of cells that surround a common core containing a meshwork of fibres similar to neuropil. HW rosettes are identified by their well-defined circular structure (marked with a full circle). Yet, spatial architectures with an almost-radial arrangement (marked with a dashed circle) may also be seen in histological records due to tissue heterogeneity. (b) shows one true ependymal rosette characterised by an empty lumen at the centre of the radial structure formed by the surrounding cells. (c) shows one neurocytic rosette, which contains a meshwork of fibres like an HW rosette but presents a more irregular and elongated structure.

and Lowengrub, 2014; Logan et al., 2010). Previous studies have employed cell-based methods to show how cells organise into circular arrangements resembling rosettes during morphogenesis of the nervous system in zebrafish to form neuromasts, which are sensory organs (Colombi et al., 2020; Di Costanzo et al., 2015; Dalle Nogare and Chitnis, 2017). In addition, new modelling frameworks have recently arisen to explore neuronal development and the formation of neuronal structures with both cell bodies and neurites through the inclusion of simulation objects with more complex geometries (Zubler, 2009; Breitwieser et al., 2022; Miller and Suter, 2018; Koene et al., 2009; Aćimović et al., 2011). However, to the best of our knowledge, there are no computational models tailored to account for the interactions between the different components of neurons and mimic the development of HW rosettes in neuroblastoma.

In this chapter, we intend to achieve a computational framework that can be used to reproduce the formation of HW rosettes. Therefore, as explained in Section 5.2, we implemented a cell-based model with an improved representation of cell morphology. Specifically, we used spherical and cylindrical objects to simulate an approximated geometry of neurons, which are composed of cell bodies and neuronal processes. In addition, we evaluated the role of mechanical interactions between model components and biological events (i.e., differentiation and proliferation) in the formation of HW rosettes. Section 5.3 presents three different configurations that were tested to identify the model parameters that promoted rosette assembly, starting with an idealized scenario and progressing towards a more realistic approximation. In the scope of this dissertation, this chapter represents the final stage of model complexity, given that we are considering a computational model with all the model components implemented in Chapters 3 and 4, namely cell motility, cell-matrix and cell-cell interactions and proliferation, and we are introducing a final element, neuronal differentiation.

Table 5.1: **Reference parameter values for the model.** Values in brackets represent ranges, showcasing that the parameter values were not fixed and varied based on the configuration of the simulation to reflect the presence and spatial location of chemotactic gradients.

Symbol	Parameter	Value	Unit	References
T_{sim}	Total simulation time	[5-24]	h	Estimated
Δt	Simulation time step	0.1	min	Estimated
μ	Drag Coefficient	10.0	Pa · s	Estimated
r_{div}	Proliferation rate	0.0007	1/min	Estimated
r_{diff}	Differentiation rate	[0.0007-0.07]	1/min	Zubler (2009), estimated
$n_{neighbours}$	Number of neighbours	19	-	Estimated
$n_{neurites}$	Number of neurites	[2-7]	-	Estimated
θ_{turn}	Neurite elongation turning angle	$\pi/30$	-	Mortimer et al. (2009)
R_c	Cell radius	7.0	μm	Ghaffarizadeh et al. (2018)
$R_{c,int}$	Cell interaction radius	$1.25R_c$	μm	Ghaffarizadeh et al. (2018)
R_n	Neurite radius	0.5	μm	Zubler (2009)
$R_{n,int}$	Neurite interaction radius	1.5	μm	Zubler (2009)
k_{spring}	Neurite spring constant	10.0	-	Zubler (2009)
L	Neurite default length	10.0	μm	Zubler (2009)
L_{min}	Neurite minimum length	3.0	μm	Zubler (2009), estimated
c_{cca}	Cell-cell adhesion coefficient	[0.0-15.0]	-	Ghaffarizadeh et al. (2018), estimated
c_{ccr}	Cell-cell repulsion coefficient	50.0	-	Ghaffarizadeh et al. (2018), estimated
α_{cc}	Cell-cell smoothness factor	1.0	-	Ghaffarizadeh et al. (2018), estimated
c_{cna}	Cell-neurite adhesion coefficient	5.0	-	Ghaffarizadeh et al. (2018), estimated
c_{cnr}	Cell-neurite repulsion coefficient	100.0	-	Ghaffarizadeh et al. (2018), estimated
α_{cn}	Cell-neurite smoothness factor	1.0	-	Ghaffarizadeh et al. (2018), estimated
c_{nna}	Neurite-neurite adhesion coefficient	5.0	-	Ghaffarizadeh et al. (2018), estimated
c_{nnr}	Neurite-neurite repulsion coefficient	25.0	-	Ghaffarizadeh et al. (2018), estimated
α_{nn}	Neurite-neurite smoothness factor	1.0	-	Ghaffarizadeh et al. (2018), estimated

5.2 Methodology

We designed an open-source modelling framework to simulate the formation of HW rosettes in neuroblastoma. We implemented different types of physical interactions and biological events, as explained in more detail below. Furthermore, we wrote our model in the Python 3 programming language due to its simple syntax and popularity in the biological sciences community and neuroscience, in particular (Muller et al., 2015). We created this framework to be intuitive, require minimal code and provide high-quality renderings of the results in real-time, making it attractive and easily useable by the general scientist. Henceforth, we simulated tissues in 2D to ease the comparison between our results and the histological records found in the literature. Furthermore, users can define model parameters in configuration files uncoupled from the main code, making the framework accessible to a general audience. The numerical values used in this work are presented in Table 1. Our package can be installed easily through Python's package installer, pip[1], and the source code files are publicly available (Gonçalves and García-Aznar, 2022).

One of our goals in the model design process was to create a framework that enabled real-time rendering of the simulation results to help users assess them visually. Therefore, we used the *vedo* package (version 2022.1.0) as our rendering engine since it creates high-quality interactive plots

[1] https://pypi.org/project/neurorosettes/

and animations and it has a shallow learning curve (Musy et al., 2022). Although we developed our framework to work primarily with this package, it is not strictly coupled to the code. Thus, it can be replaced by other rendering tools, such as *PyVista* (Sullivan and Kaszynski, 2019), according to the users' needs and preferences. In addition, screenshots of the model results can be taken, and simulation objects can be saved as meshes, to be further processed with visualisation tools such as *ParaView* (Ahrens et al., 2005).

5.2.1 Study design

Neuroblastoma develops when neuronal differentiation is blocked, which can happen at different stages of development. Due to the highly heterogeneous nature of neuroblastoma, different tissue phenotypes have been observed based on the differentiation grade of neuroblastic tumours (Westermark et al., 2011). On one end of the spectrum, undifferentiated tissues present multiple round cell bodies that remained undifferentiated. On the other end of this spectrum, differentiated neuroblastoma tissues share similarities with healthy tissues and present both cell bodies and well-differentiated neuronal processes that have a fibre-like appearance (Shimada et al., 1999b). Between these two distinct phenotypes, it is observed that poorly-diifferentiated neuroblasts that present round cell bodies with short neuronal processes organize into circular spatial arrangements. In particular, it can be observed that the cell bodies surround a common core filled with neuropil, a fibre-like material from which neuronal processes are made.

Figure 5.2a illustrates the neuronal differentiation process in healthy tissues. Conversely, as shown in Figure 5.2b, this differentiation process is blocked in neuroblastoma due to genetic factors and environmental cues. Undifferentiated tissues are composed mainly of cell bodies, whereas differentiating phenotypes have more regions filled with neuronal processes. Nonetheless, in their majority, neuroblastic cells have a cell body and a few short neurites (Kholodenko et al., 2018). The current hypothesis for explaining the emergence of HW rosettes assumes that these histological patterns are a result of the interactions between the cell bodies and neurites of neuroblastoma cells, as illustrated in Figure 5.2b (Wippold and Perry, 2006; Moss, 1983). Studies have revealed that cells undergoing neuronal differentiation secrete cell surface receptors that facilitate cell-cell recognition and adhesion (Moscona, 1976). It is therefore assumed that as cell bodies become adhered, their neurites interact and entangle, creating the radial structure of the rosette (Katsetos et al., 1988).

With this in mind, we developed a computational approach that could replicate the process of neuronal differentiation at distinct grades. In other words, we established that neuronal differentiation could occur yet it could be blocked at different development stages to simulate distinct neuroblast phenotypes. Furthermore, we designed a simplified simulation setup to reproduce the assembly of HW rosettes as represented in Fig 5.2. In particular, we investigated how this process occurred in an ideal scenario where undifferentiated cells were arranged in a circular pattern and, subsequently, compared the results obtained in this configuration to a more realistic cell distribution.

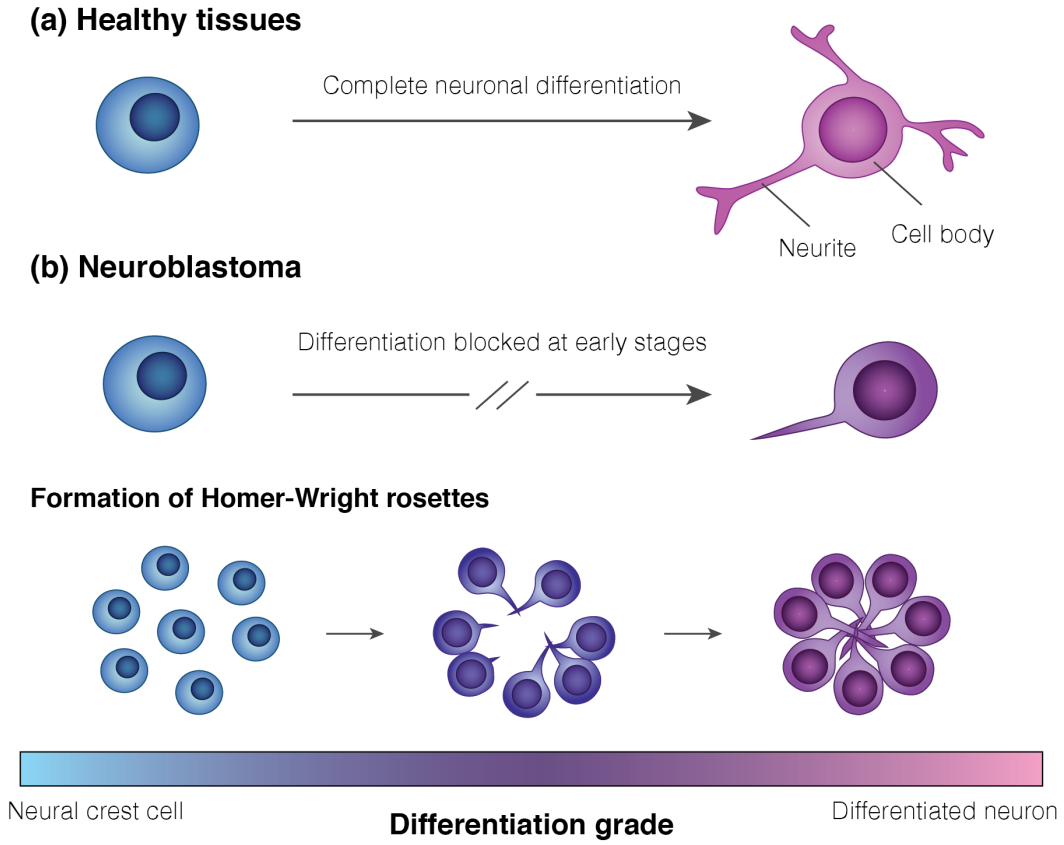

Figure 5.2: **Simplified representations of the neuronal differentiation process in healthy tissues and neuroblastoma.** On the one hand, in healthy tissues, neural crest cells differentiate into cells of the nervous system and create long neurites, adopting the structure of a neuron. On the other hand, in neuroblastoma, neuronal differentiation is blocked at the early stages. Consequently, the neuronal processes of neuroblastic cells are shorter and less differentiated than those seen in healthy tissues. Furthermore, in neuroblastoma, the morphology of poorly-differentiated neurons leads to the formation of a specific rosette subtype (Homer-Wright rosette). Specifically, the poorly-differentiated neuroblastoma cells take on a radial spatial arrangement as neurites become tangled and cell bodies adhere to each other. As a result, their neurites fill the central core of the rosette structure.

5.2.2 Neuron representation

We modelled neurons as it was previously simulated by Zubler (2009) by representing each neuron as a chain of springs and masses in series. As shown in Figure 5.3a, two types of physical objects were considered to reproduce neuronal system components and compute agent interactions: cell bodies were modelled as spheres, and neurites as cylinders. On the one hand, each cell body was represented by its central point mass and radius (R_c). On the other hand, neurites were defined by a spring and a point mass at one of their extremities, as well as their radius (R_n). When mechanical interactions were computed, the forces acting on the agent mass were calculated. Subsequently, the mass point was moved. For neurites, only one of the extremities was moved, while the other was considered to be the mass point of its adjacent object (cell body or neurite). Consequently, neurite rotation was neglected in this model. Furthermore, we assumed that each new neuron in the simulation was composed of a cell body in an undifferentiated state, whereas differentiated neurons were composed of a cell body and one or more neurites (Fletcher et al., 2000). Since we aimed to model the formation of HW rosettes and these structures are observed in poorly-differentiated tissues (Wippold and Perry, 2006), we did not implement the formation of multiple primary neurites in distinct directions. Instead, we considered the creation of a single primary neurite that could be extended by adding secondary neurites along the same outgrowth axis.

5.2.3 Cell biological events

We implemented simplified models of cell proliferation and differentiation into our computational framework. As illustrated in Figure 5.3b, we defined that only undifferentiated neurons could undergo cell division since studies have shown that proliferation stops once neurons become committed to neuronal differentiation (Hardwick and Philpott, 2014). Moreover, we assumed new neurons inherited the same attributes as the neurons from which they were created, including their physical properties and biological event rates. Hereafter, as developed in previous works (Ghaffarizadeh et al., 2018), we established that cells presented a probability to divide (P_{div}) at a given time interval Δt that was based on a user-defined proliferation rate (r_{div}) and given by Equation 5.1.

$$P_{\text{div}} = r_{\text{div}} \Delta t \tag{5.1}$$

Additionally, we introduced some control mechanisms to replicate the neurons' response to high cell densities and simulate the effect of contact inhibition on cell cycling (Pavel et al., 2018). Consequently, we defined the maximum number of neighbours cell bodies could have inside a given interaction region. If the number of cell bodies inside this area surpassed the threshold established ($n_{\text{neighbours}}$), proliferation was inhibited, as shown in Figure 5.3b. The proliferation parameter values are defined in Table 5.1.

Neuronal differentiation, represented in Figure 5.3c, was also modelled based on the probability of a neuron to differentiate at a given time interval (r_{diff}) as given by Equation 5.2.

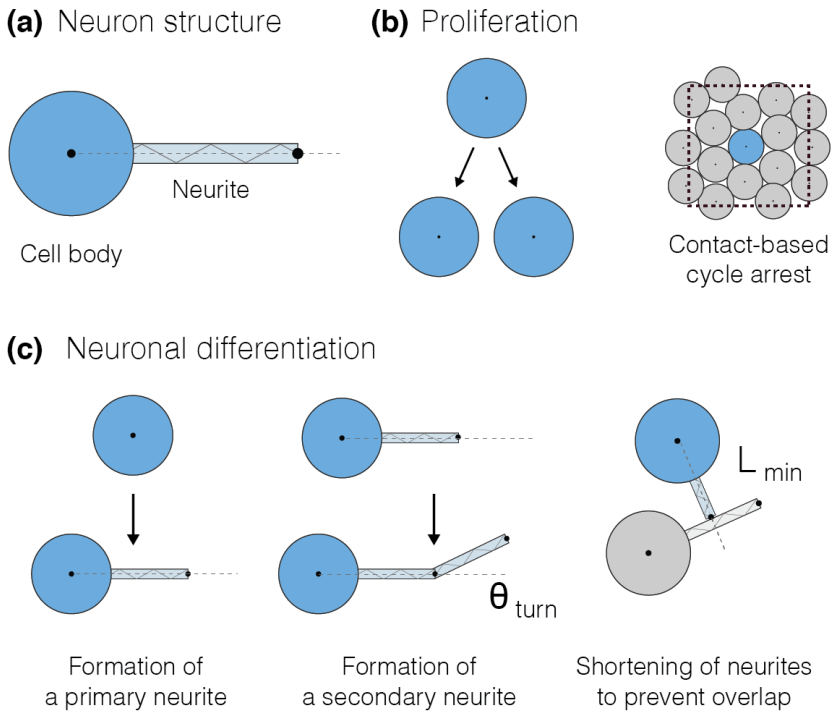

Figure 5.3: **Neuron implementation and rules.** (a) Neurons were modelled as a combination of a single cell body, represented by a sphere, and neurites, depicted as cylinders with a spring axis. In addition, neurons were able to undergo (b) proliferation or (c) differentiation based on user-defined rates. (b) During proliferation, an undifferentiated neuron originated two new cell bodies. This mechanism was arrested when the number of neighbours in a region surpassed a given threshold to replicate contact inhibition. (c) During differentiation, new primary or secondary neurites were created. The latter were able to turn slightly to mimic the turning mechanisms observed during neurite outgrowth. The degree of freedom for this process was defined by the angle θ_{turn}. Furthermore, we prevented that new neurites were placed on top of existing neurites by decreasing their length until no collisions were detected. Nonetheless, if the length of the new neurite was smaller than a defined threshold L_{min} , it would not be created

$$P_{\text{diff}} = r_{\text{diff}}\Delta t \tag{5.2}$$

Here, P_{diff} represents the probability of a neuron cell body to differentiate and create a new neurite in a time step Δt. We defined that the direction of a primary neurite, i.e., the first neurite to be created, would follow a direction specified by the user. If no preferential direction was defined, a random growth axis was chosen. Furthermore, secondary neurites tended to follow the direction of their mother's axis. Previous neurite outgrowth models considered that neurites adjusted their direction of growth by turning slightly at small angles towards the left or right in regards to its current growth axis (Mortimer et al., 2009, 2010; Catig et al., 2015). Consequently, we incorporated this turning mechanism into our model by introducing an angle of freedom (θ_{turn}), which dictated how much secondary neurites could stray away from their mother's growth axis, as illustrated in Figure 5.3c.

Furthermore, Figure 5.3c also shows how we avoided the overlap between neurites of distinct neurons, as these structures cannot physically occupy the same area in space. Specifically, when a new neurite was created, we evaluated whether it would come into contact with an already existing neurite. If this was the case, we decreased the length of the neurite to avoid collisions between the two agents. Moreover, if the minimum length required to avoid contact between the objects was smaller than a given threshold ($L_{n,min}$), the neuron would not differentiate. Lastly, we modelled the differentiation blocking process by defining a maximum number of neurites that a neuron could have ($n_{neurites}$). When a neuron differentiated up until the point where this threshold was reached, neuronal differentiation was blocked. The differentiation parameter values can be found in Table 5.1.

5.2.4 Cell mechanics and physical interactions

We introduced cell mechanics functions in our model to reproduce the physical interactions between cell bodies and neurites as proposed by Zubler (2009). Henceforth, we took into account the internal forces created in the springs that connected neurites and cell bodies, as well as the forces resulting from interactions between neighbouring objects, as represented in Figure 5.4. Interactions were computed by taking into account the forces acting on the objects' mass point (i.e., the central point of cell bodies and the mass point of neurites). An object was able to interact with a neighbour object if it was inside a radius of interaction ($R_{c,int}$ for cell bodies and $R_{n,int}$ for neurites).

Forces acting on cell bodies

Since cell bodies were represented by a single mass with no attached spring, the forces acting directly on them were, primarily, interaction forces (see Figure 5.4a). Here, we considered that object interactions were represented by sphere-sphere contacts (Zubler, 2009). Henceforth, although cell-cell interactions could be easily calculated by taking into account the central position of the cell body, the geometry of the neurites required some adjustments. Accordingly, as illustrated in

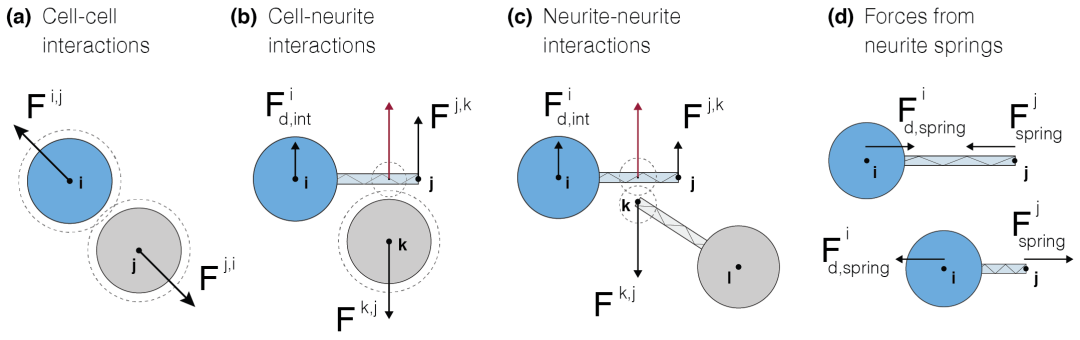

Figure 5.4: **Representation of the mechanics module which considered agent interactions based on adhesion/repulsion dynamics and neurite spring forces.** (a) Interaction forces ($F^{i,j}$) between two cell bodies, i and j, inside a radius of interaction (represented as dashed spheres) were computed through sphere-sphere contacts. (b-c) Interactions between a neurite j and a cell body/neurite k were also computed as sphere-sphere interactions. Consequently, we identified the closest point on the neurite's axis to its neighbour and, subsequently, calculated the resulting force assuming that the neurite behaved as a sphere centred on that point. The resulting force was then distributed by the neurite extremities according to their distance to the interaction point. Accordingly, a fraction of the force ($F^{j,k}$) was applied on the neurite's mass, while the remaining fraction ($F^{i}_{d,int}$) was transmitted to its mother, i. (d) As neurites were represented as Hookean springs, forces originated when they were stretched or compressed. Therefore, when the length of a neurite j changed, a force was applied on its mass point (F^{j}_{spring}) and an opposite force ($F^{i}_{d,spring}$) was transmitted to its mother, i.

Figure 5.4b, cell-neurite interactions were computed by identifying the point on the neurite axis that was closest to the cell body and computing the interaction forces between a virtual sphere, centred at this point and with the same radius as the neurite, and the cell body.

We defined that the interaction force between neighbouring objects was composed of adhesion and repulsion components. Furthermore, we modelled physical interactions with interaction potentials as implemented previously by Ghaffarizadeh et al. (2018). Hence, interactions between a cell i and its neighbour j were regulated by adhesion and repulsion coefficients and were given by Equations 5.3 and 5.4:

$$\nabla\phi(\mathbf{r}) = \begin{cases} \left(1 - \frac{|\mathbf{r}|}{R_A}\right)^2 \frac{\mathbf{r}}{|\mathbf{r}|} & \text{if } |\mathbf{r}| \leq R_A \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{5.3}$$

$$\nabla\psi(\mathbf{r}) = \begin{cases} -\left(1 - \frac{|\mathbf{r}|}{R}\right)^2 \frac{\mathbf{r}}{|\mathbf{r}|} & \text{if } |\mathbf{r}| \leq R \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{5.4}$$

where r is the distance vector between the masses of two neighbouring objects. $R_{],int}$ represents the cell body radius of interaction, which is the maximum distance at which adhesion forces are present, and R_c is the cell body radius. Furthermore, c_{adh} and c_{rep} represent the adhesion and repulsion coefficients, respectively, and α is a smoothness factor of the potentials function. In this model, we consider that cell-cell and cell-neurite interactions were defined by different coefficient values. Thus, on the one hand, in cell-cell interactions c_{adh} , c_{rep} and α represent the cell-cell adhesion coefficient c_{cca} , the cell-cell repulsion coefficient c_{ccr} and the cell-cell smoothness factor α_{cc} , respectively. On the other hand, in cell-neurite interactions these coefficients correspond to c_{cna} , c_{cnr} and α_{cn}. The values for these coefficients can be found in Table 1.

In addition, cell bodies took into account the forces transmitted by their daughter neurites, if they had any. Due to the cylindrical geometry of neurites, interaction forces were distributed between their extremities, as explained in further detail in the subsection to follow. Consequently, progenitor cell bodies received a fraction of the interaction forces acting on their daughters ($F_{d,int}$), as illustrated in Figure 5.4b. Moreover, since this model considered that cell bodies did not have springs, the internal forces for those that had daughter neurites was not computed directly. Instead, the internal forces of daughters ($F_{d,spring}$) were transmitted to the progenitors when forces were computed for the corresponding neurite agents, as illustrated in Figure 5.4d. Taking this into account, the force that was exerted on a cell body i by its daughter could be given by Equation 5.5.

$$F^i_{daughter} = F^i_{d,spring} + F^i_{d,int} \tag{5.5}$$

Consequently, the sum of the forces acting on a cell body was computed and its centre of mass was moved. To achieve this, we took the inertialess assumption and defined that the equation of motion used to update the velocity of a cell i based on all of the forces acting on it was defined by Equation 5.6:

$$\mu \mathbf{v} \approx \sum_{j \in \mathcal{N}(i)} (\mathbf{F}^{ij}_{cca} + \mathbf{F}^{ij}_{ccr}) + \mathbf{F}^{i}_{drag} + \mathbf{F}^{i}_{loc} \tag{5.6}$$

where μ represents the drag coefficient of the tissue where cells are located.

Forces acing on neurites

Neurites, just like cell bodies, were able to interact with neighbouring objects. As previously explained, since neurites were modelled as their mass point, the resulting force from the interaction between a neurite and a cell body was computed by finding the point on the neurite axis closest to the cell body and assuming that the neurite could be represented by a sphere centred at this point. A similar approach was employed to compute neurite-neurite interactions, as represented in Figure 5.4c. Therefore, as explained in the previous section, the adhesion and repulsion forces between a neurite i and its neighbour j were given by Equations 5.7 and 5.8.

$$\nabla \phi(\mathbf{r}) = \begin{cases} \left(1 - \frac{|\mathbf{r}|}{R_A}\right)^2 \frac{\mathbf{r}}{|\mathbf{r}|} & \text{if } |\mathbf{r}| \leq R_A \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{5.7}$$

$$\nabla \psi(\mathbf{r}) = \begin{cases} -\left(1 - \frac{|\mathbf{r}|}{R}\right)^2 \frac{\mathbf{r}}{|\mathbf{r}|} & \text{if } |\mathbf{r}| \leq R \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{5.8}$$

Here, $R_{n,int}$ represents the neurites' radius of interaction, R_n represents their radius and r represents the distance vector between the two neighbours. Furthermore, c_{adh} and c_{rep} are the adhesion and repulsion coefficients and α is the potentials function smoothing factor for neurite interactions. Thus, these coefficients represent c_{cna} , c_{cnr} and α_{cn} in cell-neurite interactions and c_{nna} , c_{nnr} and α_{nn} neurite-neurite interactions. The values for these coefficients can be found in Table 1. Subsequently, the interaction forces were distributed by the two extremities neurite based on their distance to the closest point to the interaction point.

In addition, given that neurons were simulated as a system of spring in series, the tension force inside a neurite i (F^{i}_{spring}) was given by Hooke's law (Zubler, 2009; Dennerll et al., 1989) as written in Equation 5.9:

$$F^{i}_{spring} = k_{spring} \frac{(L^i - L_0)}{L_0} e^i \tag{5.9}$$

where, k_{spring} is the neurite's linear spring constant (see Table 1), L^i is the current length of the neurite, and L_0 is its resting length, which is the length at which the neurite is initialized, and e^i represents the unit vector that defines the neurite's growth axis. As illustrated in Figure 4d, when computing the forces acting on this system, the spring forces were applied directly to the neurite's mass. In addition, an opposite force $F_{d,spring}$ was transmitted to their progenitors. As done for the cell bodies, the forces transmitted by daughter neurites could be defined by the sum of $F_{d,spring}$ and $F_{d,int}$, as given by Equation 5.5.

The equation of motion for a given neurite i was defined by taking into account the inertialess assumption and considering all the forces acting on the neurite's mass as defined in Equation 5.10.

$$\mu \mathbf{v} \approx \sum_{j \in \mathcal{N}(i)} (\mathbf{F}^{ij}_{cca} + \mathbf{F}^{ij}_{ccr}) + \mathbf{F}^{i}_{drag} + \mathbf{F}^{i}_{loc} \tag{5.10}$$

5.3 Results

5.3.1 Simplified model of Homer-Wright rosette formation

In this work, we designed three computational studies to identify the factors that modulate the formation of HW rosettes. Firstly, we implemented a simulation to reproduce the current theoretical model that describes the formation of HW rosettes as previously explained in Section 1. Fundamentally, some previous studies hypothesized that the neurites of poorly-differentiated neurons interact and bind to each other, causing cell bodies to arrange in a circular geometry that encloses a meshwork of neuropil fibres (Wippold and Perry, 2006). Accordingly, to replicate this scenario, we created a simulation where cell bodies were placed in a radial arrangement, as shown in Figure 5 (see the subfigure for t = 0 min). We established that neuronal differentiation could occur up to a maximum of 2 neurites per neuron ($n_{neurites}$ = 2) and at a rate r_{diff} of 0.007 1/min. We defined the differentiation rate based on previous works that showed that neurites grow at a rate of some microns per minute (Miller and Suter, 2018) and we established a maximum of 2 neurites per neuron to simulate how differentiation is blocked at early stages in neuroblastoma (Otte et al., 2021; Mora and Gerald, 2004). At the cell mechanics level, we calibrated the adhesion and repulsion parameters to promote cell-cell and neurite-neurite adhesion (Wippold and Perry, 2006; Katsetos et al., 1988; Goodman et al., 1984). These values can be found in Table 1, except for the cell-cell adhesion coefficient, which was varied in different studies. Here, it was considered that c_{cca} = 5.

Furthermore, in these simulations, we established that the growth axis of new neurites pointed towards the centre of the circular arrangement of neurons. We defined these conditions to model an ideal scenario of HW rosette formation in which neurites would be more likely to come into contact with each other and form the characteristic structure of this rosette subtype. Biologically, neurites have been shown to extend towards sources of chemical substances, such as nerve growth factors and neurotrophins (Catig et al., 2015; Liu et al., 2005). For instance, Schwann cells, which are present in some neuroblastoma subtypes, are known to increase neurite outgrowth rates (Koppes et al., 2011). Therefore, this simulation aimed to simulate neurite outgrowth towards a source of differentiation-inducing signals at the centre of the rosette.

The results obtained in this initial study are presented in Figure 5.5, from which it is possible to conclude that the simulated neurons adopted the structure and morphology of an HW rosette. We highlight that this process was promoted by the adhesion between the neurons' cell bodies and the subsequent formation and growth of neurites that became adhered as they grew into the centre of the rosette. Since our model considered that neurites could turn and change their direction of growth slightly as they grew, it was also possible to discern this behaviour in the simulation

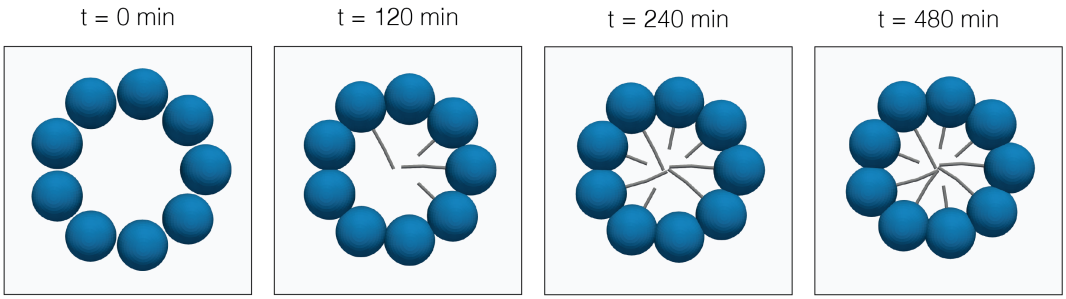

Figure 5.5: **Computational model results obtained for an HW rosette formed under ideal conditions.** Neuron cell bodies were placed in a circular arrangement, and neuronal differentiation was enabled. Yet, each neuron could only grow two neurites. We adjusted cell mechanics parameters to promote adhesion cell-cell and neurite-neurite interactions. As a result, we achieved a stable rosette pattern, as cells became adhered and neurites grew into the centre of the circular arrangement and started interacting with other neurites. The model was run for a total of 8 simulated hours.

outputs. Further turning and changes in the structure of the neurites were induced by the interactions with the neurites of other neurons. This computational simulation was also used to evaluate how the mechanical interactions between neuron cell bodies and the rates of neuronal differentiation modulated the morphology of HW rosettes. These results are presented and discussed in the section to follow.

5.3.2 Numerical sensitivity analysis studies

Having defined our base model of HW rosette formation, we designed a numerical analysis to evaluate how the model's response changed in response to different parameter values. We selected two main parameters for this study: the differentiation rate, defined by r_{diff} (see Equation 2), and the adhesion levels between cells, given by c_{cca} (see Equation 3). At the biological level, we decided to study the effect of the differentiation rate and not the proliferation rate to avoid disrupting the circular geometry of the rosette. In addition, while differentiation may occur at a time scale of minutes (Miller and Suter, 2018), proliferation should only occur after several hours or days. At the mechanics level, we focused on the interactions between cell bodies because they appeared to dominate how pattern formation occurred Since neurites were programmed to grow inwards, we observed that cell-neurite interactions did not play a major role in this system. Likewise, due to the reduced size of the rosette, it was highly probable for neurites to be in contact with each other. Hence, their interactions were regulated mainly by the adhesion forces induced by the cell bodies. Due to the lack of experimental data on HW rosettes and their origin, we could not fit these parameter values according to neuroblastoma cases. Consequently, in this work, we evaluated our results based on the morphology of the simulated HW rosettes, as there were no quantifiable data values that we could use to calibrate our model.

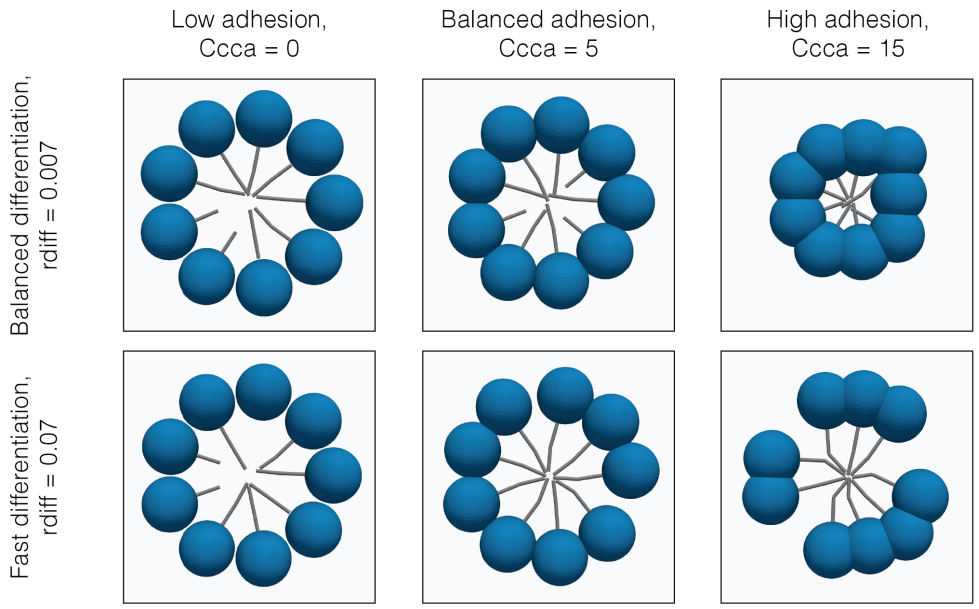

Figure 5.6: **Computational results of HW rosette formation at different differentiation rates and cell-cell adhesion levels.** At balanced differentiation rates, the adhesion between cell bodies appeared to regulate the radius and the stability of the rosette structure. Furthermore, at high differentiation rates, the radius of the rosette appeared to be maintained regardless of the adhesion levels between cell bodies. Yet, the mechanical interactions modulated the morphology of the rosette since neighbouring neurons became attached and disrupted the circular architecture of the rosette. Consequently, the rosette presented gaps between groups of cell bodies. The model was run for a total of 5 simulated hours

The results obtained for these simulations are shown in Figure 5.6. Overall, in these computational assays, we recognized that both the adhesion levels between cells and the differentiation rates of neurons caused an impact on the morphology of HW rosettes. Starting with the results obtained at balanced levels of differentiation (r_{diff} = 0.007 1/min), the model outputs suggest that balanced adhesion levels (c_{cca} = 5) promoted the formation of well-defined circular geometries with connected cell bodies. Neurites from opposite neurons also became adhered since the cell-cell adhesion levels induced a more compact radial geometry. We did not identify this behaviour when cell-cell adhesion was not present in the model (c_{cca} = 0). When only repulsion was considered, neurons were organized in a radial geometry, but the cell bodies were not connected. Also, there were fewer contact points between neurites, as the neurons were more distanced. Contrarily, at high adhesion levels (c_{cca} = 15), the circular geometry of the rosettes became smaller and disrupted. Snapshots of replicates for each condition are presented in Appendix B, Section B.3.1 to show that these results were consistent over different simulations.

In simulations performed at high differentiation rates (r_{diff} = 0.07 1/min), the simulation outputs appeared to indicate that the creation and growth of neurites at a faster rate disrupted the circular arrangement of the neurons, specifically at balanced and high cell-cell adhesion levels.

Based on the model results, neurites started to grow into the centre of the rosette before cell bodies were adhered. Hence, the interactions between neurites of neighbouring neurons induced the formation of disjointed circular architectures. This effect was promoted further at high adhesion levels, for which the results show that cell bodies were grouped into smaller clusters. We highlight that the radius of the rosette was also not decreased at this differentiation rate, as the mechanical interactions between the neurites that were already created balanced the adhesion forces between cell bodies. Conversely, at low adhesion levels, this effect was not observed, as there were no relevant cell-cell interactions in these simulations.

5.3.3 Tissue-scale simulations

Finally, in our last example application, we performed a study to confirm the relevance of the creation and extension of neuronal processes at larger scales. Hence, we designed a simulation to reproduce a tissue section comparable to those captured in histological images. We created an initial configuration where neurons with no neurites were placed in a hexagonal grid. In addition, cell clocks were initialized to model the three types of tissue in neuroblastoma: (i) undifferentiated, (ii) poorly-differentiated and (iii) differentiated (Shimada et al., 1999b). In case (i), we programmed cells to exhibit high division rates to capture the enhanced proliferative behaviour of undifferentiated tumour cells. Neuronal differentiation was completely blocked. In case (ii), we established that cells should have the same proliferation rate as in (i), but we introduced neuronal differentiation. In this case, we defined a maximum number of neurites per neuron to simulate the blocking of the differentiation process in poorly-differentiated tissues. Lastly, in case (iii), we blocked cell proliferation and kept the differentiation rates. We also increased the maximum number of neurites, thus resembling a normal process of neuronal differentiation. Given the stochastic nature of our model, we performed multiple simulations for each of the previously mentioned conditions and assessed the results to identify the conditions that enable and promote the rosette formation process.

Some representative results for this study are depicted in Figure 5.7. In undifferentiated tissues, labelled as case (i), cells formed more homogeneous tissues, with many cell bodies and a constrained amount of extracellular space. On the other end of the spectrum, differentiated tissues were characterized by regions of low cell density with abundant neuropil and extracellular space. We did not observe the presence of HW rosettes in either of these tissue types (i and iii). However, in poorly-differentiated tissues (ii), the balance between proliferation and differentiation promoted the formation of HW rosettes, as cell density levels enabled interactions between cells and neurites. Moreover, we observed that the maximum number of neurites was a relevant factor in the rosettes' radial structure maintenance.

Interestingly, we observed that cell bodies could be organised around a central empty core in poorly-differentiated and undifferentiated tissues. As previously explained in Section 1, rosette patterns that surround an empty lumen are usually classified as "true ependymal rosettes." Although these structures were the result of the mechanical interactions between cell bodies and cell proliferation in our model, it is relevant to highlight that studies attribute the formation of this

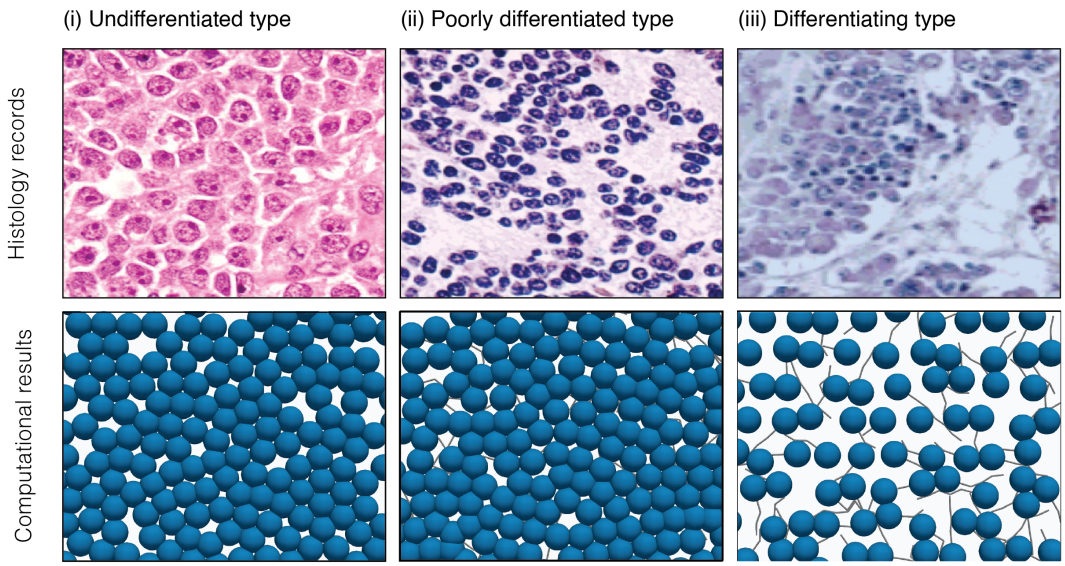

Figure 5.7: **Histological records and simulation data for different stages of neuronal differentiation.** All simulations were initialized with neurons with a cell body and no neurites in a 7x7 hexagonal grid. We analysed three groups of simulations, each corresponding to a type of tissue, namely undifferentiated, poorly-differentiated and differentiating tissues. Neuron biological clocks were programmed to create the structure of these tissues by: (a) blocking differentiation, (b) combining proliferation and differentiation, and (c) blocking proliferation. Results were evaluated by visually identifying the existence of HW rosettes which were only observed in poorly-differentiated tissues. Histology images were taken from (Wang et al., 2013) (A) and (Shimada et al., 1999a) (B, C).

rosette subtype to a more complex process. Several studies have thoroughly studied and modelled the formation of true ependymal rosettes in vitro and associated it with ependymal differentiation. During this process, cells become more elongated and present a polar geometry, forming an empty lumen in the middle of a group of cells (Wilson and Stice, 2006).

Lastly, even though HW rosettes were induced by the balanced levels of proliferation and differentiation, we still observed that this process was very stochastic and could not be predicted easily. We attribute this to the fact that, in contrast with the studies presented in the previous sections, neurite outgrowth was assumed to be a random process, i.e., the direction of growth of new neurites was chosen randomly. Thus, we did not consistently observe significant interactions between neurites, and could not confirm that neurites from different cells became tangled to form the rosette patterns (Wippold and Perry, 2006). Instead, we observed that new neurites appeared in intercellular spaces with a circular geometry. In this study, we investigated whether HW rosettes could be produced in the absence of a distinct chemical signal by establishing that neuronal development would follow no preferential orientation. Our findings led us to the conclusion that these structures could appear spontaneously due to mechanical interactions and differentiation rates, but their morphology was not as well-defined as in the studies presented in the previous sections. Accordingly, when additional data about these chemical gradients becomes available, signalling may be a pertinent component to incorporate in future iterations of this computational framework.

5.4 Summary

This chapter presents the last component of our computational approach to simulate collective cell behaviour in a bottom-up approach. In particular, we created a framework that integrated cell motility and mechanics, proliferation and neuronal differentiation. We achieved this by implementing a mechanics-based modelling framework to study the dynamics of the formation of HW rosette patterns in the early development of neuroblastoma. The occurrence of rosettes in cancer progression is still not properly understood and there is a shortage of experimental data to support the existing assumptions about this developmental process due to the difficulties in recreating these biological conditions in vitro (Wippold and Perry, 2006). Numerical and computational simulations are of great value in these scenarios because they can be used to simulate these morphological patterns in a virtual environment without these constraints. Using our computational framework, we evaluated how distinct factors modulated this process and identified that this collective behaviour was the result of an intricate interplay between cell-cell mechanics, proliferation and differentiation.

To achieve this, our model included cell mechanics and different biological events, namely cell division and neuronal differentiation. In addition, we examined distinct conditions to showcase the model's ability to simulate the formation of HW rosettes. Firstly, we simulated an ideal scenario of the assembly of these structures as described by the current theoretical hypotheses that explain their formation. Subsequently, we performed a numerical study to evaluate how mechanical interactions and differentiation rates regulated the formation of rosette-like structures using the same

ideal configuration. Lastly, we conducted some larger-scale simulations in which we investigated the presence of HW rosettes in tissues with distinct proliferation and differentiation dynamics (i.e., undifferentiated, poorly-differentiated, and differentiating tissues). We hypothesize that stages of poor neuronal differentiation in our model created a balance between proliferation and differentiation levels. Consequently, differentiation modulated the formation and presence of HW rosettes. Furthermore, model parameters that regulated tissue mechanics also played a relevant role in regulating the morphology of HW rosettes.

One of the limitations of our approach is that Python is a less powerful programming language than other low-level languages, such as C and C++. Therefore, the current implementation may require long computation times for simulations at the tissue-level. However, the scope of our study was to analyse the formation of spatial structures in small tissue regions that could be compared to histological pictures. Accordingly, we consider that our framework's flexibility, ease of use and simulation rendering capabilities outweigh its limitations.

CHAPTER 6

Conclusions

Contents

6.1 Introduction . . . 107
6.2 Main conclusions . . . 108
6.2.1 Single cell motility . . . 108
6.2.2 Tumour growth . . . 109
6.2.3 Pattern formation . . . 110
6.3 Future work . . . 111
6.3.1 Improved representations of the physical microenvironment . . . 111
6.3.2 Integrated intracellular models . . . 112
6.3.3 Optimization and parameter identifiability . . . 113
6.4 Thesis contributions . . . 114
6.4.1 Journal publications . . . 114
6.4.2 Conferences and workshops . . . 114
6.4.3 Supervising and mentoring activities . . . 115
6.4.4 Open-source software . . . 116
6.4.5 Collaborations . . . 116
6.4.6 Acknowledgements . . . 117

6.1 Introduction

This chapter gives a comprehensive summary of the dissertation's primary findings as well as the contributions made to this field of study over the course of the candidate's PhD studies. The overall conclusions of this work are discussed in Section 6.2, which also provides a summary of the specific findings presented in each chapter. Then, in Section 6.3 some of the shortcomings of the work done for this dissertation are discussed, along with suggestions for improvements and potential directions for future research.

6.2 Main conclusions

The research presented in this dissertation demonstrated the capability of computational modelling to deconstruct and comprehend challenging issues in cancer biology. We specifically demonstrated how model elements could be utilized to represent particular features of cancer biology that can be optimized and simulated using experimental data. Furthermore, the need for modelling multiscale biological events was explored, and it was demonstrated that particular cell behaviours were only observed when multiple behaviours were simulated simultaneously. In particular, we observed that tumour growth resulted from cell proliferation, cell-cell mechanics and single-cell motility. Furthermore, morphological patterns resulted from balanced mechanics, proliferation and differentiation. Accordingly, this dissertation highlights the relevance of integrating simple and well-characterized biological models into mechanistic multiscale frameworks that can combine them and simulate how complex behaviour arises.

6.2.1 Single cell motility

In Chapter 3, the first computational model of this dissertation was presented. We designed this model to simulate single-cell motility and how it is modulated by the mechanical and chemical properties of the extracellular microenvironment. In addition, we employed model calibration techniques to reproduce empirical data from *in vitro* studies, which allowed us to pose hypotheses of mechanisms that might explain the observed biological behaviours. Two studies were performed to investigate this research question, the first of which relied on qualitative data to describe the role of drag and cell-generated forces on cancer cell motility, while the latter integrated original quantitative *in vitro* results to assess how fibroblast motility changed in response to chemical factors.

In the first study presented in this chapter, we implemented the computational model that served as the first template for the subsequent iterations presented in this dissertation. We initiated this model implementation process with a straightforward approach to characterize individual cell motility, which we defined as this work's first building block. Accordingly, we simulated migration as the result of cell-generated locomotive forces regulated by the mechanical properties of the ECM. As a first approach, we assumed a simplified homogeneous representation of the microenvironment. Furthermore, we estimated and derived a force generator module to reproduce cell movement. With this approach, we were able to confirm that collagen density limited cell motility.

In the second application, we extended our model and designed a calibration pipeline to interpret experimental data from migration assays conducted with growth factors and different collagen densities. Generally, the experimental observations indicated that the presence and spatial location of PDGF-BB in the microfluidic devices regulated motility. Using statistical inference, we identified model parameters and values that can justify these findings. In particular, we predicted that collagen invasion was not solely modulated by an increase in cell forces and cell movement

due to the presence of PDGF-BB, as seen in studies that consider chemokinesis, but it was also influenced by enhanced directionality and cell persistence through chemotaxis.

Overall, these two models of single-cell motility collectively illustrate the process of developing and fine-tuning a computational framework. We could translate components from a first implementation to an improved second approach which was complemented and strengthened through model extensions and optimization techniques. We consider the second approach more complete, providing more information and, thus, more confidence to both modellers and experimentalists to interpret its outputs. Nevertheless, we also recognize that a simple and generalizable framework such as the first one can be meaningful in research, specifically when quantitative data is not accessible to permit data-driven optimization routines.

6.2.2 Tumour growth

In Chapter 4, we took the first step to move from individual to collective cell behaviour by studying tumour growth and how it is affected by cell dynamics and cell-cell interactions. With this goal in mind, we extended the computational model implemented in Chapter 3 and simulated the growth of multicellular structures from individual cells. We based this model on experimental previously published data obtained for NSCLC cells seeded in microfluidic devices with collagen matrices of different densities. We simulated this experimental setup with a simulation workflow similar to that implemented in Chapter 3. Yet, collective behaviour was investigated by enabling proliferation and cell-cell and cell-matrix mechanics. As observed in the original experiments, our results showed that high-density matrices produced large multicellular structures. Conversely, low-density matrices resulted in sparse and individual cells and small clusters. These findings were interesting since that several previous studies link high-density matrices to compressive forces that lead to limited growth.

Remarkably, our computational approach revealed that, in spite of cell-cell interaction forces and the cells' ability to proliferate, motility played a dominant role in the growth of multicellular structures. Accordingly, we concluded that cell behaviour arises from complex interactions between different model components and that it is essential to acknowledge both cell motility and cell proliferation to simulate tumour development accurately. Consequently, it is highly relevant to consider the mechanical properties of tissues when studying metastatic colonization and how they modulate cell motility, given that it can impact the development of secondary tumours. Furthermore, we observed that our computational framework had a limited ability to accurately depict cluster eccentricity results observed *in vitro*. However, we were able to take this information into account and generate new hypotheses to explain that elongated morphologies may be a result of collective cell motility patterns, which were not modelled in this work and would require cell-cell signalling. Henceforth, despite its limitations, our model was still a critical resource to identify that our assumptions were insufficient to reproduce the intended biological behaviour and that new model assumptions should be added to re-evaluate model adequacy and accuracy.

Moreover, we reviewed models of avascular tumour growth, focusing on how these previous works contemplated the presence of nutrients and chemical factors involved in glucose

metabolism. This review was relevant to understand the current state-of-the-art in tumour growth modelling, specifically in multiscale modelling of glucose metabolism and its impact on tumour development. Overall, we concluded that although several models of avascular tumour growth at the cell-level exist and network models have been employed simulate glucose metabolism at different complexity levels, few models combine these two scales. Therefore, we believe it would be highly relevant to include subcellular representations of glucose metabolism in future iterations of out computational framework.

6.2.3 Pattern formation

The model development process carried out in this dissertation culminated in Chapter 5, where we presented computational framework that integrated cell motility, cell-cell interactions, proliferation and, lastly, differentiation. We focused on modelling HW rosettes, a particular case of pattern formation observed in neuroblastoma. We implemented modules for simulating cell cycling, differentiation, contact inhibition and cell mechanics. Moreover, our approach to modelling neurons as a combination of two independent types of physical objects, namely spheres and cylinders, allowed us to obtain a more realistic depiction of neuron morphology than a standard sphere-based CBM would enable. Once more, we observed that the biological behaviour under study resulted from the interplay between different model components, namely cell-cell mechanics, proliferation and differentiation.

Moreover, in this work, we proposed some simple simulation setups to evaluate the mechanisms that lead to rosette assembly. Firstly, we implemented an ideal scenario where cells were placed in a circular arrangement since the start of the simulation and differentiation was allowed, but limited, to reproduce the formation of poorly-differentiated neurites. This simulation was designed to emulate the current theories that explain the formation of HW rosettes and could be used to evaluate if the model was behaving as expected. Furthermore, a numerical study was implemented to use this simulation design and quantify how different cell-cell and differentiation levels affected rosette morphology. With this methodology, we identified that fast differentiation rates compromised the structure of HW rosettes, as did high cell-cell adhesion levels. Lastly, we progressed from simulations with a single rosette to tissue-scale scenarios and studied more realistic cell distributions, initializing cells into a grid. In this study, we aimed to reproduce the three distinct neuroblastoma phenotypes, i.e., undifferentiated, poorly-differentiated and differentiated. Thus, we evaulated the effect of differentiation rates and maximum number of neurites. As expected, based on the reports available in the literature, we only observed HW-like structures in the poorly differentiated subtype.

We believe this model is a critical first step towards understanding HW rosettes and the mechanisms that lead to their formation. Despite being recognized in clinical diagnosis as a hallmark of poorly-differentiated neuroblastoma, HW are, overall, not sufficiently understood. For instance, only recently did *in vitro* experiments reproduce spatial patterns that share similarities with HW rosettes (Duarte Campos et al., 2019). Besides, the theoretical literature on these structures is very sparse and limited. Thus, to the best of our knowledge, this is not only the first computational

model of HW rosette formation, but also a novel resource that can be of interest to experimentalists and clinicians to visualize this process and plan new experiments to evaluate it *in vitro*, which would result in new findings in this uncharted research field.

6.3 Future work

The research questions studied in this dissertation enabled advances in the computational oncology research field. However, due to the fast developments in this research field, multiple aspects can be further investigated and enhanced. Generally, we remark that further characterization of model performance could be of value. In Chapters 3 and 4, we relied on *PhysiCell*'s ability to scale and efficiently compute the dynamics of multiple substances and cells. Overall, we did not experience a substantial increase in computational times due to our numerical implementations. In Chapter 5, we implemented a Python model, less computationally powerful than *PhysiCell*, that could be improved by further parallelization, for instance. In the context of our study, the scale of the tissues did not compromise computation times, yet this should be improved in simulations with more agents (Breitwieser et al., 2023). Furthermore, we identified three critical aspects that could be improved in future iterations of this work, as described in the following sections.

6.3.1 Improved representations of the physical microenvironment

In spite of using *PhysiCell*, a hybrid modelling framework, extensively, the focus of this dissertation was mainly on the discrete model components. In other words, we studied cell behaviour in detail and took microenvironment cues into account in our models, yet we did so through simplified representations of the environment. In the future, it would be meaningful to improve the projects developed in this dissertation to account for more realistic models of the mechanical and chemical properties of the extracellular microenvironment. Starting with the mechanical properties, we could represent the heterogeneity of collagen matrices through a model that might relate to a local collagen concentration with local viscosity values and reproduces the heterogeneous nature of the ECM. For instance, further characterization experiments as those performed by Valero et al. (2018) could be conducted to obtain more data, with which it would be possible to understand a more accurate description of this relationship. Besides matrix density, additional properties, e.g., anisotropy and orientation, could also be included in this model (Metzcar et al., 2022).

Moreover, it would be advantageous to include the effects of matrix secretion and degradation by cells, which we did not consider in this work (Shuttleworth and Trucu, 2020; Merino-Casallo et al., 2022b). In fact, it is currently possible to simulate ECM secretion and degradation with *PhysiCell*'s built-in functions. Therefore, through this model, we could also study how cells modulate the ECM's mechanical properties and how, in turn, this effect regulates cell migration and tumour formation. The idea of simulating ECM fibres as discrete agents is also of interest, despite the increased computational costs that it entails. Using a *PhysiCell* addon called *PhysiBoSS* (Letort et al., 2019), it is possible to simulate the ECM as agents with a spherical geometry. Furthermore, it could be an option to implement fibres as the cylinder objects used to simulate

neuronal processes in Chapter 5. This approach has been shown to be more realistic than previous sphere-based frameworks (Macnamara et al., 2020) and could also be applied to simulate blood vessels (Breitwieser et al., 2022).

At the chemical level, explicitly modelling reaction and diffusion could provide more realistic descriptions of the changes in substance concentration as a dynamic process in space and time. In Chapter 3, it was assumed that PDGF-BB gradients were mostly unaffected by cell consumption given that the cell number was small in comparison to the microfluidic device dimensions and because the chemotactic gradient was maintained by introducing PDGF-BB to one of the media channels every two days. Nevertheless, in the case of tumour growth and MCTSs in particular, as discussed in Chapter 4, it becomes crucial to be able to simulate local changes in concentration to reproduce the internal structure of spheroids (Murphy et al., 2022; Jagiella et al., 2016) and other spatial patterns commonly observed in nutrient deprived regions, e.g., finger-like invasion patterns that arise from migrating cells trying to escape the tumour (Rocha et al., 2021; Anderson et al., 2006; Colombi et al., 2017). In addition, if fibres were modelled as agents, it could be relevant to have these objects behave as sinks to simulate the binding of molecules to the ECM, which has been shown to occur in *in vitro* studies (Moreno-Arotzena et al., 2014) but was disregarded in this work.

6.3.2 Integrated intracellular models

In Chapter 4, we discussed tumour growth models that reproduce nutrient availability and glucose metabolic pathways. This review was a theoretical overview, and we did not model cell metabolism or the effects of chemical substances on cell cycling. However, our computational framework can be extended to include representations of glucose and its metabolic pathways. In fact, Jacquet and Stéphanou (2023) recently investigated glucose metabolic pathways in *PhysiCell* using a reduced model of anaerobic-aerobic metabolism in altered tumour microenvironments. Furthermore, recent advances in the *PhysiCell* ecosystem enable the creation and reuse of subcellular models written in Systems Biology Markup Language (SBML), a standardized format to save and share computational models of biological processes (Heiland, 2020). Therefore, previously published models can be loaded into *PhysiCell*, thus enhancing model reproducibility. Other open-source modelling platforms, e.g., CompuCell3D, offer functionalities to read and run SBML models (Swat et al., 2015) as well, which we consider being revealing of its capabilities and potential to be used in a standardized manner by the mathematical biology community.

Furthermore, intracellular networks are not limited to metabolic models. In particular, signal transduction pathways can also be simulated at the subcellular level (Letort et al., 2019; Tyson et al., 2003). For instance, neuronal differentiation in neuroblastoma is regulated by signal-transducing pathways modulated by the presence of Schwann cells and growth factors (Weiss et al., 2021; Persson et al., 2017). In addition, it has been revealed that the mechanical properties of the ECM, namely its stiffness, may promote tumour differentiation, slowing down proliferation and modulating the expression of transcription factors associated with high-risk neuroblastoma

(Lam et al., 2010). Henceforth, future iterations of our neuroblastoma model could include subcellular differentiation models, which would allow more control over which conditions could lead to poorly-differentiated states. It would be particularly relevant to simulate how these signalling pathways are activated through the inclusion of subcellular model components to account for cell-cell junctions and cell-matrix adhesion complexes (McEvoy et al., 2022; Vargas et al., 2020).

6.3.3 Optimization and parameter identifiability

Throughout this dissertation, we strived to calibrate our models with experimental data to enhance their biological relevance. Furthermore, we developed some additional work on optimization of *PhysiCell* models. We specifically developed *PhysiCOOL* (Gonçalves et al., 2022), a tool that makes the studies presented in this dissertation more accessible to the general mathematical biology community. *PhysiCOOL* is an open-source Python library tailored to create standardized calibration and optimization routines of *PhysiCell* models. Besides adding features that provide a user-friendly interface to run *PhysiCell* studies through Python, it also introduces optimization routines based on multilevel grid-search that fits model parameters to user-defined data. A detailed description of *PhysiCOOL*'s implementation and use cases is presented in Appendix C. Nonetheless, we belive this framework can still be improved, for example, through the implementation of distinct optimization techniques and the additional features to quantify parameter uncertainty and identifiability.

ABMs are highly complex and are many times characterized by a large number of parameters that describe biological phenomena yet cannot be experimentally quantified and need to be inferred (Gábor and Banga, 2015). Besides, biological data is intrinsically noisy, making parameter estimation a more challenging function. Hence, parameter identifiability analysis, which can determine if model parameters can be obtained from a given dataset (Maclaren and Nicholson, 2019), and uncertainty analysis, which quantifies how parameter uncertainty affects model outputs (Vernon et al., 2018), are increasingly relevant to provide model developers with more confidence about their models and model results. Simpler models with fewer parameters can evade these issues since they can be characterized at lower computational costs. Consequently, new surrogate modelling techniques have been developed to build representations of complex non-identifiable models that may not be directly associated with biological mechanisms, but can be tested throughoutly and provide insights translatable to the original mechanistic models (Browning and Simpson, 2023; Kennedy, 2000). In the future, it would certainly be relevant to apply these techniques to our computational frameworks to confirm our hypotheses and evaluate model accuracy.

6.4 Thesis contributions

6.4.1 Journal publications

Published work

1. Inês G. Gonçalves and José Manuel Garcia-Aznar. **"Extracellular matrix density regulates the formation of tumour spheroids through cell migration."** PLoS computational biology 17.2 (2021): e1008764. doi: 10.1371/journal.pcbi.1008764.

2. Inês G. Gonçalves and Jose Manuel Garcia-Aznar. **"Hybrid computational models of multicellular tumour growth considering glucose metabolism."** Computational and Structural Biotechnology Journal 21 (2023): 1262-1271. doi: 10.1016/j.csbj.2023.01.044.

3. Inês G. Gonçalves, Nieves Movilla, Carlos Borau and José Manuel Garcia-Aznar. **"A novel integrated experimental and computational approach to unravel fibroblast motility in response to chemical gradients in 3D environments."** Integrative Biology zyad002 (2023). doi: 10.1093/intbio/zyad002.

4. Inês G. Gonçalves, David A. Hormuth II, Sandhya Prabhakaran, Caleb M. Phillips, José Manuel García-Aznar. **"PhysiCOOL: A generalized framework for model Calibration and Optimization Of modeLing projects"** GigaByte (2023). doi: 10.46471/gigabyte.77.

Under review

1. Inês G. Gonçalves and José Manuel Garcia-Aznar. **"*neurorosettes:* A novel computational modelling framework to explore the formation of Homer-Wright rosettes in neuroblastoma."** Computational Particle Mechanics

6.4.2 Conferences and workshops

Oral communications

1. **"The role of cell migration on tumour spheroid growth"**.
   VPH2020 - Virtual Physiological Human. Paris, France (online).

2. **"Collagen Density Regulates Tumour Spheroid Growth through Cell Motility"**.
   16th U.S. National Congress on Computational Mechanics. Chicago, USA (online).

3. **"Combined *in vitro* and computational models of neuroblastoma growth"**.
   11th European Solid Mechanics Conference. Galway, Ireland.

4. **"2021 Workshop Highlights: PhysiCOOL"**.
   2022 PhysiCell Workshop and Hackathon. USA (online).
   Invited oral communication.

5. **"Locomotive forces and chemical gradients regulate human fibroblast motility in 3D environments: Insights from an integrated experimental and computational approach"**.
   Hausdorff Summer School "Inverse problems for multi-scale models". Bonn, Germany.
   Recipient of a travel grant.

6. **"Agent-based models for biomedical applications: An introduction to PhysiCell"**.
   XI MEETING OF THE SPANISH CHAPTER OF THE EUROPEAN SOCIETY OF BIOMECHANICS. Zaragoza, Spain.
   Lecturer of a hands-on workshop.

Poster communications

1. **"Collagen Density Regulates Tumour Spheroid Growth Through Cell Motility: A Computational Study"**.
   II Jornada del I3A y X Jornada de Jóvenes Investigadores. Zaragoza, Spain.

Participation

1. *PDBEB Courses 2021/2022 - Computational Biology*. Coimbra, Portugal (online).
   Presentation of a short project titled "Using text-mining techniques in literature reviews".

2. *2021 Network Modeling Virtual Summer School & Hackathon.* USA (online).

3. *2021 PhysiCell Workshop and Hackathon.* USA (online).
   Part of the runner-up team for the "Best Tool" prize in the Hackathon.

4. *Short Course on Particle-Based Methods in Engineering and Applied Science.* Hamburg, Germany (online)

6.4.3 Supervising and mentoring activities

MSc thesis

1. "*In silico* modeling of single-cell fibroblast migration in response to 3D chemical gradients." Ijeoma Charity Asilebo. 2020/2021.

2. "Computational simulation of the T cell migration process using a discrete model." Imán Laga Boul-Atarass. 2021/2022.

3. "Modelling of tumoral cells metabolism in tumoral spheroids formation and development." Alejandro Modrego Bravo. 2022/2023

BSc final project

1. "Simulation by methods based on agents of immunotherapy techniques against cancer." Guillermo Villamor Aparicio. 2022/2023

6.4.4 Open-source software

1. **"*PhysiCell-ECM*:** A PhysiCell extension to account for the extracellular matrix".
   Available on GitHub[1] (MIT License).

2. **"*pdgf-induced-motility*:** Computational model of chemotaxis calibrated with Bayesian Optimization."
   Available on GitHub[2] (BSD-3-Clause License).

3. **"*neurorosettes*:** An agent-based framework to model the formation of rosette patterns in tissues of the nervous system."
   Available on Zenodo (BSD-3-Clause License) (Gonçalves and García-Aznar, 2022).

4. **"*PhysiCOOL*:** A generalized framework for model Calibration and Optimization Of modeLing projects".
   Available on Zenodo (BSD-3-Clause License) (Gonçalves et al., 2023).

6.4.5 Collaborations

PhysiCell Workshop 2021

During their doctoral studies, in 2021, the PhD candidate participated in an online workshop and hackathon organized by the team responsible for developing *PhysiCell* (Paul Macklin's Math Cancer Lab, Indiana University). During this hackathon, the PhD candidate joined a team to implement *PhysiCOOL*, a tool designed to calibrate and optimize *PhysiCell* models. The team was composed by the PhD candidate, Inês G. Gonçalves, David Hormuth II and Caleb M. Phillips (members of the Center for Computational Oncology group at the Oden Institute in Texas) and Sandhya Prabhakaran (member of the Integrated Mathematical Oncology department at H.Lee Moffitt Cancer Center and Research Institute in Florida) and they were awarded the "Best Tool" runner-up prize at the end of the hackathon. Since then, the team has collaborated to improve this tool and make it a standard framework in the *PhysiCell* ecosystem through the development of new features, creation of documentation and interactive examples and publishing and distribution of the software as a publicly-available a Python library. Moreover, this work has been accepted for publication and can be found on bioRxiv (Gonçalves et al., 2022).

MAtrix Group - KU Leuven

Furthermore, during the first year of their PhD, the candidate collaborated with the research group where they completed their MSc thesis, supervised by Diego A. Vargas and Hans Van Oosterwyck (members of the Mechanobiology and Tissue Engineering group at KU Leuven), to publish this work. In particular, the candidate studied the role of mechanosensing mechanisms in single-cell motility and how substrate stiffness and cell-substrate adhesions influenced the cell's migration

[1] https://github.com/m2be-igg/PhysiCell-ECM
[2] https://github.com/m2be-igg/pdgf-induced-motility

phenotype. Therefore, the knowledge acquired in this project was of particular relevance to this PhD dissertation and could be applied to model and evaluate migration patterns in single-cell motility as a response to the properties of the surrounding environment. Furthermore, this publication resulted in the publication of an original article:

1. Vargas, D. A., Gonçalves, I. G., Heck, T., Smeets, B., Lafuente-Gracia, L., Ramon, H., & Van Oosterwyck, H. **"Modeling of mechanosensing mechanisms reveals distinct cell migration modes to emerge from combinations of substrate stiffness and adhesion receptor–ligand affinity."** Frontiers in Bioengineering and Biotechnology 8 (2020): 459.

LEPABE/BioSIM - Faculty of Engineering of University of Porto

Lastly, during the first two years of their doctoral studies, the PhD candidate established a collaboration to continue and finish a previous research project developed under the supervision of Manuel Simões (member of the LEPABE research group at the Faculty of Engineering of the University of Porto). This collaboration was not directly linked to the candidate's doctoral studies and was developed in parallel to the PhD thesis but was relevant to the candidate's academic path. In this work, the PhD candidate aimed to build an online database with data from peer-reviewed reports of *Legionella* outbreaks to identify the environmental factors that may have promoted these outbreaks. To achieve this, the student gathered and analysed peer-reviewed data on *Legionella* cases and worked together with members of the Biomolecular SIMulations (BioSIM) research group to implement and publish the online database, *LegionellaDB*[3]. This collaboration also resulted in the publication of one original article and a review paper, listed below:

1. Gonçalves, I. G., Fernandes, H. S., Melo, A., Sousa, S. F., Simões, L. C., & Simões, M. **"LegionellaDB–A database on Legionella outbreaks."** Trends in Microbiology 29.10 (2021): 863-866.

2. Gonçalves, I. G., Simões, L. C., & Simões, M. **"Legionella pneumophila."** In Trends in Microbiology 29.9 (2021): 860–861.

6.4.6 Acknowledgements

This PhD thesis is the result of projects that received funding from the European Research Council (ERC), under the Advanced Grant programme, and the European Union's Horizon 2020 research and innovation programme: CADENCE project (ERC-2016-ADG- grant agreement no. 742684), PRIMAGE project (SC1-DTH-07-2018-grant agreement no. 826494), and ICoMICS project (ERC-2020-ADG-grant agreement no. 101018587).

[3] https://legionelladb.biosim.pt/

APPENDIX A

Conclusiones

Contents

A.1 Introducción . . . 119

A.2 Conclusiones generales . . . 120

A.2.1 Migración celular . . . 120

A.2.2 Crecimiento tumoral . . . 121

A.2.3 Formación de patrones . . . 122

A.3 Trabajo futuro . . . 123

A.3.1 Modelos intracelulares . . . 124

A.3.2 Optimización y identificabilidad de parámetros . . . 125

A.4 Contribuciones . . . 126

A.4.1 Publicaciones . . . 126

A.4.2 Participación en congresos y workshops . . . 126

A.4.3 Dirección de trabajos . . . 127

A.4.4 Software libre . . . 128

A.4.5 Colaboraciones . . . 128

A.4.6 Agradecimientos . . . 130

A.1 Introducción

Este capítulo ofrece un resumen completo de los principales resultados de la tesis, así como las contribuciones realizadas a este campo de estudio a lo largo de los estudios de doctorado de la candidata. Las conclusiones generales de este trabajo se discuten en la Sección A.2, que también proporciona un resumen de los resultados específicos presentados en cada capítulo. Luego, en la Sección 6.3 se discuten algunas de las limitaciones del trabajo realizado para esta tesis, junto con sugerencias para mejoras y propuestas para futuras investigaciones.

A.2 Conclusiones generales

La investigación presentada en esta tesis demostró la capacidad del modelado computacional para deconstruir y comprender cuestiones desafiantes en la biología del cáncer. Se demonstra específicamente cómo los elementos de los modelos se pueden utilizar para representar características particulares de la biología del cáncer, los cuales se pueden optimizar y simular utilizando datos experimentales. Además, se exploró la necesidad de modelar eventos biológicos multiescala y se demostró que solo se observaban comportamientos celulares particulares cuando se simulaban múltiples comportamientos simultáneamente. En particular, se observa que el crecimiento tumoral se debió a la proliferación celular, la mecánica celular y la motilidad unicelular. Además, los patrones morfológicos observados en neuroblastoma resultaron de la mecánica, proliferación y diferenciación. En consecuencia, esta tesis destaca la relevancia de integrar modelos biológicos simples y bien caracterizados en herramientas de múltiples escalas que puedan combinarlos y simular cómo surge un comportamiento complejo.

A.2.1 Migración celular

En el capítulo 3, se presentó el primer modelo computacional de esta tesis. Este modelo fue diseñado para simular la motilidad de una sola célula y como las propiedades mecánicas y químicas del entorno extracelular modulan su comportamiento. Además, se utilizaron técnicas de calibración para reproducir datos empíricos de estudios *in vitro*, lo que nos permitió formular hipótesis sobre mecanismos que podrían explicar los comportamientos biológicos observados. Se realizaron dos estudios para investigar esto, el primero de los cuales se basó en datos cualitativos para describir el papel del arrastre y las fuerzas generadas por las células en la motilidad de las células cancerosas, mientras que el último integró resultados cuantitativos originales *in vitro* para evaluar cómo la motilidad de fibroblastos cambió en respuesta a factores químicos.

En el primer estudio presentado en este capítulo, se implementó el modelo computacional que sirvió como primera instancia para las iteraciones posteriores presentadas en esta tesis. Este proceso de implementación del modelo comenzó con un enfoque para caracterizar la motilidad celular individual, que se definió como el primer bloque de construcción de este trabajo. En consecuencia, se simuló la migración como resultado de las fuerzas locomotoras generadas por las células reguladas por las propiedades mecánicas de la matriz extracelular. Como primera aproximación, se asumió una representación homogénea simplificada del entorno. Además, se estimó y derivó un módulo de generación de fuerza para reproducir el movimiento celular. Con este enfoque, fue posible confirmar que la densidad de colágeno limitaba la motilidad celular.

En la segunda aplicación, este modelo fue ampliado y se diseñó una batería de calibración para interpretar datos experimentales de ensayos de migración realizados con factores de crecimiento y diferentes densidades de colágeno. En general, las observaciones experimentales indicaron que la presencia y la ubicación espacial de PDGF-BB en los dispositivos de microfluídica regulaban la motilidad. Utilizando la inferencia estadística, se identificaron los parámetros y valores del modelo que pueden justificar estos hallazgos. En particular, fue posible predecir que la invasión

de colágeno no solo fue modulada por un aumento en las fuerzas celulares y el movimiento celular debido a la presencia de PDGF-BB, como se vio en estudios que consideraron la quimiocinesis, sino que también estuvo influenciada por una direccionalidad mejorada. persistencia celular por quimiotaxis.

En general, estos dos modelos de motilidad unicelular ilustran colectivamente el proceso de desarrollo y ajuste de un marco computacional. Componentes de una primera implementación fueron aplicadas a un segundo enfoque mejorado que se complementó y fortaleció a través de extensiones de modelo y técnicas de optimización. Este segundo enfoque es más completo, ya que proporciona más información y, por lo tanto, más confianza tanto a los modeladores como a los experimentadores para interpretar sus resultados. Sin embargo, un modelo simple y generalizable como el primero puede ser significativo en la investigación, específicamente cuando no se puede acceder a los datos cuantitativos para permitir las rutinas de optimización basadas en datos.

A.2.2 Crecimiento tumoral

En el Capítulo 4, se definió el primer paso para pasar del comportamiento celular individual al colectivo, estudiando el crecimiento tumoral y cómo se ve afectado por la dinámica celular y las interacciones célula-célula. Con este objetivo en mente, se amplió el modelo computacional implementado en el Capítulo 3 y se simuló el crecimiento de estructuras multicelulares a partir de células individuales. Este modelo se basó en datos experimentales publicados previamente obtenidos para células tumorales sembradas en dispositivos microfluídicos con matrices de colágeno de diferentes densidades. Esta configuración experimental se simuló con un flujo de trabajo de simulación similar al implementado en el Capítulo 3. Sin embargo, se investigó el comportamiento colectivo que permite la proliferación y la mecánica célula-célula y célula-matriz. Como se observó en los experimentos originales, los resultados mostraron que las matrices de alta densidad produjeron grandes estructuras multicelulares. Por el contrario, las matrices de baja densidad dieron como resultado células individuales dispersas y la formación de pequeños grupos celulares. Estos resultados son relevantes, ya que varios estudios previos relacionan matrices de alta densidad con fuerzas de compresión que conducen a un crecimiento limitado.

Sorprendentemente, el enfoque computacional reveló que, a pesar de las fuerzas de interacción célula-célula y la capacidad de las células para proliferar, la motilidad desempeñó un papel dominante en el crecimiento de las estructuras multicelulares. En consecuencia, el comportamiento celular surge de interacciones complejas entre diferentes componentes del modelo y que es esencial reconocer tanto la motilidad celular como la proliferación celular para simular el desarrollo tumoral con precisión. En consecuencia, es muy relevante considerar las propiedades mecánicas de los tejidos al estudiar la colonización metastásica y cómo modulan la motilidad celular, dado que puede impactar en el desarrollo de tumores secundarios. Además, este modelo computacional tenía una capacidad limitada para representar con precisión los resultados de excentricidad de cúmulos observados *in vitro*. Sin embargo, esta información permitió generar nuevas hipótesis para explicar que las morfologías alargadas pueden ser el resultado de patrones de motilidad celular colectivos, que no fueron modelados en este trabajo y requerirían señalización celular. A pesar de

sus limitaciones, este modelo seguía siendo un recurso fundamental para identificar que nuestras hipótesis eran insuficientes para reproducir el comportamiento biológico previsto y que debían agregarse nuevas suposiciones del modelo para reevaluar la adecuación y precisión del modelo.

Además, se revisaron modelos de crecimiento tumoral avascular, centrándonos en cómo estos trabajos previos contemplaban la presencia de nutrientes y factores químicos implicados en el metabolismo de la glucosa. Esta revisión fue relevante para comprender el estado del arte actual en el modelado del crecimiento tumoral, específicamente en el modelado multiescala del metabolismo de la glucosa y su impacto en el desarrollo tumoral. En general, aunque existen varios modelos de crecimiento tumoral avascular a nivel celular y se han empleado modelos de red para simular el metabolismo de la glucosa en diferentes niveles de complejidad, pocos modelos combinan estas dos escalas. Por lo tanto, sería muy relevante incluir representaciones subcelulares del metabolismo de la glucosa en futuras iteraciones de este modelo computacional.

A.2.3 Formación de patrones

El proceso de desarrollo del modelo llevado a cabo en esta tesis culminó en el Capítulo 5, donde se presenta un modelo computacional que integraba la motilidad celular, las interacciones célula-célula, la proliferación y, por último, la diferenciación. En este capítulo se modeló la formación de las rosetas HW, un caso particular de formación de patrones observado en el neuroblastoma. Se implementaron módulos para simular ciclos celulares, diferenciación, inhibición por contacto y mecánica celular. Además, el enfoque para modelar neuronas como una combinación de dos tipos independientes de objetos físicos, a saber, esferas y cilindros, nos permitió obtener una representación más realista de la morfología de las neuronas que la que permitiría un modelo basado en esferas estándar. Una vez más, se observó que el comportamiento biológico bajo estudio resultó de la interacción entre los diferentes componentes del modelo, a saber, la mecánica celular, la proliferación y la diferenciación.

Además, en este trabajo, algunas configuraciones de simulación simples fueron utilizadas para evaluar los mecanismos que conducen al ensamblaje de rosetas. En primer lugar, se implementó un escenario ideal donde las células se colocaron en un arreglo circular desde el inicio de la simulación y se permitió la diferenciación, pero limitada, para reproducir la formación de neuritas pobremente diferenciadas. Esta simulación fue diseñada para emular las teorías actuales que explican la formación de rosetas HW y podría usarse para evaluar si el modelo se estaba comportando como se esperaba. Además, se implementó un estudio numérico para cuantificar cómo los diferentes niveles de diferenciación y interacciones célula-célula afectaban la morfología de las rosetas. Con esta metodología, se identificó que las tasas de diferenciación rápida comprometían la estructura de las rosetas HW, al igual que los altos niveles de adhesión célula-célula. Por último, se pasó de simulaciones con una sola roseta a escenarios a escala de tejido y se estudiaron distribuciones de células más realistas. En este estudio, el objetivo fue reproducir los tres fenotipos distintos de neuroblastoma, es decir, indiferenciado, pobremente diferenciado y diferenciado. Por lo tanto, se evaluó el efecto de las tasas de diferenciación y el número máximo de neuritas. Como era de

esperar, según los informes disponibles en la literatura, solo se observaron estructuras similares a HW en el subtipo pobremente diferenciado.

Este modelo es un primer paso crítico hacia la comprensión de las rosetas HW y los mecanismos que conducen a su formación. A pesar de ser reconocido en el diagnóstico clínico como un sello distintivo del neuroblastoma pobremente diferenciado, las rosetas HW, en general, no se comprenden lo suficiente. Por ejemplo, solo recientemente los experimentos *in vitro* reprodujeron patrones espaciales que comparten similitudes con HW rosetas (Duarte Campos et al., 2019). Además, la literatura teórica sobre estas estructuras es muy escasa y limitada. Por lo tanto, hasta donde se conoce hoy en día, este no es solo el primer modelo computacional de formación de rosetas HW, sino también un recurso novedoso que puede ser de interés para experimentadores y clínicos para visualizar este proceso y planificar nuevos experimentos *in vitro*, lo que daría lugar a nuevos hallazgos en este campo de investigación desconocido.

A.3 Trabajo futuro

A pesar de usar *PhysiCell*, una herramienta de trabajo de modelado híbrido, el enfoque de esta tesis se centró principalmente en los componentes discretos del modelo. En otras palabras, se estudiaron el comportamiento celular en detalle y se tomarón en cuenta las señales del microentorno en estos modelos, pero a través de representaciones simplificadas del entorno. En el futuro, sería significativo mejorar los proyectos desarrollados en esta tesis para dar cuenta de modelos más realistas de las propiedades mecánicas y químicas del entorno celular. Comenzando con las propiedades mecánicas, se podría representar la heterogeneidad de las matrices de colágeno a través de un modelo que podría relacionarse con una concentración de colágeno local con valores de viscosidad y reproducir la naturaleza heterogénea de la matriz. Por ejemplo, se podrían realizar más experimentos de caracterización como los realizados por Valero et al. (2018) para obtener más datos, con los cuales sería posible comprender una descripción más precisa de esta relación. Además de la densidad de la matriz, también podrían incluirse en este modelo propiedades adicionales, por ejemplo, anisotropía y orientación de las fibras (Metzcar et al., 2022).

Además, sería ventajoso incluir los efectos de la secreción y degradación de la matriz por parte de las células, que no fueron consideradas en este trabajo (Shuttleworth and Trucu, 2020; Merino-Casallo et al., 2022b). De hecho, actualmente es posible simular la secreción y degradación de matriz con las funciones integradas de *PhysiCell*. Por lo tanto, a través de este modelo, también se podría estudiar cómo las células modulan las propiedades mecánicas de la matriz y cómo, a su vez, este efecto regula la migración celular y la formación de tumores. La idea de simular fibras de colágeno como agentes discretos también es interesante, a pesar de los mayores costos computacionales que implica. Usando un complemento de *PhysiCell* llamado *PhysiBoSS* (Letort et al., 2019), es posible simular la matriz extracelular como agentes con una geometría esférica. Además, podría ser una opción implementar fibras como objetos cilíndricos utilizados para simular procesos neuronales en el Capítulo 5. Se ha demostrado que este enfoque es más realista que los

marcos anteriores basados en esferas (Macnamara et al., 2020) y también podría aplicarse para simular vasos sanguíneos (Breitwieser et al., 2022).

A nivel químico, modelar explícitamente la reacción y la difusión de las substancias del entorno podría añadir descripciones más realistas de los cambios en la concentración como un proceso dinámico en el espacio y el tiempo. En el capítulo 3, se asumió que los gradientes de PDGF-BB no se vieron afectados en su mayoría por el consumo de células, dado que el número de células era pequeño en comparación con las dimensiones del dispositivo de microfluídica y porque el gradiente quimiotáctico se mantuvo constante. Sin embargo, en el caso del crecimiento tumoral, y en esferoides en particular, como se discutió en el Capítulo 4, se vuelve crucial poder simular cambios locales en la concentración para reproducir la estructura interna de los esferoides (Murphy et al., 2022; Jagiella et al., 2016), bien como otros patrones espaciales comúnmente observados en regiones privadas de nutrientes, por ejemplo, patrones de invasión similares a protusiones que surgen de células migratorias que intentan escapar del tumor (Rocha et al., 2021; Anderson et al., 2006; Colombi et al., 2017). Además, si las fibras se modelaran como agentes, podría ser relevante que estos objetos se comportaran como sumideros para simular la unión de moléculas a la ECM, lo que se ha demostrado que ocurre en estudios *in vitro* (Moreno-Arotzena et al., 2014), pero no fue considerado en este trabajo.

A.3.1 Modelos intracelulares

En el Capítulo 4, los modelos de crecimiento tumoral que reproducen la disponibilidad de nutrientes y las vías metabólicas de la glucosa fueron revisados. Esta revisión fue una descripción general teórica y el metabolismo celular no fue modelado, ni los efectos de las sustancias químicas en el ciclo celular. Sin embargo, nuestro modelo computacional se puede ampliar para incluir representaciones de la glucosa y sus vías metabólicas. De hecho, Jacquet and Stéphanou (2023) investigó recientemente las vías metabólicas de la glucosa en *PhysiCell* usando un modelo reducido de metabolismo anaeróbico-aeróbico en entornos tumorales. Además, los avances recientes en el ecosistema *PhysiCell* permiten la creación y reutilización de modelos subcelulares escritos en SBML, un formato estandarizado para guardar y compartir modelos computacionales de procesos biológicos (Heiland, 2020). Por lo tanto, los modelos publicados anteriormente se pueden cargar en *PhysiCell*, mejorando su reproducibilidad. Otras plataformas de modelado de código abierto, por ejemplo, CompuCell3D, también ofrecen funcionalidades para leer y ejecutar modelos SBML (Swat et al., 2015), lo que es revelador de sus capacidades y potencial para ser utilizado de manera estandarizada.

Además, las redes intracelulares no se limitan a modelos metabólicos. En particular, las vías de transducción de señales también se pueden simular a nivel subcelular (Letort et al., 2019; Tyson et al., 2003). Por ejemplo, la diferenciación neuronal en el neuroblastoma está regulada por vías de transducción de señales moduladas por la presencia de células de Schwann y factores de crecimiento (Weiss et al., 2021; Persson et al., 2017). Además, se ha revelado que las propiedades mecánicas de la matriz extracelular, como su rigidez, pueden promover la diferenciación tumoral, ralentizar la proliferación y modular la expresión de factores de transcripción asociados con el

neuroblastoma de alto riesgo (Lam et al., 2010). De ahora en adelante, las iteraciones futuras de nuestro modelo de neuroblastoma podrían incluir modelos de diferenciación subcelular, lo que permitiría un mayor control sobre qué condiciones podrían conducir a estados pobremente diferenciados. Sería particularmente relevante simular cómo se activan estas vías de señalización mediante la inclusión de componentes del modelo subcelular para dar cuenta de las uniones célula-célula y los complejos de adhesión célula-matriz (McEvoy et al., 2022; Vargas et al., 2020).

A.3.2 Optimización y identificabilidad de parámetros

A lo largo de esta tesis, nuestros modelos fueron calibrados con datos experimentales para mejorar su relevancia biológica. Además, se desarrolloró un trabajo adicional sobre la optimización de los modelos, llamado *PhysiCOOL* (Gonçalves et al., 2022). Esta herramienta hace que los estudios presentados en esta tesis sean más accesibles para la comunidad general de biología matemática. *PhysiCOOL* es una biblioteca Python de código abierto diseñada para crear rutinas estandarizadas de calibración y optimización de modelos *PhysiCell*. Además de agregar características que brindan una interfaz fácil de usar para ejecutar estudios de *PhysiCell* a través de Python, también presenta rutinas de optimización que ajustan los parámetros del modelo a los datos definidos por el usuario. En el Apéndice C se presenta una descripción detallada de su implementación y sus casos de uso. No obstante, esta herramienta aún se puede mejorar, por ejemplo, mediante la implementación de distintas técnicas de optimización y de funciones adicionales para cuantificar la incertidumbre y la identificabilidad de los parámetros.

Los modelos basados en agentes son altamente complejos y muchas veces se caracterizan por una gran cantidad de parámetros que describen fenómenos biológicos que aún no pueden cuantificarse experimentalmente y deben inferirse (Gábor and Banga, 2015). Además, los datos biológicos son intrínsecamente ruidosos, lo que hace que la estimación de parámetros sea un reto. Por lo tanto, el análisis de identificabilidad de parámetros, que puede determinar si los parámetros del modelo se pueden obtener de un conjunto de datos determinado (Maclaren and Nicholson, 2019), y el análisis de incertidumbre, que cuantifica cómo la incertidumbre de los parámetros afecta los resultados del modelo (Vernon et al., 2018), son cada vez más relevantes para proporcionar a los desarrolladores de modelos con más confianza acerca de sus modelos y resultados del modelo. Los modelos más simples y con menos parámetros pueden evadir estos problemas, ya que se pueden caracterizar a costos computacionales más bajos. En consecuencia, se han desarrollado nuevas técnicas de modelado para construir representaciones de modelos complejos no identificables. Estas herramientas pueden no estar directamente asociadas con mecanismos biológicos, pero pueden probarse en su totalidad y proporcionar información traducible a los modelos mecanicistas originales (Browning and Simpson, 2023; Kennedy, 2000). En el futuro, sin duda sería relevante aplicar estas técnicas a nuestros modelos computacionales para confirmar nuestras hipótesis y evaluar su precisión.

A.4 Contribuciones

A.4.1 Publicaciones

Artículos publicados

1. Inês G. Gonçalves and José Manuel Garcia-Aznar. **"Extracellular matrix density regulates the formation of tumour spheroids through cell migration."** PLoS computational biology 17.2 (2021): e1008764. doi: 10.1371/journal.pcbi.1008764.

2. Inês G. Gonçalves and Jose Manuel Garcia-Aznar. **"Hybrid computational models of multicellular tumour growth considering glucose metabolism."** Computational and Structural Biotechnology Journal 21 (2023): 1262-1271. doi: 10.1016/j.csbj.2023.01.044.

3. Inês G. Gonçalves, Nieves Movilla, Carlos Borau and José Manuel Garcia-Aznar. **"A novel integrated experimental and computational approach to unravel fibroblast motility in response to chemical gradients in 3D environments."** Integrative Biology zyad002 (2023). doi: 10.1093/intbio/zyad002.

4. Inês G. Gonçalves, David A. Hormuth II, Sandhya Prabhakaran, Caleb M. Phillips, José Manuel García-Aznar. **"PhysiCOOL: A generalized framework for model Calibration and Optimization Of modeLing projects"** GigaByte (2023). doi: 10.46471/gigabyte.77.

Artículos en revisión

1. Inês G. Gonçalves and José Manuel Garcia-Aznar. **"*neurorosettes:* A novel computational modelling framework to explore the formation of Homer-Wright rosettes in neuroblastoma."** Computational Particle Mechanics

A.4.2 Participación en congresos y workshops

Participaciones con presentación oral

1. **"The role of cell migration on tumour spheroid growth"**.
   VPH2020 - Virtual Physiological Human. Paris, France (online).

2. **"Collagen Density Regulates Tumour Spheroid Growth through Cell Motility"**.
   16th U.S. National Congress on Computational Mechanics. Chicago, USA (online).

3. **"Combined *in vitro* and computational models of neuroblastoma growth"**.
   11th European Solid Mechanics Conference. Galway, Ireland.

4. **"2021 Workshop Highlights: PhysiCOOL"**.
   2022 PhysiCell Workshop and Hackathon. USA (online).
   Comunicación oral invitada.

5. **"Locomotive forces and chemical gradients regulate human fibroblast motility in 3D environments: Insights from an integrated experimental and computational approach"**.
   Hausdorff Summer School "Inverse problems for multi-scale models". Bonn, Germany.
   Destinataria de una beca de viaje.

6. **"Agent-based models for biomedical applications: An introduction to PhysiCell"**.
   XI MEETING OF THE SPANISH CHAPTER OF THE EUROPEAN SOCIETY OF BIOMECHANICS. Zaragoza, Spain.
   Presentadora de un taller práctico.

Participación con póster

1. **"Collagen Density Regulates Tumour Spheroid Growth Through Cell Motility: A Computational Study"**.
   II Jornada del I3A y X Jornada de Jóvenes Investigadores. Zaragoza, Spain.

Participaciones

1. *PDBEB Courses 2021/2022 - Computational Biology*. Coimbra, Portugal (online).
   Presentación de un proyecto corto llamado "Using text-mining techniques in literature reviews".

2. *2021 Network Modeling Virtual Summer School & Hackathon.* USA (online).

3. *2021 PhysiCell Workshop and Hackathon.* USA (online).
   Parte del equipo finalista del premio "Mejor Herramienta" en el Hackathon.

4. *Short Course on Particle-Based Methods in Engineering and Applied Science.* Hamburg, Germany (online)

A.4.3 Dirección de trabajos

Trabajos Fin de Master (TFM)

1. "*In silico* modeling of single-cell fibroblast migration in response to 3D chemical gradients."
   Ijeoma Charity Asilebo. 2020/2021.

2. "Computational simulation of the T cell migration process using a discrete model."
   Imán Laga Boul-Atarass. 2021/2022.

3. "Modelling of tumoral cells metabolism in tumoral spheroids formation and development."
   Alejandro Modrego Bravo. 2022/2023

Trabajos Fin de Grado (TFG)

1. "Simulation by methods based on agents of immunotherapy techniques against cancer." Guillermo Villamor Aparicio. 2022/2023

A.4.4 Software libre

1. **"*PhysiCell-ECM*:** A PhysiCell extension to account for the extracellular matrix". Available on GitHub[1] (MIT License).

2. **"*pdgf-induced-motility*:** Computational model of chemotaxis calibrated with Bayesian Optimization." Available on GitHub[2] (BSD-3-Clause License).

3. **"*neurorosettes*:** An agent-based framework to model the formation of rosette patterns in tissues of the nervous system." Available on Zenodo (BSD-3-Clause License) (Gonçalves and García-Aznar, 2022).

4. **"*PhysiCOOL*:** A generalized framework for model Calibration and Optimization Of modeLing projects". Available on Zenodo (BSD-3-Clause License) (Gonçalves et al., 2023).

A.4.5 Colaboraciones

PhysiCell Workshop 2021

Durante sus estudios de doctorado, en 2021, la candidata a doctorada participó en un taller/*hackathon* en línea, organizado por el equipo responsable del desarrollo de *PhysiCell* (Paul Macklin's Math Cancer Lab, Universidad de Indiana). Durante este *hackathon*, la candidata se unió a un equipo para implementar *PhysiCOOL*, una herramienta diseñada para calibrar y optimizar los modelos de *PhysiCell*. El equipo estuvo compuesto por la doctoranda Inês G. Gonçalves, David Hormuth II y Caleb M. Phillips (miembros del grupo del Centro de Oncología Computacional del Instituto Oden de Texas) y Sandhya Prabhakaran (miembro del departamento de Oncología Matemática Integrada del H.Lee Moffitt Cancer Center and Research Institute en Florida) y fueron galardonados con el segundo premio de "Mejor herramienta" al final del *hackathon*. Desde entonces, el equipo ha colaborado para mejorar esta herramienta y convertirla en un estándar en el ecosistema de *PhysiCell*, a través del desarrollo de nuevas funciones, la creación de documentación y ejemplos interactivos, bien como la publicación y distribución del software como un recurso libre. Además, este trabajo ha sido aceptado para su publicación y se puede encontrar en bioRxiv (Gonçalves et al., 2022).

[1] https://github.com/m2be-igg/PhysiCell-ECM
[2] https://github.com/m2be-igg/pdgf-induced-motility

MAtrix - KU Leuven

Además, durante el primer año de su doctorado, la candidata colaboró con el grupo de investigación donde completó su TFM, supervisado por Diego A. Vargas y Hans Van Oosterwyck (miembros del grupo de Mecanobiología e Ingeniería de Tejidos en KU Leuven), para publicar este trabajo. En particular, la candidata estudió el papel de los mecanismos de reconocimiento mecánico del entorno celular en la motilidad de una célula. Además, ha estudiado cómo la rigidez del sustrato y las adherencias célula-sustrato influyeron en el fenotipo de migración de la célula. Por lo tanto, el conocimiento adquirido en este proyecto fue de particular relevancia para esta tesis doctoral y fue aplicado para modelar y evaluar patrones de migración en la motilidad celular como respuesta a las propiedades del entorno circundante. Esta publicación dio lugar a la publicación de un artículo original:

1. Vargas, D. A., Gonçalves, I. G., Heck, T., Smeets, B., Lafuente-Gracia, L., Ramon, H., & Van Oosterwyck, H. **"Modeling of mechanosensing mechanisms reveals distinct cell migration modes to emerge from combinations of substrate stiffness and adhesion receptor–ligand affinity."** Frontiers in Bioengineering and Biotechnology 8 (2020): 459.

LEPABE/BioSIM - Facultad de Ingeniería de la Universidad de Porto

Por último, durante los dos primeros años de sus estudios de doctorado, la doctoranda estableció una colaboración para continuar y finalizar un proyecto de investigación anterior desarrollado bajo la dirección de Manuel Simões (miembro del grupo de investigación LEPABE de la Facultad de Ingeniería de la Universidad de Porto). Esta colaboración no estuvo directamente vinculada a los estudios de doctorado de la candidata y se desarrolló en paralelo a la tesis doctoral, pero fue relevante para la trayectoria académica de la candidata. En este trabajo, la candidata a doctorada tenía como objetivo construir una base de datos en línea con datos de informes revisados por pares de brotes de *Legionella* para identificar los factores ambientales que pueden haber promovido estos brotes. Para lograr esto, la estudiante reunió y analizó datos publicados sobre casos de *Legionella* y trabajó junto con miembros del grupo de investigación Biomolecular SIMulations (BioSIM) para implementar y publicar la base de datos en línea, *LegionellaDB*[3]. Esta colaboración también resultó en la publicación de un artículo original y un documento de revisión, que se enumeran a continuación:

1. Gonçalves, I. G., Fernandes, H. S., Melo, A., Sousa, S. F., Simões, L. C., & Simões, M. **"LegionellaDB–A database on Legionella outbreaks."** Trends in Microbiology 29.10 (2021): 863-866.

2. Gonçalves, I. G., Simões, L. C., & Simões, M. **"Legionella pneumophila."** In Trends in Microbiology 29.9 (2021): 860–861.

[3]https://legionelladb.biosim.pt/

A.4.6 Agradecimientos

Esta tesis doctoral es el resultado de proyectos que han recibido financiación del Consejo Europeo de Investigación (ERC), en el marco del programa Advanced Grant, y del programa de investigación e innovación Horizonte 2020 de la Unión Europea: proyecto CADENCE (ERC-2016-ADG-grant accord no. 742684), proyecto PRIMAGE (SC1-DTH-07-2018-acuerdo de subvención n.º 826494), y proyecto ICoMICS (ERC-2020-ADG-acuerdo de subvención n.º 101018587).

APPENDIX B

Supplementary data

Contents

B.1 Single cell migration 132

B.1.1 Study 1: Model identification for the locomotive forces generator function 132

B.1.2 Study 1: Sensitivity analysis 132

B.1.3 Study 2: Relevance of the administration methods of chemotactic substances . . 134

B.1.4 Study 2: Direct comparison between 2.5 mg/mL and 4.0 mg/mL matrices 136

B.1.5 Study 2: Spatial organization and intercellular distance 136

B.1.6 Study 2: Model predictions evaluation of the estimated cell persistence time . . . 138

B.2 Tumour growth 139

B.2.1 Automatic cluster classification and quantification 139

B.2.2 Oxygen diffusion and consumption dynamics 143

B.3 Pattern formation 144

B.3.1 Rosette formation at distinct adhesion and differentiation levels 144

This appendix is based on the supplementary files of:

Inês Godinho Gonçalves and Jose Manuel Garcia-Aznar. **"Extracellular matrix density regulates the formation of tumour spheroids through cell migration."** PLoS computational biology 17.2 (2021): e1008764.

Inês Godinho Gonçalves, Nieves Movilla, Carlos Borau and Jose Manuel Garcia-Aznar. **"A novel integrated experimental and computational approach to unravel fibroblast motility in response to chemical gradients in 3D environments."** Integrative Biology (2023)

nês Godinho Gonçalves and Jose Manuel Garcia-Aznar. **"*neurorosettes:* A novel computational modelling framework to explore the formation of Homer-Wright rosettes in neuroblastoma."** Computational Particle Mechanics (under review)

B.1 Single cell migration

B.1.1 Study 1: Model identification for the locomotive forces generator function

We identified that the model parameters that play a significant role in our obtained results are related to the cell-generated locomotive forces. Taking this into consideration, we will mainly focus on how the function we have used to generate these cell-generated locomotive forces affects the model outputs. Since our work regarded the qualitative trends observed experimentally, we have mostly focused on studying how different forces distributions influence our results, focusing on the shape of the distributions. Based on the assumption that the main factors acting on cell velocity (in the single-cell setup) are the cell-generated locomotive forces, as well as the drag forces determined by the ECM, and to avoid the computational cost of the model, we have built a simplified script to study the effect of these forces. In particular, we built a Python script that implements the equation below, taking into account values from a given force distribution and the effect of the ECM, through the dynamic viscosity of the matrix.

$$\mathbf{v}_i \approx \frac{1}{\mu}\mathbf{F}^i_{loc} \tag{B.1}$$

Here, $\mathbf{F}^i_{loc}$ is a value chosen at random from the chosen distribution and μ is the dynamic viscosity of the collagen matrix. For this sensitivity study, we have considered a normal distribution with a mean value, μ, of 0.15 and a standard deviation, σ, of 0.2; a normal distribution with μ = 0.0, σ = 0.45; and a lognormal distribution with μ = 0.0, σ = 0.7. The parameters of each function were fitted to provide the best possible results, and we compiled a number of velocity values comparable to that measured experimentally. The results obtained for this study are summarized in Fig , from which we conclude that the shape of the chosen distribution is highly significant and can lead to different cell behaviours.

Extending this analysis, we have also studied how the coefficients of our force generator function may influence our results. For our implementation, which was based on the empirical velocity distributions and fitted accordingly, the velocity values depend on the coefficients of a function given by a general third-degree polynomial form

$$y(x) = ax^3 + bx^2 + cx + d \tag{B.2}$$

which we fitted to become

$$F_{loc}(x) = 1.56x^3 + 3.27x^2 + 0.07x + 0.06 \tag{B.3}$$

B.1.2 Study 1: Sensitivity analysis

Particularly, we studied how these coefficients influence the mean, median and maximum cell velocity values. We have chosen to use the cells grown in medium-density matrices as an example since the viscosity of the matrix affects all these parameters equally, and, thus, we expect the

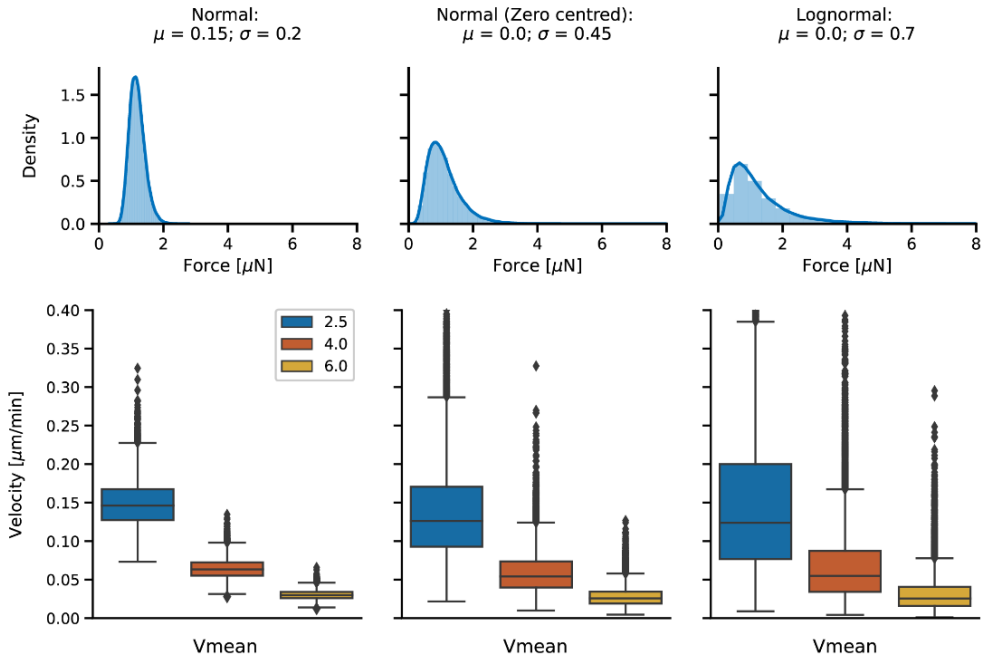

Figure B.1: **Estimated results for cell velocities in function of different force generator functions.** (A) Estimated velocity distributions for a normal distribution, a normal distribution centred at zero and a lognormal distribution (top). To study how the force generator influences the computational results, we have studied multiple force distribution functions and their effect on cell velocity. It can be concluded that, despite capturing the median velocities, these distributions fail to replicate the range of the experimental values, as well as the lack of outliers. Hence, we conclude that the chosen distribution highly impacts the obtained results.

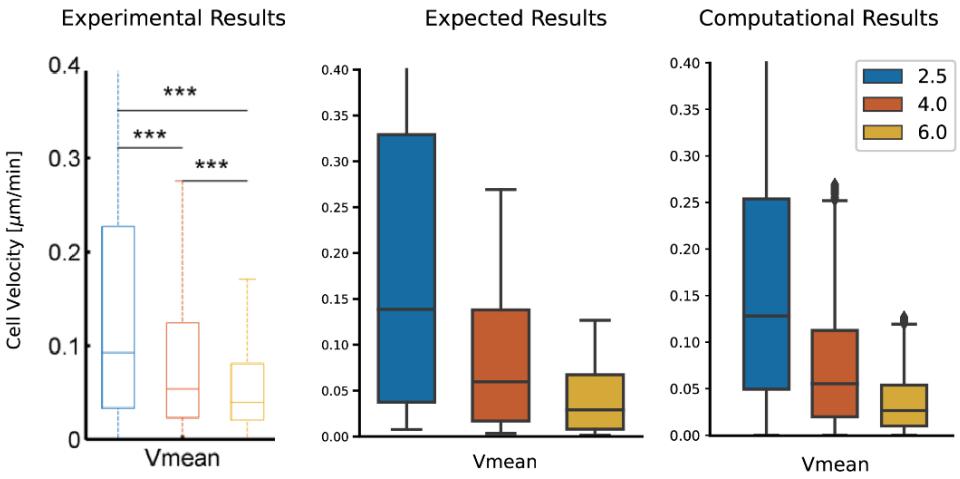

Figure B.2: **Comparison between a simplified model of cell motility based on experimental data and *PhysiCell* results.** Experimental results (left), the simplified model (centre) and the computational results obtained with the actual model (right). Although the simplified model predicted a broader distribution for cells grown in low-density matrices, we consider that the conclusions obtained with this implementation apply to our proposed model.

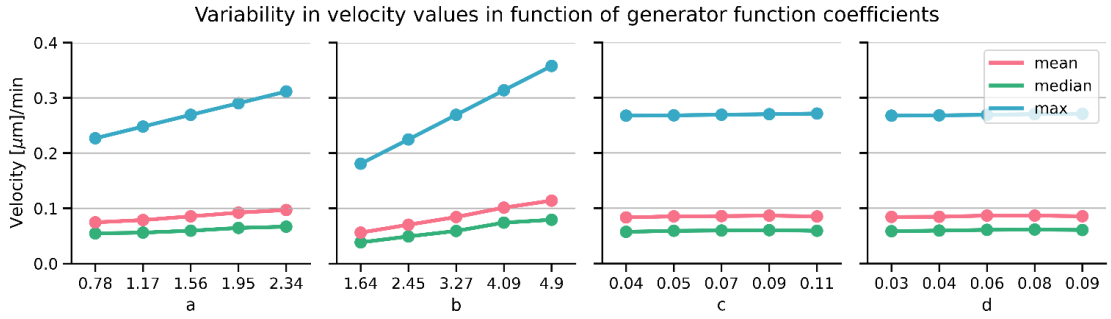

Figure B.3: **Effect of the generator function coefficients on the cell velocity results.** Mean (red), median (green) and maximum (blue) values for the instantaneous cell velocities of cells grown in medium-density matrices, for different parameter values. For each plot, a single parameter was changed, while the others were kept at the values presented in the paper (a = 1.56, b = 3.27, c = 0.07, d = 0.06). The coefficients were varied based on their magnitude. For small changes, especially those regarding coefficients c and d, the velocity values do not appear to be largely affected. However, for more significant increases both in a and b, the velocities values increase with these parameters, in particular the maximum velocity value.

changes to be comparable between matrix densities. The results of this study are summarized in Fig B.3. Based on this figure, we conclude that our model is robust to small variations in these coefficients.

B.1.3 Study 2: Relevance of the administration methods of chemotactic substances

We performed some additional assays to assess whether (i) increasing the concentration of PDGF-BB or (ii) discontinuing its administration would modify the fibroblasts' chemotactic response. We performed these studies with the experimental setup described as Condition 1 in our manuscript. That is, we used a microfluidic device with a single chamber, surrounded by two media channels. Subsequently, we created a chemotactic gradient through the introduction of PDGF-BB in the channel opposite to the cell monolayer. In the first study, we considered a PDGF-BB concentration of 50.0 ng/mL and compared the results to those obtained for a concentration of 5.0 ng/mL. These results are presented in Table B.1.

Overall, we observed that the changes in the chemotactic response were almost negligible and we could not identify any trend originated by the concentration change. We believe that these

Table B.1: **Relative increase of the mean distance travelled by fibroblasts in response to different concentrations of PDGF-BB.** Results were obtained by comparing the mean travelled distance in the presence of PDGF-BB to those obtained under control conditions.

Time (hours)	5 ng/mL	50 ng/mL
48	1.98	1.89
96	1.83	1.87

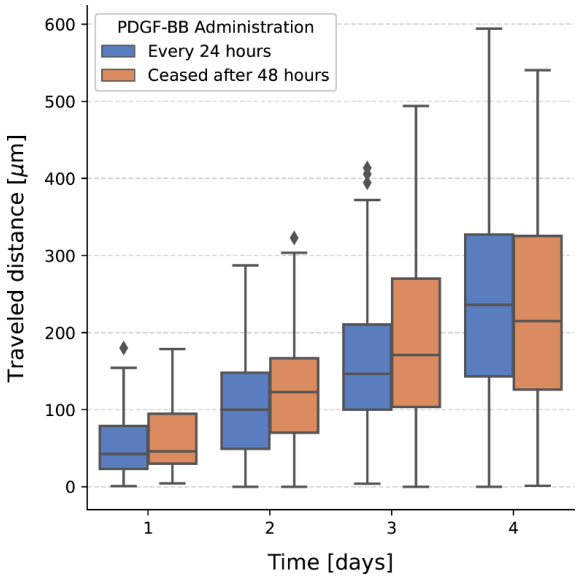

Figure B.4: **Displacement values for migrating fibroblasts over 4 days when PDGF-BB administration was ceased after 48 hours of experiments.** Distances were computed based on the perpendicular distance between the current position of each cell and the position of the monolayer. Results were taken at 24, 48, 72 and 96 hours.

findings are in accordance with previous studies that have shown that cell velocity is proportional to the magnitude of a chemotactic gradient but only up to a certain concentration value, at which cell velocity values saturate (Chertock et al., 2012). Henceforth, it is possible that fibroblasts are already experiencing this saturation limit at a concentration of 5.0 ng/mL.

In regard the second study, which aimed to identify changes in the chemotactic behaviour when PDGF-BB stopped being administrated, we used the same experimental setup and stopped adding PDGF-BB to the cell media after 48 hours. These results are shown in Fig R1, shown below. Fundamentally, we observed that, even though we removed PDGF-BB after 48 hours, the fibroblasts travelled as much as those that had access to PDGF-BB during 96 hours. We believe that this may have occurred due to the binding of chemical factor to the collagen matrix, which has been shown to occur in previous studies (Moreno-Arotzena et al., 2014). Consequently, fibroblast motility might have been a result of haptotaxis even after PDGF-BB was no longer administrated and the chemical gradient was reduced.

In future experimental studies, the relevance of these haptotactic cues could be further assessed. We propose that the computational model parameters found to best fit this experimental setup could be used to simulate the experiments performed under control conditions, and, in addition, a model extension could be introduced to modify the parameters responsible for simulating the cells' response to migration-inducing factors, namely the forward bias. We believe that it would be expected that forward bias value would decrease, as it represents the cell's ability to follow a given direction and the chemotactic gradient would no longer be present. The factor by which the value should decrease could be manually varied and, subsequently, the value that best fitted the experimental results performed with PDGF-BB would be chosen. This would be particu-

larly helpful to further support the hypothesis that haptotactic cues were important (we believe that a slight decrease would signify a less powerful signal), or if cells adopted a more random pattern, similar to the cells seeded under control conditions (i.e., a large decrease would signify that there were no relevant migration-inducing signals).

Furthermore, at the experimental level, a new experimental configuration could be implemented where, instead of stopping the administration of PDGF-BB after 48 hours, the placement of the PDGF-BB source would be changed instead. In other words, a new experimental design could be considered in which PDGF-BB would be introduced in the opposite channel until 48 hours and, at this timepoint, stop the administration in this channel and, instead, introduce PDGF-BB in the monolayer channel. Henceforth, it could be evaluated whether fibroblasts showed a preferential migration response to the initial chemotactic gradient and the haptotactic cues generated by it, or if the new stimulus introduced by the new chemotactic signal in the opposite direction would lead to cell arrest or even a change in direction. These results could also be integrated into the computational approach to further validate the model.

B.1.4 Study 2: Direct comparison between 2.5 mg/mL and 4.0 mg/mL matrices

As presented in the main text of this work, our experimental results show that the final displacement values for fibroblasts seeded in collagen matrices of distinct collagen densities were relatively similar. Taking into consideration the effect of the ECM and its mechanical properties on cell motility, and how an increased matrix density may impair cell movement due to reduced pore sizes, we include a direct analysis between these datasets to better understand why the displacement values were not largely affected by the differences in the matrices' structure. These results are shown in Fig B.5. It is apparent that, for both fibroblasts seeded under control conditions and those seeded in the presence of PDGF-BB, the displacement values start by being different during the first two days between the two collagen densities. Yet, this difference is reduced over time.

B.1.5 Study 2: Spatial organization and intercellular distance

In addition to the distance travelled by cells from the position of the initial fibroblast monolayer, we have computed the intercellular distance for the single-channel studies. We considered the average distance between each fibroblast and its five nearest neighbours. The obtained results are shown in Fig B.6. Overall, it is apparent that, although cells migrated individually, some common patterns were shared between cells, even under different experimental conditions. In particular, we concluded that the intercellular distances in the control and MC experiments tended to be similar. However, we observed that these values were slightly smaller for the OC experiments. Furthermore, we obtained larger distance values in the experiments with higher collagen density.

Similarly to what was seen for less dense matrices, the presence of PDGF-BB in the channel opposite to the monolayer promoted a slight decrease in the intercellular distance values. Comparing the experimental and computational data, we conclude that the experimental results differed less between conditions than the computational results. Moreover, the computational results

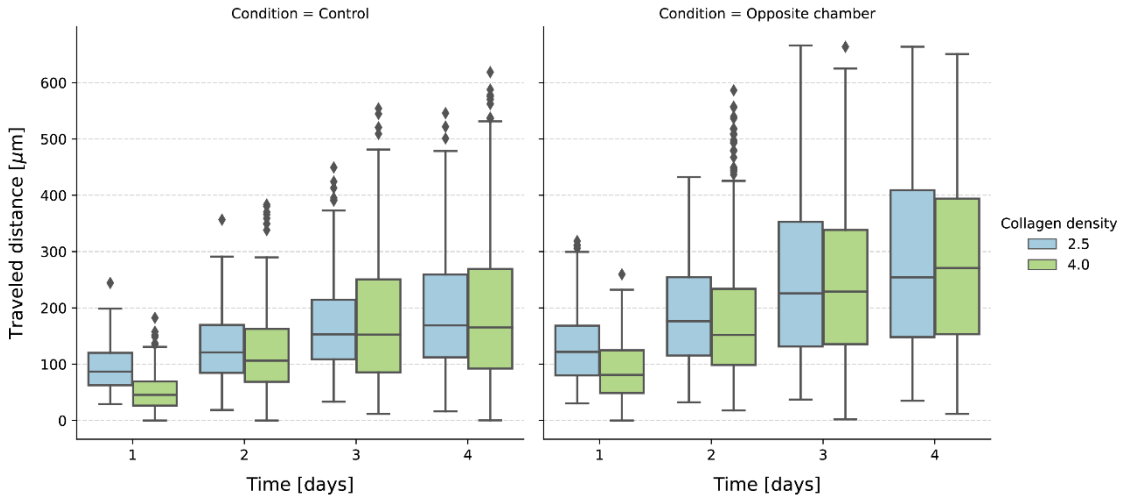

Figure B.5: **Displacement values measured for migrating fibroblasts over 4 days in 2.5 mg/mL and 4.0 mg/mL collagen matrices.** Distances were computed based on the perpendicular distance between the current position of each cell and the position of the monolayer. Results were taken at 24, 48, 72 and 96 hours. Results are statistically different ($p < 0.001$) between the two matrix densities for the first day of experiments.

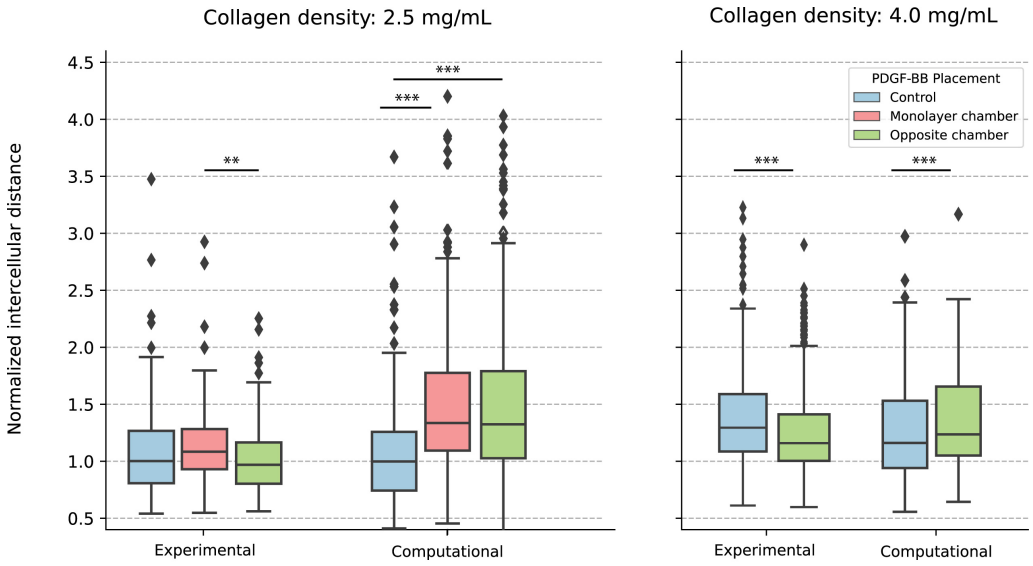

Figure B.6: **Experimental and computational results for the intercellular distances between fibroblasts seeded in single-chamber devices.** Normalized intercellular distance values for the experiments and simulations performed with single-chamber devices. Distances were computed as the average distance between each cell and a number of neighbours, $n = 5$, at day 4 of experiments. Values were normalized using the median distance value obtained for the corresponding control study as the reference value. Differences between the experimental results obtained for 2.5 mg/mL matrices (represented on the left) are statistically significant between the cell channel condition and opposite channel condition ($p < 0.01$). Differences between the experimental results obtained for 4.0 mg/mL matrices are also statistically significant ($p < 0.001$).

suggest that the enhanced motility of fibroblasts seeded in the presence of PDGF-BB produced increased intercellular distance values. However, given that these results differed from those obtained in vitro, we suggest that proliferation may have been promoted by the presence of PDGF-BB in the experiments, which resulted in new cells appearing in the intercellular space (47). Subsequently, the intercellular distance values decreased due to the new cells. Since we did not include the effect of PDGF-BB on cell proliferation in our computational model, as this was beyond the scope of our work and we focused primarily on modelling single-cell motility, this behaviour was not captured in the computational results.

B.1.6 Study 2: Model predictions evaluation of the estimated cell persistence time

As an example of how our model predictions can be used to plan future validation experiments to confirm them, we quantified the persistence time of fibroblasts' migration under different conditions by tracking the cell trajectories at a shorter time scale. To do so, we used the experimental configuration described as Condition 1 in our manuscript and, at 96 hours of settling time, we tracked the cells' positions every 10 minutes for a period of 4 hours. To obtain the numerical values, we fitted the Mean Squared Displacement (MSD) of the individual cell trajectories to an Anisotropic Persistent Random Walk model as described by Wu et al. (2015) .The MSD of individual trajectories at different time lags was computed from the following equation:

$$MSD(\tau) = (x(t+\tau) - x(t))^2 + (y(t+\tau) - y(t))^2 \tag{B.4}$$

where $\tau = n\Delta t$ and $n = 1, 2, ..., N_{frames-1}$. Δt is the time step used during image acquisition, and then fitted to the following model:

$$MSD(\tau) = 2S^2P(\tau - P(1 - e^{(} - \tau/P))) + 4\sigma^2 \tag{B.5}$$

where S is the cell speed, P is its persistence time, and $4\sigma^2$ is the noise (error) in the position of the cell. The results for this study are show in Fig B.7.

Note that these assays were posterior to the model calculations and that the involved parameter values were nor directed nor limited in any way during the optimization process. Thus, even though the experimentally measured persistence time median values were slightly smaller than those obtained with the computational model, we verified that they are on the same order of magnitude, which is relevant considering that the experimental data used to calibrate the model was obtained on a scale of days (every 24h), and yet we were able to estimate proper values through our optimization pipeline. Furthermore, we identified that there was an increase in the persistence times both in the presence of PDGF-BB and in denser collagen matrices, again in accordance with our model predictions.

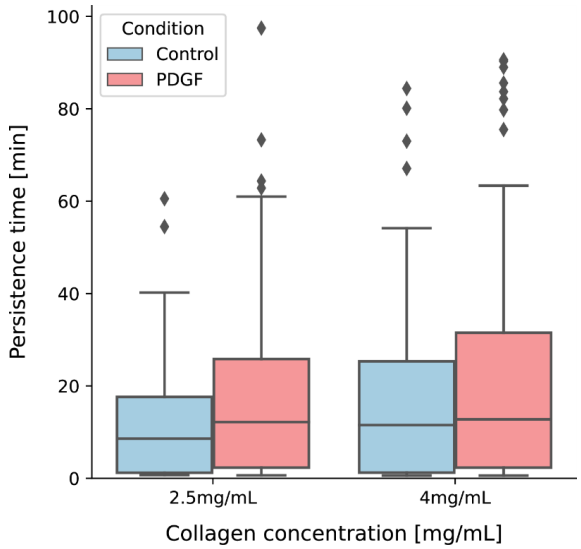

Figure B.7: **Distribution of the fibroblast persistence times quantified experimentally at 96 hours of experiments.** Results were obtained for both 2.5 and 4.0 mg/mL collagen matrices and in the presence of PDGF-BB, which was placed in the channel opposite to the initial fibroblast monolayer. The quantified values were in the order of minutes, as it had been predicted by our computational model. Furthermore, it is apparent that both the addition of PDGF-BB and the use of denser collagen matrices enhanced fibroblast persistence, which had also been predicted by the model.

B.2 Tumour growth

B.2.1 Automatic cluster classification and quantification

Apart from the model parameters, some aspects of the data processing methodology used may influence our results, particularly for the multicellular setup. In order to quantify the area and eccentricity of the multicellular clusters observed after some days of growth, cells must be classified into clusters. Although we have tried to replicate the type of processing used in the experimental results, this was not completely possible, as there are some differences the experimental and image-based data and our computational results, that are based on the coordinates of the centres of the cells. Henceforth, some parameters had to be chosen and fitted, namely the radius and the minimum number of cells considered by the clustering algorithm, as well as the height of interest that we have defined to replicate the effect of an image-based analysis.

We have chosen to classify cells into clusters through the implementation of the DBSCAN (Ram et al., 2010; Pedregosa et al., 2012), which requires the user to define the radius of interaction and the minimum number of cells in a cluster. For these parameters, we have aimed to choose a radius of interaction that was slightly larger than the radius of two cells, so that only cells that were close to each other were selected. Moreover, we initially defined that the minimum of clusters of cells in the radius of interaction should be 3. However, given that the experimental results suggest that there are clusters at day 1, and cells are only able to replicate once in that period, we have defined that, at day 1, clusters could be composed by a minimum number of cells of 2. We

have only used this value for day 1, though, as we have observed that this also promoted the classification of single cells into clusters. More information on the effect of these parameters can be found in Fig B.8.

We quantified cluster area by assuming that multicellular clusters had a spherical geometry. Accordingly, we found the approximate radius of the multicellular cluster and computed its area as a circle with that radius. However, other options are available to obtain this metric, assuming other geometries. For example, an ellipsoid shape could be considered to account for cluster elongation along one of its axis. In addition, the convex hull could be evaluated to provide a more accurate area description considering the clusters' specific geometry. However, it must be taken into account that this detailed representation is based on the cells' centre positions and imaging data does not have this resolution level, which can lead to differences between computational and experimental data. Examples of alternative cluster area classification metrics and its corresponding values can be seen in Fig B.9.

Cluster eccentricity

Regarding the eccentricity values, we have used the same parameter values for our cluster area study, as we have observed that changes in these parameters would not lead to significant differences in eccentricity values, as presented in Fig B.8. We would also like to comment on the fact that cluster eccentricity values are highly sensitive to small variations in cluster dimensions. Accordingly, although the computational results are robust to small perturbations in parameter values, there are some differences between the experimental and computational results, as can be seen in Fig . In particular, we have observed that it was very difficult to obtain eccentricity values as low as those seen experimentally.

Cluster eccentricity is computed based on the dimensions of the ellipse that best fits the cluster. In particular, taking an ellipse's major, a, and minor, b, axis, we compute its eccentricity using the following equation

$$eccentricity = \frac{\sqrt{(a^2 - b^2)}}{a^2} \tag{B.6}$$

Therefore, we can plot cluster eccentricity in function of the ratio between a and b, as presented in Fig B.11. This plot allows us to confirm that cluster eccentricity is particularly sensitive at low eccentricity values. In this range (between 0-0.3, approximately), very small differences between the length of the ellipse axes greatly influence the eccentricity values. We note that our computational data is very sensitive to these differences, which are in the order of <10 μm, as we can keep track of the coordinates of each cell. However, the experimental results, which are image-based, present a lower sensitivity, as they are limited by the size of each pixel. Henceforth, experimental results present lower eccentricity values.

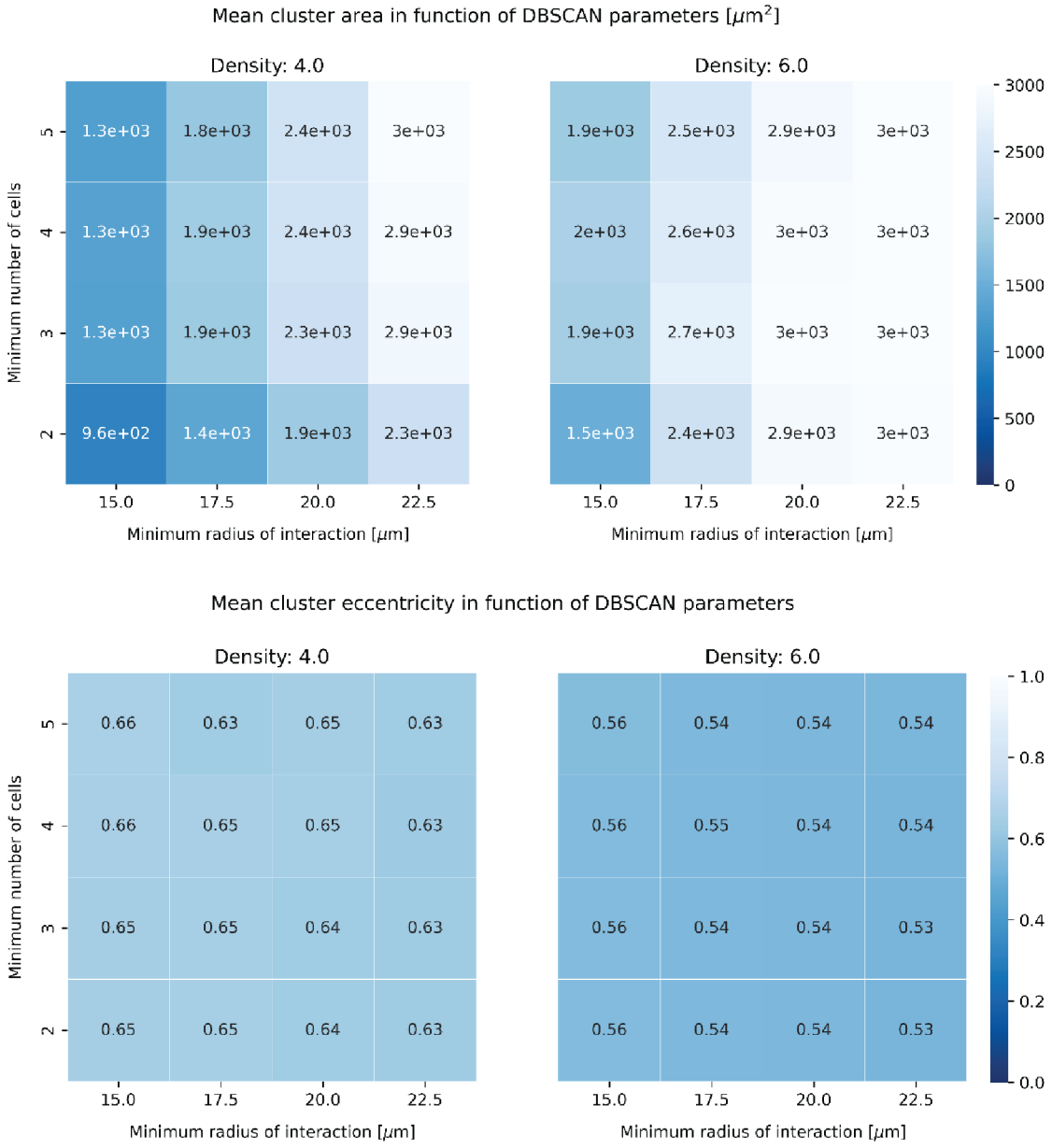

Figure B.8: **Effect of the DBSCAN parameters on cluster metrics.** Mean cluster area for different values of the radius of interaction and the minimum number of cells in said radius, at day 5. On the one hand, the effect of different values of the minimum number of cells is mainly noticeable between 2 and 3. A smaller minimum number of cells leads to the detection of small aggregates, that do not truly classify as clusters, reducing the value of the mean cluster area. On the other hand, the value of the radius of interaction appears to have a more significant effect, particularly for clusters grown in medium-density matrices. This is probably explained by having more cells scattered through the domain, which may be detected through an increase in the radius magnitude. Contrarily, high-density matrices produce large clusters that are fairly distanced from each other and do not present individual cells. Changes in the DBSCAN parameters do not produce significant changes in the computed eccentricity values. Therefore, we opted to keep the same values as those used to compute cluster area, to keep our methods consistent.

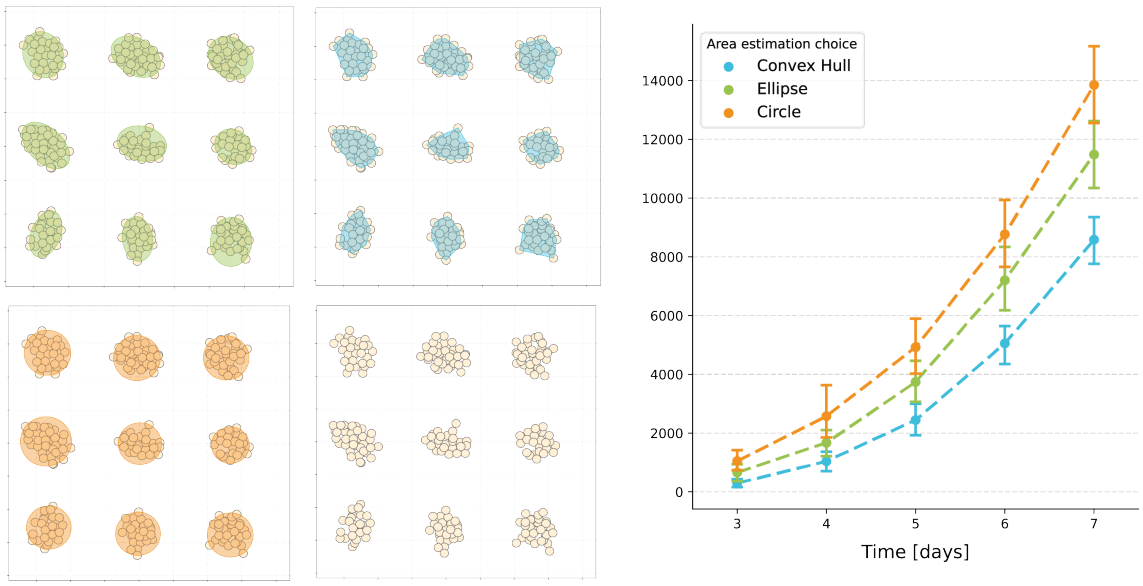

Figure B.9: **Cluster area quantification approaches.** Examples of alternate area quantification approaches, such as the smallest enclosing circle (orange), smallest enclosing ellipse (green) and convex hull (blue) and their evolution over time. The convex hull provides the most accurate area depiction based on the cluster specific geometry, yet its values may vary when compared with imaging data with lower resolution. In said cases, the smallest enclosing ellipse or circle may prove to be better options.

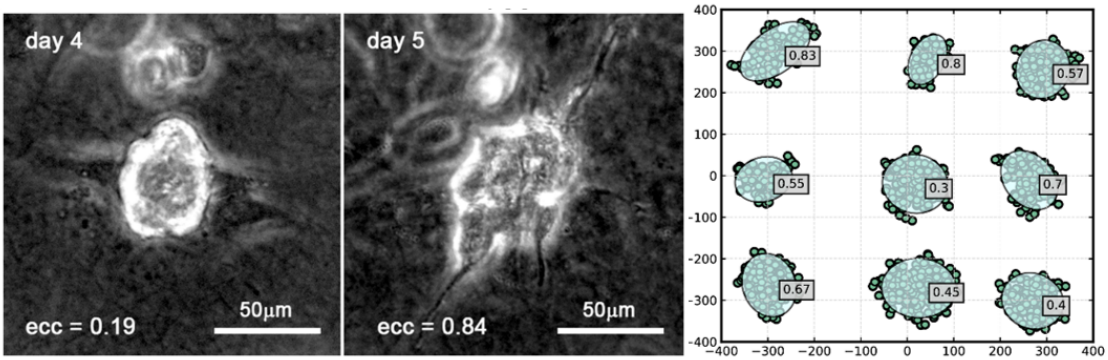

Figure B.10: **Differences between the experimental and computational values for cluster eccentricity.** Experimental (left) and computational (right) results for cluster eccentricity (the computational results also present the equivalent ellipse from which these values have been calculated). It must be noted that an eccentricity of zero indicates a round cluster, whereas an eccentricity of one indicates a cluster that resembles a line. The experimental and computational results present differences in cluster eccentricity, although a visual analysis may suggest comparable eccentricity values. For instance, the centre cluster in the computational dataset presents an eccentricity of 0.3. Compared to the experimental results for day 4, it seems to be as round, if not more, than the experimental cluster. Yet, the eccentricity value of the latter is of 0.19. Similarly, we observed this pattern for several of the computational clusters.

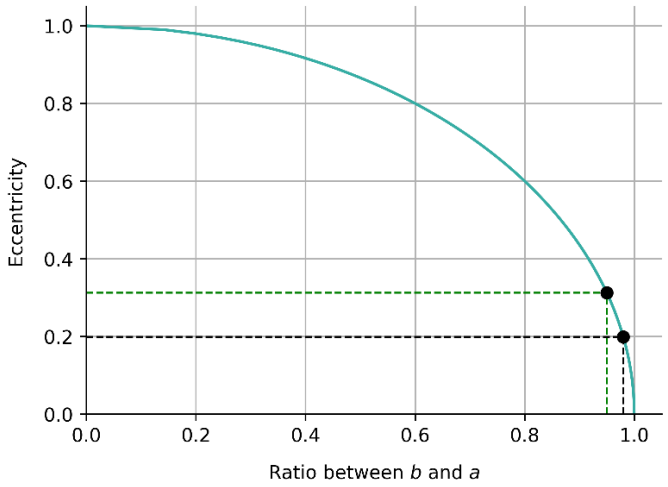

Figure B.11: **Effect of the ratio between the major and minor axes of an ellipse on cluster eccentricity.** The relationship between ellipse eccentricity and the ratio between its major and minor axes dictates that, for low eccentricity values, the outcome is highly sensitive to differences in the dimensions of the ellipse. For a cluster with a major axis of 100 µm, a difference of just 5% between a and b (which, in this case, is just around 5 µm) produces an eccentricity of around 0.3 (green lines). Furthermore, a difference of just 2 µm between the two axes of this cluster produces an eccentricity value of 0.2 (black lines). Contrarily, as the eccentricity values increase, the output becomes less sensitive to these differences.

B.2.2 Oxygen diffusion and consumption dynamics

Although we have included oxygen and the effects that this substance may have on cells into our model, we do not consider it to play a significant role in our simulations. Considering the presence of oxygen and the cells' dependence on oxygen allows for a more accurate representation of cellular systems and can be useful to study other experimental configurations, to which our model can easily be adapted. Nevertheless, the experimental settings in which we base our simulations for this particular study are not expected to promote hypoxia, nor do we believe that changes in oxygen levels affected cell migration. Consequently, we do not consider oxygen to have a relevant effect on cell cycle or cell motility, either.

Here we aim to present a study conducted with the simulated data for our model, to show that oxygen is present but does not play a significant role in our simulations. Hypoxia levels are defined in the model as oxygen pressure values below 15 mmHg, and cell necrosis starts occurring when pressure levels reach values lower than 5 mmHg. Furthermore, we set an initial oxygen concentration of 38 mmHg, which we expect to be lower than that found in experimental conditions. To evaluate oxygen dynamics, we considered only the simulations that refer to multicellular cluster growth, as the oxygen consumption by a single cell in the individual cell motility study is not significant. We selected the middle plane (z=0) of the domain, as we expect this to be the plane with the highest cell density and, consequently, where oxygen consumption is maximal. Subsequently, we plotted a heatmap representing the oxygen concentration at each voxel for each of the studied collagen densities after seven days of growth, as seen in Fig B.12. From this data, it is possible to

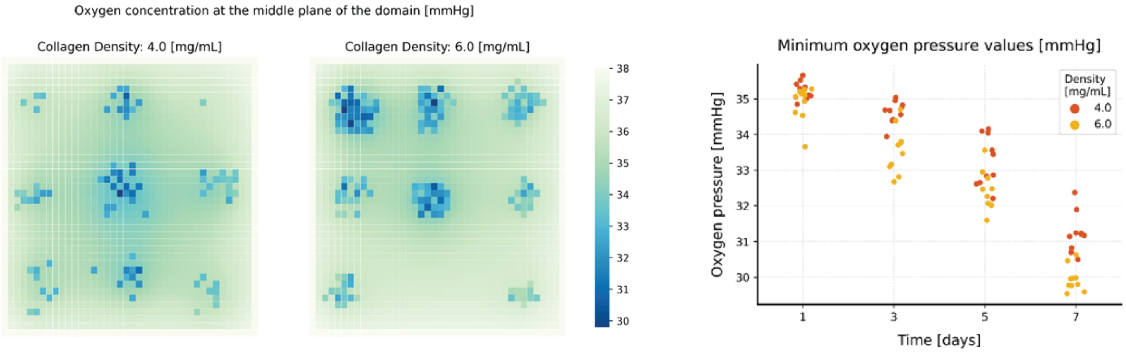

Figure B.12: **Oxygen levels over the simulation timespan.** Oxygen levels decrease over time and in response to differences in the matrix properties. Specifically, matrices of high density present lower oxygen levels, which can be attributed to the cells' restrained motility. As cells adopt a packed distribution, oxygen diffusion is hindered at the centre of the cell cluster. Contrarily, in lower density matrices, cells can migrate through the matrix and oxygen consumption does not overpower its diffusion. We also present a quantitative analysis of the minimum concentration values at the middle plane of the simulated domain. Our results show that the minimum values decrease through time but never reach pressure levels lower than 29 mmHg.

conclude that oxygen levels decrease over time, but never reach values lower than 30 mmHg.

Finally, for an extended quantitative analysis, Fig B.12 showcases how the minimum oxygen concentrations at the middle plane change over time, for all 10 replicates. From this, we can conclude that the results are coherent throughout all of the replicates and that minimum values never adopt values lower than 29 mmHg, which is higher than the hypoxic threshold that activates cell necrosis and hindered cell proliferation. In conclusion, we infer that oxygen does not play a role in our simulations, due to our initial conditions, but its effect could be studied through our model, for different experimental configurations.

B.3 Pattern formation

B.3.1 Rosette formation at distinct adhesion and differentiation levels

In the numerical study performed to evaluate how cell-cell adhesion and neuronal differentiation, three replicates were considered, even though only one was shown in the main text. Fig B.13 shows the results for all the replicates evaluated, showing that the results were consistent in all the simulations.

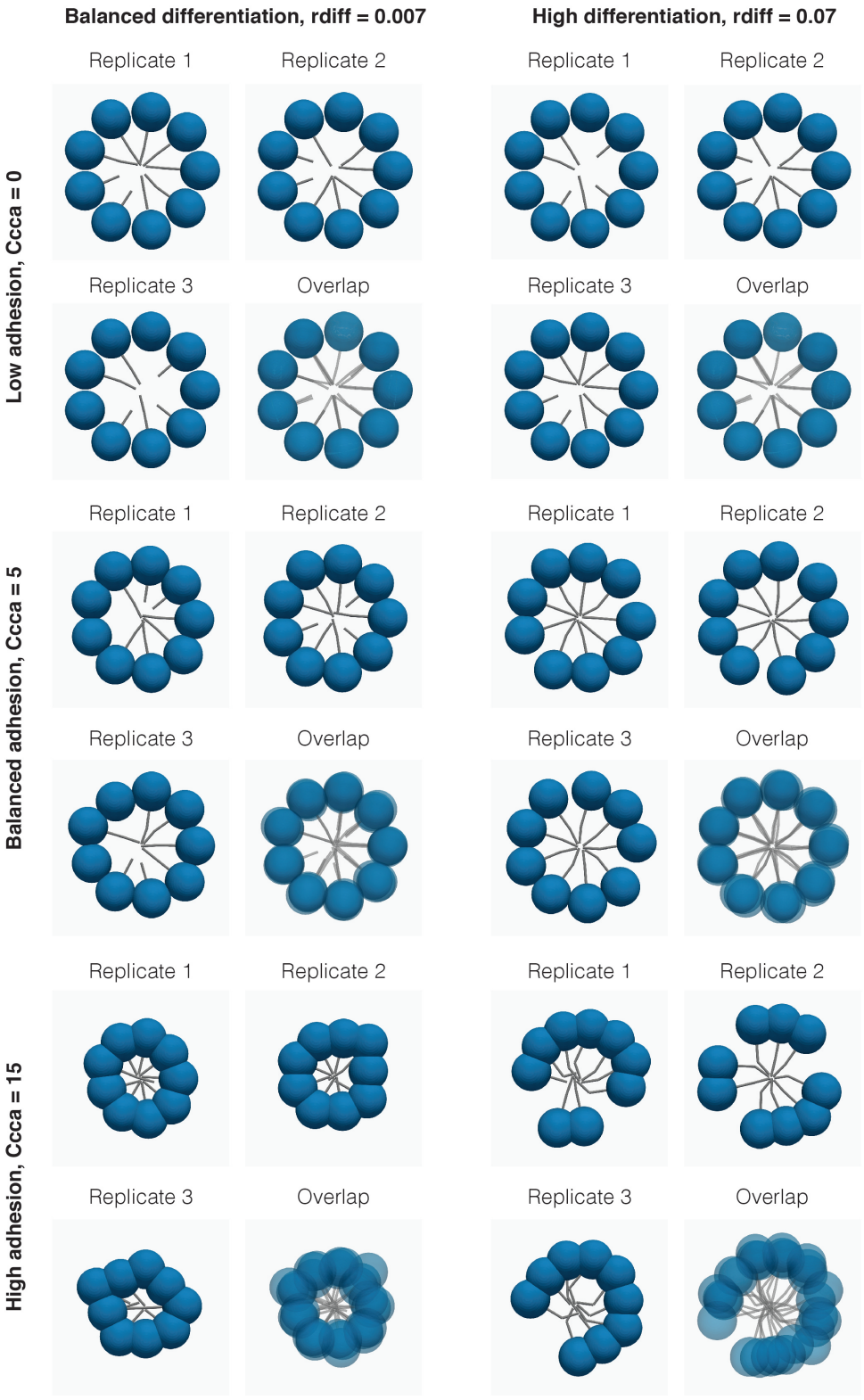

Figure B.13: **Computational results of HW rosette formation at different differentiation rates and cell-cell adhesion levels.** Results are shown for three replicates and the overlap of all replicates is also presented, which shows that results are consistent over multiple replicates.

APPENDIX C

PhysiCOOL: An open-source optimization Python library for PhysiCell

Contents

C.1 Introduction 148
C.2 Statement of Need 148
C.3 Implementation 149
C.3.1 Configuration file parser 149
C.3.2 Black-box models 150
C.3.3 Multilevel parameter sweeps 151
C.4 Examples 152
C.4.1 Simple model of logistic growth 152
C.4.2 PhysiCell chemotaxis model 154
C.4.3 Connecting to third-party libraries 154
C.5 Future directions 156

This appendix is based on:

> Inês Godinho Gonçalves, David A. Hormuth II, Sandhya Prabhakaran, Caleb M. Phillips and José Manuel Garcia-Aznar. **"*PhysiCOOL*: A generalized framework for model Calibration and Optimization Of modeLing projects"** GigaByte (under review)

C.1 Introduction

In silico models of biological systems are usually very complex and rely on a large number of parameters describing physical and biological properties that require validation. As such, exploration of parameter space is an essential component of computational model development to fully characterize and validate simulation results. Experimental data may also be used to constrain parameter space (or enable model calibration) to enhance the biological relevance to model parameters. One widely used computational platform in the mathematical biology community is *PhysiCell* which provides a standardized approach to agent-based models of biological phenomena at different time and spatial scales. One limitation of *PhysiCell*, however, is that there has not been a generalized approach for parameter space exploration and calibration that can be run without high-performance computing access. Taking this into account, we present *PhysiCOOL*, an open-source Python library tailored to create standardized calibration and optimization routines of *PhysiCell* models.

C.2 Statement of Need

Mathematical biology is a field of study that aims to represent biological systems through the language of mathematics as a set of mathematical rules which can be used to test hypotheses and make predictions (Clermont and Zenker, 2015). There are several types of mathematical models that can be used to simulate biological systems at varying complexity levels. Agent-based models are one of the most popular implementations to develop models that consider the cellular and sub-cellular scales. Currently, multiple computational frameworks are available to facilitate the creation of agent-based models based on previously built templates, making mathematical biology more accessible to researchers from different backgrounds (Metzcar et al., 2019). Among these platforms, *PhysiCell* (Ghaffarizadeh et al., 2018) is an open-source hybrid framework that is able to simulate cells as discrete agents and model the reaction-diffusion dynamics of the substances present in the surrounding microenvironment through a continuous approach. Furthermore, recent add-ons have been developed to introduce new biological processes into the *PhysiCell* ecosystem (Letort et al., 2019; Bergman et al., 2022; Gonçalves and Garcia-Aznar, 2021).

Despite the recent advances in the development of additional *PhysiCell* plugins, the new modules are mostly centred around model extensions. Nevertheless, model exploration can be as important as model development to validate results and evaluate whether the model predictions about the underlying biological mechanisms are plausible (Hasenauer et al., 2015). Furthermore, experimental data could be used to provide biological and/or physical constraints on model parameters to validate whether the model captures the range of expected biological behaviors (Kazerouni et al., 2020), they can be used in computational biology to understand which model parameters maximize the similarity between the model results and a target data set, for example experimentally measured data. Subsequently, model developers may take these optimal solutions into account

to identify which biological mechanisms captured by the computational model may explain the experimental data.

We highlight that previous works have developed model exploration routines with *PhysiCell* (Duggan et al., 2021; Ozik et al., 2018), but these were specifically designed for high-performance computing (HPC) and distributed systems. Hence, currently, general *PhysiCell* users without access to such resources, or whose needs do not require them, must develop their own scripts to process simulation results and perform model exploration studies. As well as introducing a barrier to scientific progress depending on the researchers' programming knowledge level and computing resources, bespoke HPC work flows lack standardization that may enable a widespread use in the mathematical biology community (Banga, 2008).

Taking into account that there is still a need in the *PhysiCell* community for a standardized tool that implements calibration and optimization routines, we present *PhysiCOOL*, a generalized framework for model calibration and optimization of modeling projects written in *PhysiCell*. *PhysiCOOL* aims to be model agnostic, in the sense that models are treated as a black-box that can be executed through Python, making it suitable to several kinds of biological problems. Moreover, our library includes a built-in multilevel optimization routine for parameter estimation that is constrained by target output (experimental or otherwise). A visual representation of the new functionalities added by *PhysiCOOL* to the *PhysiCell* ecossystem is shown in Fig C.1. We also provide two practical examples of how *PhysiCOOL* can be used, showcasing *PhysiCOOL*'s optimization routine at two distinct complexity levels. Furthermore, we show how *PhysiCOOL* black-boxes can be used to couple *PhysiCell* with other publicly-available Python libraries for model optimization.

C.3 Implementation

PhysiCOOL is a Python library that requires Python 3.8 or higher. This package was created to work specifically with *PhysiCell* models and it fully supports *PhysiCell* v.1.10.2 and lower, and has partial support for *PhysiCell* v1.10.3 and higher. Furthermore, *PhysiCOOL* has been tested extensively and includes several unit tests to assure that its modules are working as expected and that it can be used in different platforms.

C.3.1 Configuration file parser

As for many several computational modelling frameworks, *PhysiCell* models are initialized with values stored in a text-based configuration file, namely a Extensible Markup Language (XML) file (Ghaffarizadeh et al., 2018). Thus, in parameter sweeps and sensitivity analysis studies it is necessary to open these files and modify the parameter values to be studied every time a new simulation is run. This process can be done manually, either by editing the XML file directly or using graphical user interface (GUI) tools such as *xml2jupyter* (Heiland et al., 2019). However, it becomes unfeasible to repeat this action several times in large scale studies. Henceforth, it is crucial to automate this process to run optimization and calibration workflows. Although it is possible to create Python scripts that will edit these files automatically with a standard module

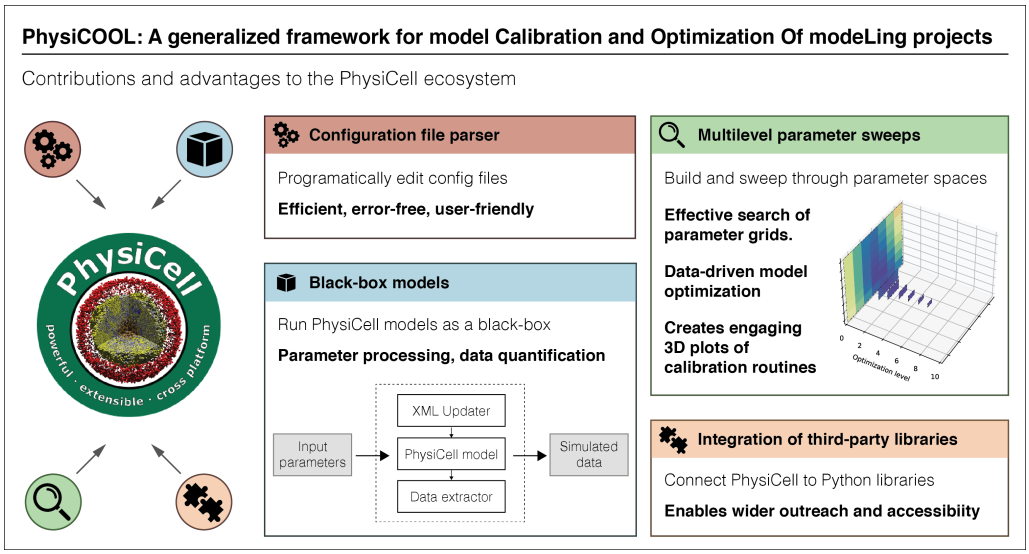

Figure C.1: *PhysiCOOL*'s contributions and advantages to the *PhysiCell* ecossystem. PhysiCOOL aims to improve the way researchers design and implement their parameter and calibration studies for models written in *PhysiCell*. To this end, *PhysiCOOL* introduces new functionalities, such as configuration file parser that updates configuration files in an error-free and user-friendly manner. *PhysiCOOL* also enables users to turn models into black-box models, making the optimization pipeline model-agnostic, and it implements a multilevel parameter sweep routine to optimize models using some target data. Lastly, *PhysiCOOL* facilitates the integration of third-party libraries, which makes *PhysiCell* more accessible.

such as *ElementTree*[1], doing so requires users to identify the values to be updated with long strings that reflect the structure of the XML file, as shown in the code snippet below.

Here, we aimed to develop a Python class that enables users to read the data from these configuration files in a more efficient manner, making this process less prone to errors. We implemented a *ConfigurationFileParser* class that reads the data from the configuration file into custom Python objects that follow the expected structure and data requirements defined in the XML file. Variable types and numerical constraints are validated when new instances of these data classes are created and when their values are updated. To achieve this, we implemented our classes using *Pydantic*[2] which improves data validation in Python. An example of how the task described in the code snippet presented previously may be implemented in a more user-friendly way with *PhysiCOOL* can be found below.

C.3.2 Black-box models

In complex and large computational models, it may be challenging or even impossible to estimate the model outputs analytically. Consequently, it is common to conduct calibration and optimization studies by running several simulations and performing sensitivity analysis studies to identify

[1] https://docs.python.org/3/library/xml.etree.elementtree.html
[2] https://github.com/samuelcolvin/pydantic

how model outputs change in response to different input parameter values. This process is recognized as simulation-based optimization or black-box optimization Alarie et al. (2021). *PhysiCell* models are written in C++ and should be compiled to produce an executable file that can be run to produce simulation results. In order to test and characterize the response of these models, it is generally necessary to conduct three tasks:

1. Update the PhysiCell configuration file with input parameters values;
2. Run the PhysiCell model;
3. Read the model outputs and compute a desired output metric.

These tasks can be performed manually. Nonetheless, it is not feasible or productive to do so in large computational studies, specifically when trying to characterize the model's response to a large number of input parameter values that can be inside a large range and require multiple simulation runs. Hence, PhysiCOOL allows users to create black-box models using the *PhysiCell-BlackBox* class and automatically perform the aforementioned tasks through Python.

These black-box models are modular in the sense that the users can select what functions to use to update the configuration file (i) and to process the results (iii). For instance, users can decide to change the cells' motility parameters and evaluate the effect on the distance traveled by cells over time. Alternatively, the cell cycling rates could be varied to analyze the evolution of the number of cells. Furthermore, it is not essential that both (i) and (iii) are defined in the black-box. In fact, users can also create black-box models composed only of the PhysiCell executable and use our approach to run multiple simulation replicates.

PhysiCOOL offers some built-in data quantification methods that can be used to extract and process data in step (iii). For example, functions are provided to obtain the final number of cells in a simulation, the final coordinates of the cells and the concentration of a given substance over the simulation domain. Furthermore, these methods can be employed by users to process simulation results and generate 2D and 3D plots of the cells and the microenvironment.

C.3.3 Multilevel parameter sweeps

Parameter optimization studies require the definition of a search space, which defines the range of the parameter values that will be studied. There are multiple approaches to defining this space and how to explore it. For example, random search algorithms can be employed to randomly sampling points within a defined bounded parameter space. Alternatively, a grid search, while a more computationally expensive option, systematically samples every point within a defined parameter grid space providing a more comprehensive overview of the model's response than that offered by random search.

PhysiCOOL implements a multilevel parameter sweep class (*MultiLevelSweep*) that is aimed at identifying the parameters that best fit a target data set through a grid search. In this example, the parameter sweep considers two PhysiCell parameters for which the user should provide initial

Table C.1: Parameter values used in the multilevel optimization examples.

Example	Initial point	Points	Percentage	Levels	Estimated	Target
Logistic growth	[0.15, 1000.0]	8	50 %	7	[0.10, 994.7]	[0.10, 1000.0]
Chemotaxis	[3.0, 0.7]	6	50 %	4	[3.0, 0.75]	[2.0, 0.9]

values. At each level, *MultiLevelSweep* creates a search grid based on these two values, the number of points per direction and the percentage per direction. These values should be configured by the user.

The results for each simulation are compared to the target data and the error between both datasets is computed and stored. At the end of the level, the parameters that provided the minimum error value are selected as the center of the parameter exploration grid for the next level and the parameter bounds are updated accordingly. The number of levels can be defined by the user.

C.4 Examples

C.4.1 Simple model of logistic growth

The first example was implemented to calibrate two parameters of a simple model of logistic growth based on some target data that defines a generated growth curve. Therefore, it serves as an introduction to this *PhysiCOOL* feature, as users are able to fully understand the behaviour of this simple model. We modelled the number of agents in a population, N, over a period of time t through a logistic function given by Eq C.1:

$$N(t) = \frac{KN_0}{N_0 + (K - N_0)\exp(-rt)} \tag{C.1}$$

where K represents the carrying capacity, i.e., the maximum population size, N_0 represents the number of initial agents and r is the proliferation rate. In this study, we fixed the initial number of agents and evaluated how the carrying capacity and the proliferation rate regulated the growth curve of a population. An example of two growth curves obtained for different model parameters is shown in Fig C.2(a).

We generated some target data using this model ($K = 1000, r = 0.1$) and, subsequently, we used *PhysiCOOL*'s multilevel sweep algorithm to evaluate if we could estimate these model parameters based on their resulting growth curve. To do so, we first created a search grid based on a set of user-defined values: an initial estimate for both parameters, the number of points to search in each direction of the search grid, the percentage to vary in each direction and the number of levels to search. These values can be found in Table C.1.

Fig C.2(b) shows the error between the target and simulated datasets for every cell of the parameter space after one level of the multilevel search. At this point, a new point estimate was calculated based on the parameter values that minimized the error between the two datasets. Likewise, the parameter space was adjusted to the area of interest and the process was repeated in the

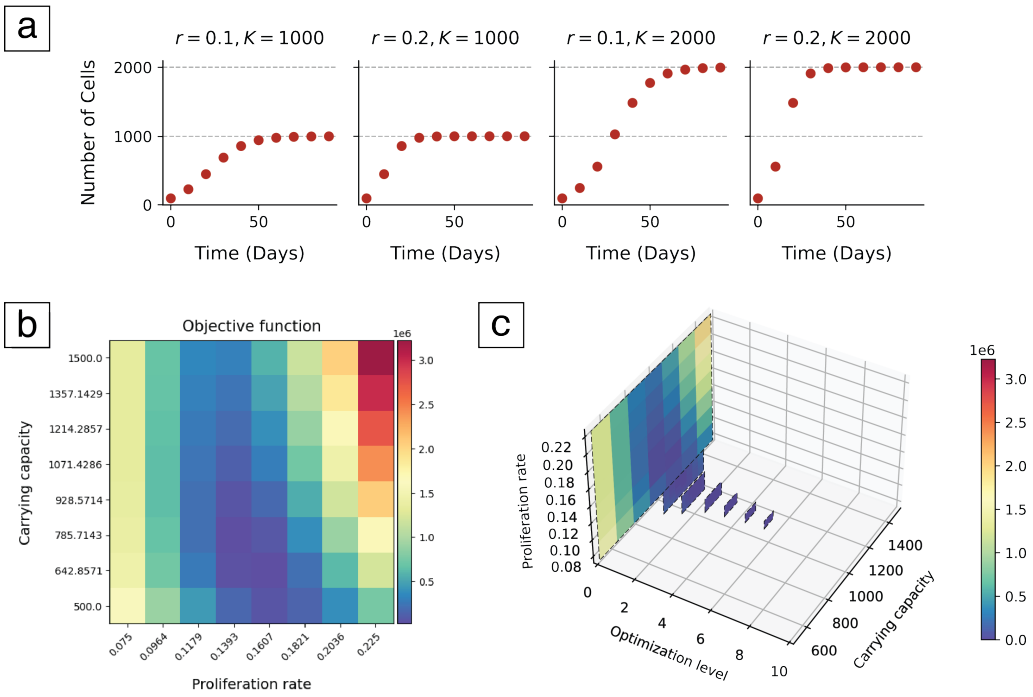

Figure C.2: Model and optimization results for the logistic growth example. (a) Growth curves obtained for different parameter sets (carrying capacity, K, and proliferation rate, r). (b) Optimization results after the completion of the first level of the multilevel optimization algorithm. The heatmap shows the difference, as given by the summed squared error, between the target data and the data produced by each cell's input parameters. (c) Optimization results after 7 levels of the multilevel optimization algorithm. Results converged to the parameters that originated the target data.

new parameter grid. This process was repeated for each level of the search and the results are shown in Fig C.2(c).

C.4.2 PhysiCell chemotaxis model

The second example can be classified as a more complex problem, since it was developed to calibrate a chemotaxis model written in *PhysiCell*. In this modelling framework, the cells' chemotactic response, i.e., the ability to migrate along a substance gradient, is dictated by a bias value defined between 0 and 1 (Ghaffarizadeh et al., 2018). When cells have a migration bias of 0, they move according to a random walk. Conversely, if the value is set to 1, cells follow the substance gradient in a deterministic manner. Therefore, we developed a model to estimate the migration bias of the cells in response to an oxygen gradient, as well as their speed, based on their travelled distances.

We implemented a 2D simulation with an oxygen source on one of the domain walls, as defined by the model's boundary conditions, and a group of cells placed on the opposite wall, as shown in Fig C.3(a). We expected that the cells' final position would be modulated by the cells' sensitivity to the oxygen chemotactic gradient. On the one hand, if a cell population had low sensitivity and, thus, moved randomly, they would likely remain close to their initial position as they would move around without following any specific direction. On the other hand, cells that followed oxygen would move towards the opposite wall, as seen in panel C.3(b).

We generated some target data by running a simulation with a migration bias of 0.9 and a speed value of 2.0 μm/min and storing the final y coordinates of the cells. Subsequently, we ran our multilevel sweep pipeline to evaluate whether we could estimate the parameter values that originated this data with a set of initial points different from the target parameter values. The results for this study are shown in Fig C.3(c).

C.4.3 Connecting to third-party libraries

Given that *PhysiCOOL* makes it possible for users to turn their *PhysiCel*l models into black-box models that receive some input parameters and return an output metric, it is straightforward to couple them with third-party Python libraries that accept this kind of models. For example, *psweep* (Schmerler, 2022) is a Python library developed to run parameter studies that saves the input parameters values and the returned output metrics into a database. Users must define a set of parameters and, for each of the defined values, *psweep* will (i) run a given user-defined function that takes these parameters as input, and (ii) save the input and output values returned by this function into the database. Therefore, a PhysiCOOL black-box could seamlessly be integrated into step (i).

In addition, more sophisticated libraries could be considered to perform advanced optimization studies such as Approximate Bayesian Computation (ABC) and Bayesian Optimization for Likelihood-Free Inference (BOLFI) (Lintusaari et al., 2018). Henceforth, although *PhysiCOOL* offers built-in optimization routines, it can be used in a modular way to take advantage of other

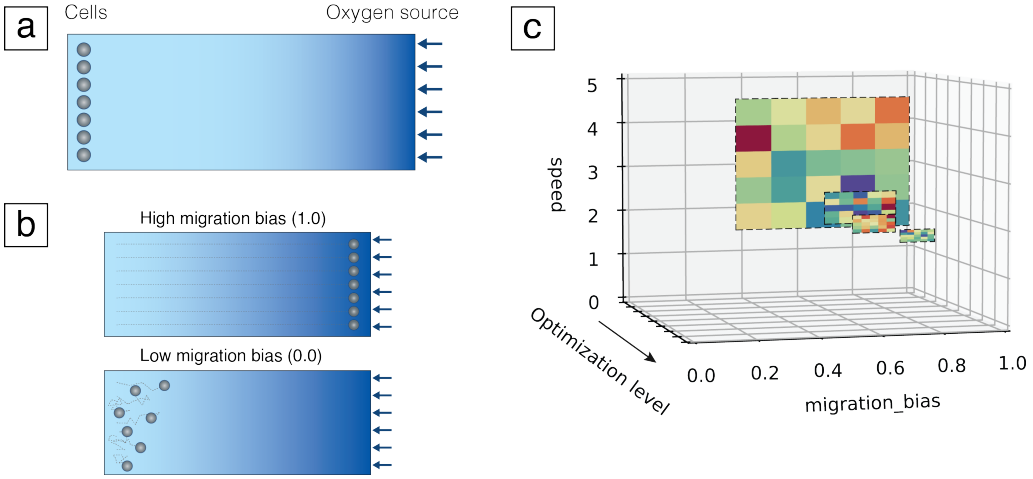

Figure C.3: Model and optimization results for the chemotaxis example. (a) Initial model configuration design. Cells (represented as grey circles) were placed close to a domain wall and an oxygen source (represented by the blue arrows) was simulated on the opposite wall, creating a chemotactic gradient that cells could follow. This gradient is illustrated by the colour gradient shown in the figure. (b) Expected model results for cells with different migration bias values. High migration bias populations were expected to migrate in a deterministic manner and follow the oxygen gradient, crossing the domain and arriving at the opposite wall, as shown by their trajectories, shown as grey dashed lines. On the other hand, cells with low migration bias were expected to move randomly and, thus, present low net displacement values. (c) Optimization results after 4 levels of the multilevel optimization algorithm. Results converged to the parameters that originated the target data. The colormap was updated for each level, describing the minimum and maximum error values at the current level.

libraries that may be more appropriate to a certain study or type of research, without the need to implement these optimization algorithms from scratch.

C.5 Future directions

At its current state of development, we believe that *PhysiCOOL* will already improve *PhysiCell*'s accessibility as it provides an intuitive interface to run studies in Python, which is more popular among biology researchers than C++, in which *PhysiCell* was originally written. Additionally, this standardized approach provides a straightforward workflow for integrating target data (defined from simulations or biological observations) to constrain parameter space for agent-based models. In the future, new features can be added to *PhysiCOOL*, such as the ability to generate non-linear parameter spaces and employ alternative optimization algorithms. Although future iterations of this library may include different optimization approaches, its modular design assures that advanced users are still able to build pipelines that suit their needs.

References

Acar, E., Plopper, G. E., and Yener, B. (2012). Coupled Analysis of In Vitro and Histology Tissue Samples to Quantify Structure-Function Relationship. *PLoS ONE*, 7(3):e32227.

Aceto, N., Bardia, A., Miyamoto, D., Donaldson, M., Wittner, B., Spencer, J., Yu, M., Pely, A., Engstrom, A., Zhu, H., Brannigan, B., Kapur, R., Stott, S., Shioda, T., Ramaswamy, S., Ting, D., Lin, C., Toner, M., Haber, D., and Maheswaran, S. (2014). Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell*, 158(5):1110–1122.

Ahrens, J., Geveci, B., and Law, C. (2005). ParaView: An End-User Tool for Large-Data Visualization. In *Visualization Handbook*, pages 717–731. Elsevier.

Alarie, S., Audet, C., Gheribi, A. E., Kokkolaras, M., and Le Digabel, S. (2021). Two decades of blackbox optimization applications. *EURO Journal on Computational Optimization*, 9:100011.

Alfarouk, K. O., Muddathir, A. K., and Shayoub, M. E. A. (2011). Tumor Acidity as Evolutionary Spite. *Cancers*, 3(1):408–414.

Almagro, J., Messal, H. A., Elosegui-Artola, A., van Rheenen, J., and Behrens, A. (2022). Tissue architecture in tumor initiation and progression. *Trends in Cancer*, 8(6):494–505.

Altrock, P. M., Liu, L. L., and Michor, F. (2015). The mathematics of cancer: integrating quantitative models. *Nature Reviews Cancer*, 15(12):730–745.

Anderson, A. R., Weaver, A. M., Cummings, P. T., and Quaranta, V. (2006). Tumor Morphology and Phenotypic Evolution Driven by Selective Pressure from the Microenvironment. *Cell*, 127(5):905–915.

Anderson, A. R. A. and Quaranta, V. (2008). Integrative mathematical oncology. *Nature Reviews Cancer*, 8(3):227–234.

Anderson, A. R. A., Rejniak, K. A., Gerlee, P., and Quaranta, V. (2009). Microenvironment driven invasion: a multiscale multimodel investigation. *Journal of Mathematical Biology*, 58(4-5):579–624.

Antonopoulos, M., Dionysiou, D., Stamatakos, G., and Uzunoglu, N. (2019). Three-dimensional tumor growth in time-varying chemical fields: a modeling framework and theoretical study. *BMC Bioinformatics*, 20(1):442.

Araujo, R. (2004). A history of the study of solid tumour growth: the contribution of mathematical modelling. *Bulletin of Mathematical Biology*, 66(5):1039–1091.

Astanin, S. and Preziosi, L. (2009). Mathematical modelling of the Warburg effect in tumour cords. *Journal of Theoretical Biology*, 258(4):578–590.

Aćimović, J., Mäki-Marttunen, T., Havela, R., Teppola, H., and Linne, M.-L. (2011). Modeling of Neuronal Growth In Vitro: Comparison of Simulation Tools NETMORPH and CX3D. *EURASIP Journal on Bioinformatics and Systems Biology*, 2011(1):616382.

Baker, R. E., Peña, J.-M., Jayamohan, J., and Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters*, 14(5):20170660.

Baker, R. E. and Simpson, M. J. (2010). Correcting mean-field approximations for birth-death-movement processes. *Physical Review E*, 82(4):041905.

Banga, J. R. (2008). Optimization in computational systems biology. *BMC Systems Biology*, 2(1):47.

Banwarth-Kuhn, M. and Sindi, S. (2020). How and why to build a mathematical model: A case study using prion aggregation. *Journal of Biological Chemistry*, 295(15):5022–5035.

Barnes, J. L., Zubair, M., John, K., Poirier, M. C., and Martin, F. L. (2018). Carcinogens and DNA damage. *Biochemical Society Transactions*, 46(5):1213–1224.

Barney, L., Jansen, L., Polio, S., Galarza, S., Lynch, M., and Peyton, S. (2016). The predictive link between matrix and metastasis. *Current Opinion in Chemical Engineering*, 11:85–93.

Basan, M., Prost, J., Joanny, J.-F., and Elgeti, J. (2011). Dissipative particle dynamics simulations for biological tissues: rheology and competition. *Physical Biology*, 8(2):026014.

Bayless, K. J., Kwak, H.-I., and Su, S.-C. (2009). Investigating endothelial invasion and sprouting behavior in three-dimensional collagen matrices. *Nature Protocols*, 4(12):1888–1898.

Bear, J. E. and Haugh, J. M. (2014). Directed migration of mesenchymal cells: where signaling and the cytoskeleton meet. *Current Opinion in Cell Biology*, 30:74–82.

Bekisz, S. and Geris, L. (2020). Cancer modeling: From mechanistic to data-driven approaches, and from fundamental insights to clinical applications. *Journal of Computational Science*, 46:101198.

Benzekry, S., Lamont, C., Beheshti, A., Tracz, A., Ebos, J. M. L., Hlatky, L., and Hahnfeldt, P. (2014). Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth. *PLoS Computational Biology*, 10(8):e1003800.

Bergman, D., Marazzi, L., Chowkwale, M., Maheshvare M, D., Bidanta, S., Mapder, T., and Li, J. (2022). PhysiPKPD: A pharmacokinetics and pharmacodynamics module for PhysiCell. *Gigabyte*, 2022:1–11.

Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).

Bertuzzi, A., Fasano, A., Gandolfi, A., and Sinisgalli, C. (2010). Necrotic core in EMT6/Ro tumour spheroids: Is it caused by an ATP deficit? *Journal of Theoretical Biology*, 262(1):142–150.

Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*, 15(2):107–120.

Bose, S., Zhang, C., and Le, A. (2021). Glucose Metabolism in Cancer: The Warburg Effect and Beyond. In Le, A., editor, *The Heterogeneity of Cancer Metabolism*, volume 1311, pages 3–15. Springer International Publishing, Cham. Series Title: Advances in Experimental Medicine and Biology.

Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356):791–799.

Braicu, Buse, Busuioc, Drula, Gulei, Raduly, Rusu, Irimie, Atanasov, Slaby, Ionescu, and Berindan-Neagoe (2019). A Comprehensive Review on MAPK: A Promising Therapeutic Target in Cancer. *Cancers*, 11(10):1618.

Bravo, R. R., Baratchart, E., West, J., Schenck, R. O., Miller, A. K., Gallaher, J., Gatenbee, C. D., Basanta, D., Robertson-Tessi, M., and Anderson, A. R. A. (2020). Hybrid Automata Library: A flexible platform for hybrid modeling with real-time visualization. *PLOS Computational Biology*, 16(3):e1007635.

Bray, F., Laversanne, M., Weiderpass, E., and Soerjomataram, I. (2021). The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer*, 127(16):3029–3030.

Breitwieser, L., Hesam, A., de Montigny, J., Vavourakis, V., Iosif, A., Jennings, J., Kaiser, M., Manca, M., Di Meglio, A., Al-Ars, Z., Rademakers, F., Mutlu, O., and Bauer, R. (2022). BioDynaMo: a modular platform for high-performance agent-based simulation. *Bioinformatics*, 38(2):453–460.

Breitwieser, L., Hesam, A., Rademakers, F., Luna, J. G., and Mutlu, O. (2023). High-Performance and Scalable Agent-Based Simulation with BioDynaMo. Publisher: arXiv Version Number: 1.

Breslin, S. and O'Driscoll, L. (2013). Three-dimensional cell culture: the missing link in drug discovery. *Drug Discovery Today*, 18(5-6):240–249.

Bronzert, D. A., Pantazis, P., Antoniades, H. N., Kasid, A., Davidson, N., Dickson, R. B., and Lippman, M. E. (1987). Synthesis and secretion of platelet-derived growth factor by human breast cancer cell lines. *Proceedings of the National Academy of Sciences*, 84(16):5763–5767.

Browning, A. P. and Simpson, M. J. (2023). Geometric analysis enables biological insight from complex non-identifiable models using simple surrogates. *PLOS Computational Biology*, 19(1):e1010844.

Bull, J. A. and Byrne, H. M. (2022). The Hallmarks of Mathematical Oncology. *Proceedings of the IEEE*, 110(5):523–540.

Bull, J. A., Mech, F., Quaiser, T., Waters, S. L., and Byrne, H. M. (2020). Mathematical modelling reveals cellular dynamics within tumour spheroids. *PLOS Computational Biology*, 16(8):e1007961.

Burton, A. C. (1966). Rate of growth of solid tumours as a problem of diffusion. *Growth*, 30(2):157–176.

Bustamante, D. J., Basile, E. J., Hildreth, B. M., Browning, N. W., Jensen, S. A., Moldovan, L., Petrache, H. I., and Moldovan, N. I. (2021). Biofabrication of spheroids fusion-based tumor models: computational simulation of glucose effects. *Biofabrication*, 13(3):035010.

Butcher, D. T., Alliston, T., and Weaver, V. M. (2009). A tense situation: forcing tumour progression. *Nature Reviews Cancer*, 9(2):108–122.

Buttenschön, A. and Edelstein-Keshet, L. (2020). Bridging from single to collective cell migration: A review of models and links to experiments. *PLOS Computational Biology*, 16(12):e1008411.

Byrne, H. M. (2010). Dissecting cancer through mathematics: from the cell to the animal model. *Nature Reviews Cancer*, 10(3):221–230.

Camley, B. A. and Rappel, W.-J. (2017). Physical models of collective cell motility: from cell to tissue. *Journal of Physics D: Applied Physics*, 50(11):113002.

Campos, D., Méndez, V., and Llopis, I. (2010). Persistent random motion: Uncovering cell migration dynamics. *Journal of Theoretical Biology*, 267(4):526–534.

Canel, M., Serrels, A., Frame, M. C., and Brunton, V. G. (2013). E-cadherin–integrin crosstalk in cancer invasion and metastasis. *Journal of Cell Science*, 126(2):393–401.

Cannata, N., Corradini, F., Merelli, E., and Tesei, L. (2013). Agent-Based Models of Cellular Systems. In Reisfeld, B. and Mayeno, A. N., editors, *Computational Toxicology*, volume 930, pages 399–426. Humana Press, Totowa, NJ. Series Title: Methods in Molecular Biology.

Cao, Y. (2017). Tumorigenesis as a process of gradual loss of original cell identity and gain of properties of neural precursor/progenitor cells. *Cell & Bioscience*, 7(1):61.

Carey, S. P., D'Alfonso, T. M., Shin, S. J., and Reinhart-King, C. A. (2012). Mechanobiology of tumor invasion: Engineering meets oncology. *Critical Reviews in Oncology/Hematology*, 83(2):170–183.

Casciari, J. J., Sotirchos, S. V., and Sutherland, R. M. (1988). Glucose diffusivity in multicellular tumor spheroids. *Cancer Research*, 48(14):3905–3909.

Casciari, J. J., Sotirchos, S. V., and Sutherland, R. M. (1992a). Mathematical modelling of microenvironment and growth in EMT6/Ro multicellular tumour spheroids. *Cell Proliferation*, 25(1):1–22.

Casciari, J. J., Sotirchos, S. V., and Sutherland, R. M. (1992b). Variations in tumor cell growth rates and metabolism with oxygen concentration, glucose concentration, and extracellular pH. *Journal of Cellular Physiology*, 151(2):386–394.

Casás-Selves, M. and DeGregori, J. (2011). How Cancer Shapes Evolution and How Evolution Shapes Cancer. *Evolution: Education and Outreach*, 4(4):624–634.

Catig, G. C., Figueroa, S., and Moore, M. J. (2015). Experimental and computational models of neurite extension at a choice point in response to controlled diffusive gradients. *Journal of Neural Engineering*, 12(4):046012.

Chaffer, C. L. and Weinberg, R. A. (2015). How Does Multistep Tumorigenesis Really Proceed? *Cancer Discovery*, 5(1):22–24.

Chaplain, M. (1996). Avascular growth, angiogenesis and vascular growth in solid tumours: The mathematical modelling of the stages of tumour development. *Mathematical and Computer Modelling*, 23(6):47–87.

Charras, G. and Sahai, E. (2014). Physical influences of the extracellular environment on cell migration. *Nature Reviews Molecular Cell Biology*, 15(12):813–824.

Cheng, G., Tse, J., Jain, R. K., and Munn, L. L. (2009). Micro-Environmental Mechanical Stress Controls Tumor Spheroid Size and Morphology by Suppressing Proliferation and Inducing Apoptosis in Cancer Cells. *PLoS ONE*, 4(2):e4632.

Chertock, A., Kurganov, A., Wang, X., Wu, Y., ,Department of Mathematics, North Carolina State University, Raleigh, NC 27695, ,Mathematics Department, Tulane University, New Orleans, LA 70118, and ,Department of Mathematics, Capital Normal University, Beijing 100048 (2012). On a chemotaxis model with saturated chemotactic flux. *Kinetic & Related Models*, 5(1):51–95.

Chow, C. and Brittingham, J. (1987). Perivascular pseudorosettes in childhood brain tumours: An ultrastructural and immunohistochemical study. *Pathology*, 19(1):12–16.

Clark, A. G. and Vignjevic, D. M. (2015). Modes of cancer cell invasion and the role of the microenvironment. *Current Opinion in Cell Biology*, 36:13–22.

Cleri, F. (2019). Agent-based model of multicellular tumor spheroid evolution including cell metabolism. *The European Physical Journal E*, 42(8):112.

Clermont, G. and Zenker, S. (2015). The inverse problem in mathematical biology. *Mathematical Biosciences*, 260:11–15.

Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of The Royal Society Interface*, 5(25):813–834.

Colombi, A., Scianna, M., and Preziosi, L. (2017). Coherent modelling switch between pointwise and distributed representations of cell aggregates. *Journal of Mathematical Biology*, 74(4):783–808.

Colombi, A., Scianna, M., and Preziosi, L. (2020). Collective migration and patterning during early development of zebrafish posterior lateral line. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1807):20190385.

Comaniciu, D., Ramesh, V., and Meer, P. (2000). Real-time tracking of non-rigid objects using mean shift. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, volume 2, pages 142–149, Hilton Head Island, SC, USA. IEEE Comput. Soc.

Cooper, G. M. (2000). *The cell: a molecular approach*. ASM Press [u.a.], Washington, DC, 2. ed edition.

Costa, E., Ferreira-Gonçalves, T., Chasqueira, G., Cabrita, A. S., Figueiredo, I. V., and Reis, C. P. (2020). Experimental Models as Refined Translational Tools for Breast Cancer Research. *Scientia Pharmaceutica*, 88(3):32.

Craver, C. F. (2006). When mechanistic models explain. *Synthese*, 153(3):355–376.

Cress, D., Engel, B., and Santiago-Cardona, P. (2014). The retinoblastoma protein: a master tumor suppressor acts as a link between cell cycle and cell adhesion. *Cell Health and Cytoskeleton*, page 1.

Cóndor, M., Mark, C., Gerum, R. C., Grummel, N. C., Bauer, A., García-Aznar, J. M., and Fabry, B. (2019). Breast Cancer Cells Adapt Contractile Forces to Overcome Steric Hindrance. *Biophysical Journal*, 116(7):1305–1312.

Cóndor, M., Rüberg, T., Borau, C., Piles, J., and García-Aznar, J. (2018). A web-based application for automated quantification of chemical gradients induced in microfluidic devices. *Computers in Biology and Medicine*, 95:118–128.

Dada, J. O. and Mendes, P. (2011). Multi-scale modelling and simulation in systems biology. *Integrative Biology*, 3(2):86.

Dai, Z., Shestov, A., Lai, L., and Locasale, J. (2016). A Flux Balance of Glucose Metabolism Clarifies the Requirements of the Warburg Effect. *Biophysical Journal*, 111(5):1088–1100.

Dalle Nogare, D. and Chitnis, A. B. (2017). A framework for understanding morphogenesis and migration of the zebrafish posterior Lateral Line primordium. *Mechanisms of Development*, 148:69–78.

Danhier, P., Bański, P., Payen, V. L., Grasso, D., Ippolito, L., Sonveaux, P., and Porporato, P. E. (2017). Cancer metabolism in space and time: Beyond the Warburg effect. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1858(8):556–572.

Das, D., Bhattacharjee, K., Barthakur, S., Tahiliani, P., Deka, P., Bhattacharjee, H., Deka, A., and Paul, R. (2014). A new rosette in retinoblastoma. *Indian Journal of Ophthalmology*, 62(5):638.

de Montigny, J., Iosif, A., Breitwieser, L., Manca, M., Bauer, R., and Vavourakis, V. (2021). An in silico hybrid continuum-agent-based procedure to modelling cancer development: Interrogating the interplay amongst glioma invasion, vascularity and necrosis. *Methods*, 185:94–104.

De Pieri, A., Korman, B. D., Jüngel, A., and Wuertz-Kozak, K. (2021). Engineering Advanced In Vitro Models of Systemic Sclerosis for Drug Discovery and Development. *Advanced Biology*, 5(4):2000168.

DeBerardinis, R. J. and Chandel, N. S. (2016). Fundamentals of cancer metabolism. *Science Advances*, 2(5):e1600200.

Del Amo, C., Borau, C., Movilla, N., Asín, J., and García-Aznar, J. M. (2017). Quantifying 3D chemotaxis in microfluidic-based chips with step gradients of collagen hydrogel concentrations. *Integrative Biology*, 9(4):339–349.

Del Amo, C., Olivares, V., Cóndor, M., Blanco, A., Santolaria, J., Asín, J., Borau, C., and García-Aznar, J. M. (2018). Matrix architecture plays a pivotal role in 3D osteoblast migration: The effect of interstitial fluid flow. *Journal of the Mechanical Behavior of Biomedical Materials*, 83:52–62.

Delile, J., Herrmann, M., Peyriéras, N., and Doursat, R. (2017). A cell-based computational model of early embryogenesis coupling mechanical behaviour and gene regulation. *Nature Communications*, 8(1):13929.

Dell'Aquila, M., Musarra, T., Fiorentino, V., Brunelli, C., De Marco, C., Raffaelli, M., Traini, E., Pio Lombardi, C., Fadda, G., Larocca, L. M., Pantanowitz, L., and Rossi, E. D. (2020). Relevance of rosette patterns in variants of papillary thyroid carcinoma. *Cytopathology*, 31(6):533–540.

Dennerll, T. J., Lamoureux, P., Buxbaum, R. E., and Heidemann, S. R. (1989). The cytomechanics of axonal elongation and retraction. *Journal of Cell Biology*, 109(6):3073–3083.

Deuel, T. F., Kawahara, R. S., Mustoe, T. A., and Pierce, G. F. (1991). Growth Factors and Wound Healing: Platelet-Derived Growth Factor as a Model Cytokine. *Annual Review of Medicine*, 42(1):567–584.

Dhillon, A. S., Hagan, S., Rath, O., and Kolch, W. (2007). MAP kinase signalling pathways in cancer. *Oncogene*, 26(22):3279–3290.

Di Costanzo, E., Natalini, R., and Preziosi, L. (2015). A hybrid mathematical model for self-organizing cell migration in the zebrafish lateral line. *Journal of Mathematical Biology*, 71(1):171–214.

Dickinson, R. B., Guido, S., and Tranquillo, R. T. (1994). Biased cell migration of fibroblasts exhibiting contact guidance in oriented collagen gels. *Annals of Biomedical Engineering*, 22(4):342–356.

Discher, D. E., Janmey, P., and Wang, Y.-l. (2005). Tissue Cells Feel and Respond to the Stiffness of Their Substrate. *Science*, 310(5751):1139–1143.

Dormann, S. and Deutsch, A. (2002). Modeling of self-organized avascular tumor growth with a hybrid cellular automaton. *In Silico Biology*, 2(3):393–406.

Doyle, A. D., Sykora, D. J., Pacheco, G. G., Kutys, M. L., and Yamada, K. M. (2021). 3D mesenchymal cell migration is driven by anterior cellular contraction that generates an extracellular matrix prestrain. *Developmental Cell*, 56(6):826–841.e4.

Drasdo, D. (2007). Center-based Single-cell Models: An Approach to Multi-cellular Organization Based on a Conceptual Analogy to Colloidal Particles. In Anderson, A. R. A., Chaplain, M. A. J., and Rejniak, K. A., editors, *Single-Cell-Based Models in Biology and Medicine*, pages 171–196. Birkhäuser Basel, Basel. Series Title: Mathematics and Biosciences in Interaction.

Drasdo, D. and Höhme, S. (2005). A single-cell-based model of tumor growth *in vitro* : monolayers and spheroids. *Physical Biology*, 2(3):133–147.

Drasdo, D. and Loeffler, M. (2001). Individual-based models to growth and folding in one-layered tissues: intestinal crypts and early development. *Nonlinear Analysis: Theory, Methods & Applications*, 47(1):245–256.

Du, J., Zhou, Y., Jin, L., and Sheng, K. (2022). Gell: A GPU-powered 3D hybrid simulator for large-scale multicellular system. preprint, Systems Biology.

Duarte Campos, D. F., Bonnin Marquez, A., O'Seanain, C., Fischer, H., Blaeser, A., Vogt, M., Corallo, D., and Aveic, S. (2019). Exploring Cancer Cell Behavior In Vitro in Three-Dimensional Multicellular Bioprintable Collagen-Based Hydrogels. *Cancers*, 11(2):180.

Duggan, B., Metzcar, J., and Macklin, P. (2021). DAPT: A package enabling distributed automated parameter testing. *Gigabyte*, 2021:1–10.

Duhaylongsod, F. G., Lowe, V. J., Patz, E. F., Vaughn, A. L., Coleman, R. E., and Wolfe, W. G. (1995). Lung tumor growth correlates with glucose metabolism measured by fluoride-18 fluorodeoxyglucose positron emission tomography. *The Annals of Thoracic Surgery*, 60(5):1348–1352.

Dyba, T., Randi, G., Bray, F., Martos, C., Giusti, F., Nicholson, N., Gavin, A., Flego, M., Neamtiu, L., Dimitrova, N., Negrão Carvalho, R., Ferlay, J., and Bettio, M. (2021). The European cancer burden in 2020: Incidence and mortality estimates for 40 countries and 25 major cancers. *European Journal of Cancer*, 157:308–347.

D'Arcy, M. S. (2019). Cell death: a review of the major forms of apoptosis, necrosis and autophagy. *Cell Biology International*, 43(6):582–592.

Edelman, L. B., Eddy, J. A., and Price, N. D. (2010). *In silico* models of cancer. *WIREs Systems Biology and Medicine*, 2(4):438–459.

Ehrbar, M., Sala, A., Lienemann, P., Ranga, A., Mosiewicz, K., Bittermann, A., Rizzi, S., Weber, F., and Lutolf, M. (2011). Elucidating the Role of Matrix Stiffness in 3D Cell Migration and Remodeling. *Biophysical Journal*, 100(2):284–293.

Enderling, H. and Rejniak, K. A. (2013). Simulating Cancer: Computational Models in Oncology. *Frontiers in Oncology*, 3.

Enderling, H. and Wolkenhauer, O. (2021). Are all models wrong? *Computational and Systems Oncology*, 1(1).

Escribano, J., Chen, M. B., Moeendarbary, E., Cao, X., Shenoy, V., Garcia-Aznar, J. M., Kamm, R. D., and Spill, F. (2019). Balance of mechanical forces drives endothelial gap formation and may facilitate cancer and immune-cell extravasation. *PLOS Computational Biology*, 15(5):e1006395.

Fachada, N., Lopes, V. V., Martins, R. C., and Rosa, A. C. (2017). Parallelization Strategies for Spatial Agent-Based Models. *International Journal of Parallel Programming*, 45(3):449–481.

Feitelson, M. A., Arzumanyan, A., Kulathinal, R. J., Blain, S. W., Holcombe, R. F., Mahajna, J., Marino, M., Martinez-Chantar, M. L., Nawroth, R., Sanchez-Garcia, I., Sharma, D., Saxena, N. K., Singh, N., Vlachostergios, P. J., Guo, S., Honoki, K., Fujii, H., Georgakilas, A. G., Bilsland, A., Amedei, A., Niccolai, E., Amin, A., Ashraf, S. S., Boosani, C. S., Guha, G., Ciriolo, M. R., Aquilano, K., Chen, S., Mohammed, S. I., Azmi, A. S., Bhakta, D., Halicka, D., Keith, W. N., and Nowsheen, S. (2015). Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. *Seminars in Cancer Biology*, 35:S25–S54.

Fernández Slezak, D., Suárez, C., Cecchi, G. A., Marshall, G., and Stolovitzky, G. (2010). When the Optimal Is Not the Best: Parameter Estimation in Complex Biological Models. *PLoS ONE*, 5(10):e13283.

Fidler, I. J. (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nature Reviews Cancer*, 3(6):453–458.

Fletcher, A. G. and Osborne, J. M. (2022). Seven challenges in the multiscale modeling of multicellular tissues. *WIREs Mechanisms of Disease*, 14(1).

Fletcher, L., Mcfarlane, E., Fletcher, K., Grant, P. A., and Campbell, I. C. (2000). A Mathematical Model for Assessing Changes in Neurofilament Protein Levels in Neurites and Cell Bodies of Differentiating Neuroblastoma Cells. *Journal of Theoretical Biology*, 205(2):241–252.

Ford Versypt, A. N. (2021). Multiscale modeling in disease. *Current Opinion in Systems Biology*, 27:100340.

Fraley, S. I., Wu, P.-h., He, L., Feng, Y., Krisnamurthy, R., Longmore, G. D., and Wirtz, D. (2015). Three-dimensional matrix fiber alignment modulates cell migration and MT1-MMP utility by spatially and temporally directing protrusions. *Scientific Reports*, 5(1):14580.

Frantz, C., Stewart, K. M., and Weaver, V. M. (2010). The extracellular matrix at a glance. *Journal of Cell Science*, 123(24):4195–4200.

Freyer, J. P. and Sutherland, R. M. (1985). A reduction in the in situ rates of oxygen and glucose consumption of cells in EMT6/Ro spheroids during growth. *Journal of Cellular Physiology*, 124(3):516–524.

Friedl, P. and Wolf, K. (2003). Proteolytic and non-proteolytic migration of tumour cells and leucocytes. *Biochemical Society Symposia*, 70:277–285.

Friedl, P. and Wolf, K. (2010). Plasticity of cell migration: a multiscale tuning model. *Journal of Cell Biology*, 188(1):11–19.

Fusco, G. and Minelli, A. (2010). Phenotypic plasticity in development and evolution: facts and concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1540):547–556.

Garay, T., Juhász, , Molnár, E., Eisenbauer, M., Czirók, A., Dekan, B., László, V., Hoda, M. A., Döme, B., Tímár, J., Klepetko, W., Berger, W., and Hegedűs, B. (2013). Cell migration or cytokinesis and proliferation? – Revisiting the "go or grow" hypothesis in cancer cells in vitro. *Experimental Cell Research*, 319(20):3094–3103.

Gatenby, R. A. and Maini, P. K. (2003). Mathematical oncology: Cancer summed up. *Nature*, 421(6921):321–321.

Gatto, F., Miess, H., Schulze, A., and Nielsen, J. (2015). Flux balance analysis predicts essential genes in clear cell renal cell carcinoma metabolism. *Scientific Reports*, 5(1):10738.

Ge, H., Tian, M., Pei, Q., Tan, F., and Pei, H. (2021). Extracellular Matrix Stiffness: New Areas Affecting Cell Metabolism. *Frontiers in Oncology*, 11:631991.

Gerlee, P. (2013). The Model Muddle: In Search of Tumor Growth Laws. *Cancer Research*, 73(8):2407–2411.

Gerlee, P. and Anderson, A. (2007). An evolutionary hybrid cellular automaton model of solid tumour growth. *Journal of Theoretical Biology*, 246(4):583–603.

Germann, P., Marin-Riera, M., and Sharpe, J. (2019). yalla: GPU-Powered Spheroid Models for Mesenchyme and Epithelium. *Cell Systems*, 8(3):261–266.e3.

Ghaffarizadeh, A., Friedman, S. H., and Macklin, P. (2016). BioFVM: an efficient, parallelized diffusive transport solver for 3-D biological simulations. *Bioinformatics*, 32(8):1256–1258.

Ghaffarizadeh, A., Heiland, R., Friedman, S. H., Mumenthaler, S. M., and Macklin, P. (2018). PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems. *PLOS Computational Biology*, 14(2):e1005991.

Giese, A., Bjerkvig, R., Berens, M., and Westphal, M. (2003). Cost of Migration: Invasion of Malignant Gliomas and Implications for Treatment. *Journal of Clinical Oncology*, 21(8):1624–1636.

Gombert, A. K. and Nielsen, J. (2000). Mathematical modelling of metabolism. *Current Opinion in Biotechnology*, 11(2):180–186.

Gonzalez-de Aledo, P., Vladimirov, A., Manca, M., Baugh, J., Asai, R., Kaiser, M., and Bauer, R. (2018). An optimization approach for agent-based computational models of biological development. *Advances in Engineering Software*, 121:262–275.

González-Valverde, I. and García-Aznar, J. M. (2017). A hybrid computational model to explore the topological characteristics of epithelial tissues. *International Journal for Numerical Methods in Biomedical Engineering*, 33(11):e2877.

González-Valverde, I. and García-Aznar, J. M. (2018). Mechanical modeling of collective cell migration: An agent-based and continuum material approach. *Computer Methods in Applied Mechanics and Engineering*, 337:246–262.

González-Valverde, I., Semino, C., and García-Aznar, J. (2016). Phenomenological modelling and simulation of cell clusters in 3D cultures. *Computers in Biology and Medicine*, 77:249–260.

Gonçalves, I. G. and Garcia-Aznar, J. M. (2021). Extracellular matrix density regulates the formation of tumour spheroids through cell migration. *PLOS Computational Biology*, 17(2):e1008764.

Gonçalves, I. G. and García-Aznar, J. M. (2022). m2be-igg/neurorosettes: v0.1.2.

Gonçalves, I. G., Gonçalves, I. G., and David (2023). IGGoncalves/PhysiCOOL: v0.2.4.

Gonçalves, I. G., Hormuth, D. A., Prabhakaran, S., Phillips, C. M., and García-Aznar, J. M. (2022). PhysiCOOL: A generalized framework for model Calibration and Optimization Of modeLing projects. preprint, Bioengineering.

Goodman, C. S., Bastiani, M. J., Doe, C. Q., du Lac, S., Helfand, S. L., Kuwada, J. Y., and Thomas, J. B. (1984). Cell Recognition During Neuronal Development. *Science*, 225(4668):1271–1279.

Graziani, V., Rodriguez-Hernandez, I., Maiques, O., and Sanz-Moreno, V. (2022). The amoeboid state as part of the epithelial-to-mesenchymal transition programme. *Trends in Cell Biology*, 32(3):228–242.

Greenspan, H. P. (1972). Models for the Growth of a Solid Tumor by Diffusion. *Studies in Applied Mathematics*, 51(4):317–340.

Griffith, L. G. and Swartz, M. A. (2006). Capturing complex 3D tissue physiology in vitro. *Nature Reviews Molecular Cell Biology*, 7(3):211–224.

Guimarães, C. F., Gasperini, L., Marques, A. P., and Reis, R. L. (2020). The stiffness of living tissues and its implications for tissue engineering. *Nature Reviews Materials*, 5(5):351–370.

Gupta, P. B., Pastushenko, I., Skibinski, A., Blanpain, C., and Kuperwasser, C. (2019). Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. *Cell Stem Cell*, 24(1):65–78.

Gupta, S., Gupta, N., Ahlawat, S., Kumar, A., Taneja, R., Sharma, R., Sunder, S., and Tantia, M. (2005). In Vitro Culture of Skin Fibroblast Cells for Potential Cloning by Nuclear Transfer. In Makkar, H. P. and Viljoen, G. J., editors, *Applications of Gene-Based Technologies for Improving Animal Production and Health in Developing Countries*, pages 631–640. Springer-Verlag, Berlin/Heidelberg.

Gwangwa, M. V., Joubert, A. M., and Visagie, M. H. (2018). Crosstalk between the Warburg effect, redox regulation and autophagy induction in tumourigenesis. *Cellular & Molecular Biology Letters*, 23(1):20.

Gábor, A. and Banga, J. R. (2015). Robust and efficient parameter estimation in dynamic models of biological systems. *BMC Systems Biology*, 9(1):74.

Haeger, A., Krause, M., Wolf, K., and Friedl, P. (2014). Cell jamming: Collective invasion of mesenchymal tumor cells imposed by tissue confinement. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(8):2386–2395.

Hagios, C., Lochter, A., and Bissell, M. J. (1998). Tissue architecture: the ultimate regulator of epithelial function? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353(1370):857–870.

Hagiwara, M. and Koh, I. (2020). Engineering approaches to control and design the in vitro environment towards the reconstruction of organs. *Development, Growth & Differentiation*, 62(3):158–166.

Han, S. J., Kwon, S., and Kim, K. S. (2021). Challenges of applying multicellular tumor spheroids in preclinical phase. *Cancer Cell International*, 21(1):152.

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discovery*, 12(1):31–46.

Hanahan, D. and Weinberg, R. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, 144(5):646–674.

Hanahan, D. and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell*, 100(1):57–70.

Hardwick, L. J. and Philpott, A. (2014). Nervous decision-making: to divide or differentiate. *Trends in Genetics*, 30(6):254–261.

Hasenauer, J., Jagiella, N., Hross, S., and Theis, F. J. (2015). Data-Driven Modelling of Biological Multi-Scale Processes. *Journal of Coupled Systems and Multiscale Dynamics*, 3(2):101–121.

Hashemzadeh, S., Shahmorad, S., Rafii-Tabar, H., and Omidi, Y. (2020). Computational modeling to determine key regulators of hypoxia effects on the lactate production in the glycolysis pathway. *Scientific Reports*, 10(1):9163.

Hassanpour, S. H. and Dehghani, M. (2017). Review of cancer from perspective of molecular. *Journal of Cancer Research and Practice*, 4(4):127–129.

Hatzikirou, H., Basanta, D., Simon, M., Schaller, K., and Deutsch, A. (2012). 'Go or Grow': the key to the emergence of invasion in tumour progression? *Mathematical Medicine and Biology*, 29(1):49–65.

Hay, N. (2016). Reprogramming glucose metabolism in cancer: can it be exploited for cancer therapy? *Nature Reviews Cancer*, 16(10):635–649.

Heiland, R. (2020). PhysiCell model with SBML intracellular model. Language: en.

Heiland, R., Mishler, D., Zhang, T., Bower, E., and Macklin, P. (2019). xml2jupyter: Mapping parameters between XML and Jupyter widgets. *Journal of Open Source Software*, 4(39):1408.

Heldin, C.-H. and Westermark, B. (1999). Mechanism of Action and In Vivo Role of Platelet-Derived Growth Factor. *Physiological Reviews*, 79(4):1283–1316.

Helmlinger, G., Netti, P. A., Lichtenbeld, H. C., Melder, R. J., and Jain, R. K. (1997). Solid stress inhibits the growth of multicellular tumor spheroids. *Nature Biotechnology*, 15(8):778–783.

Hirschhaeuser, F., Menne, H., Dittfeld, C., West, J., Mueller-Klieser, W., and Kunz-Schughart, L. A. (2010). Multicellular tumor spheroids: An underestimated tool is catching up again. *Journal of Biotechnology*, 148(1):3–15.

Hoehme, S. and Drasdo, D. (2010). A cell-based simulation software for multi-cellular systems. *Bioinformatics*, 26(20):2641–2642.

Hofmarcher, T., Lindgren, P., Wilking, N., and Jönsson, B. (2020). The cost of cancer in Europe 2018. *European Journal of Cancer*, 129:41–49.

Hyun, A. Z. and Macklin, P. (2013). Improved patient-specific calibration for agent-based cancer modeling. *Journal of Theoretical Biology*, 317:422–424.

Jacob, M., Chang, L., and Pure, E. (2012). Fibroblast Activation Protein in Remodeling Tissues. *Current Molecular Medicine*, 12(10):1220–1243.

Jacquet, P. and Stéphanou, A. (2023). A reduced model of cell metabolism to revisit the glycolysis-OXPHOS relationship in the deregulated tumor microenvironment. *Journal of Theoretical Biology*, 562:111434.

Jafari Nivlouei, S., Soltani, M., Shirani, E., Salimpour, M. R., Travasso, R., and Carvalho, J. (2022). A multiscale cell-based model of tumor growth for chemotherapy assessment and tumor-targeted therapy through a 3D computational approach. *Cell Proliferation*, 55(3).

Jagiella, N., Müller, B., Müller, M., Vignon-Clementel, I. E., and Drasdo, D. (2016). Inferring Growth Control Mechanisms in Growing Multi-cellular Spheroids of NSCLC Cells from Spatial-Temporal Image Data. *PLOS Computational Biology*, 12(2):e1004412.

Jamshidi, N. and Palsson, B. O. (2010). Mass Action Stoichiometric Simulation Models: Incorporating Kinetics and Regulation into Stoichiometric Models. *Biophysical Journal*, 98(2):175–185.

Janmey, P. A. and McCulloch, C. A. (2007). Cell Mechanics: Integrating Cell Responses to Mechanical Stimuli. *Annual Review of Biomedical Engineering*, 9(1):1–34.

Jeggo, P. A., Pearl, L. H., and Carr, A. M. (2016). DNA repair, genome stability and cancer: a historical perspective. *Nature Reviews Cancer*, 16(1):35–42.

Jeon, J., Quaranta, V., and Cummings, P. T. (2010). An Off-Lattice Hybrid Discrete-Continuum Model of Tumor Growth and Invasion. *Biophysical Journal*, 98(1):37–47.

Jessen, K., Mirsky, R., and Arthur-Farraj, P. (2015). The Role of Cell Plasticity in Tissue Repair: Adaptive Cellular Reprogramming. *Developmental Cell*, 34(6):613–620.

Jia, D., Lu, M., Jung, K. H., Park, J. H., Yu, L., Onuchic, J. N., Kaipparettu, B. A., and Levine, H. (2019). Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. *Proceedings of the National Academy of Sciences*, 116(9):3909–3918.

Jiang, Y., Pjesivac-Grbovic, J., Cantrell, C., and Freyer, J. P. (2005). A Multiscale Model for Avascular Tumor Growth. *Biophysical Journal*, 89(6):3884–3894.

Johnsen, J. I., Dyberg, C., and Wickström, M. (2019). Neuroblastoma—A Neural Crest Derived Embryonal Malignancy. *Frontiers in Molecular Neuroscience*, 12:9.

Jones, R. G. and Thompson, C. B. (2009). Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes & Development*, 23(5):537–548.

Jubelin, C., Muñoz-Garcia, J., Griscom, L., Cochonneau, D., Ollivier, E., Heymann, M.-F., Vallette, F. M., Oliver, L., and Heymann, D. (2022). Three-dimensional in vitro culture models in oncology research. *Cell & Bioscience*, 12(1):155.

Jögi, A., Vaapil, M., Johansson, M., and Påhlman, S. (2012). Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors. *Upsala Journal of Medical Sciences*, 117(2):217–224.

Jørgensen, A. C. S., Ghosh, A., Sturrock, M., and Shahrezaei, V. (2022). Efficient Bayesian inference for stochastic agent-based models. *PLOS Computational Biology*, 18(10):e1009508.

Júlvez, J. and Oliver, S. G. (2020). A unifying modelling formalism for the integration of stoichiometric and kinetic models. *Journal of The Royal Society Interface*, 17(169):20200341.

Kalluri, R. (2016). The biology and function of fibroblasts in cancer. *Nature Reviews Cancer*, 16(9):582–598.

Kalluri, R. and Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation*, 119(6):1420–1428.

Kalluri, R. and Zeisberg, M. (2006). Fibroblasts in cancer. *Nature Reviews Cancer*, 6(5):392–401.

Kamiyama, K., Iguchi, I., Wang, X., and Imanishi, J. (1998). Effects of PDGF on the migration of rabbit corneal fibroblasts and epithelial cells. *Cornea*, 17(3):315–325.

Kang, S., Kahan, S., McDermott, J., Flann, N., and Shmulevich, I. (2014). *Biocellion* : accelerating computer simulation of multicellular biological system models. *Bioinformatics*, 30(21):3101–3108.

Kansal, A., Torquato, S., Harsh, G., Chiocca, E., and Deisboeck, T. (2000). Simulated Brain Tumor Growth Dynamics Using a Three-Dimensional Cellular Automaton. *Journal of Theoretical Biology*, 203(4):367–382.

Karolak, A., Markov, D. A., McCawley, L. J., and Rejniak, K. A. (2018). Towards personalized computational oncology: from spatial models of tumour spheroids, to organoids, to tissues. *Journal of The Royal Society Interface*, 15(138):20170703.

Katsetos, C. D., Liu, H. M., and Zacks, S. I. (1988). Immunohistochemical and ultrastructural observations on homer wright (neuroblastic) rosettes and the "pale islands" of human cerebellar medulloblastomas. *Human Pathology*, 19(10):1219–1227.

Kazerouni, A. S., Gadde, M., Gardner, A., Hormuth, D. A., Jarrett, A. M., Johnson, K. E., Lima, E. A. F., Lorenzo, G., Phillips, C., Brock, A., and Yankeelov, T. E. (2020). Integrating Quantitative Assays with Biologically Based Mathematical Modeling for Predictive Oncology. *iScience*, 23(12):101807.

Kempf, H., Bleicher, M., and Meyer-Hermann, M. (2010). Spatio-temporal cell dynamics in tumour spheroid irradiation. *The European Physical Journal D*, 60(1):177–193.

Kennedy, M. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13.

Khalil, A. A. and Friedl, P. (2010). Determinants of leader cells in collective cell migration. *Integrative Biology*, 2(11-12):568.

Kholodenko, I. V., Kalinovsky, D. V., Doronin, I. I., Deyev, S. M., and Kholodenko, R. V. (2018). Neuroblastoma Origin and Therapeutic Targets for Immunotherapy. *Journal of Immunology Research*, 2018:1–25.

Kim, S.-H., Turnbull, J., and Guimond, S. (2011). Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. *Journal of Endocrinology*, 209(2):139–151.

Kim, Y., Kang, H., and Lawler, S. (2014). The Role of the miR-451-AMPK Signaling Pathway in Regulation of Cell Migration and Proliferation in Glioblastoma. In Eladdadi, A., Kim, P., and Mallet, D., editors, *Mathematical Models of Tumor-Immune System Dynamics*, volume 107, pages 125–155. Springer New York, New York, NY. Series Title: Springer Proceedings in Mathematics & Statistics.

Klipp, E. and Liebermeister, W. (2006). Mathematical modeling of intracellular signaling pathways. *BMC Neuroscience*, 7(S1):S10.

Koene, R. A., Tijms, B., van Hees, P., Postma, F., de Ridder, A., Ramakers, G. J. A., van Pelt, J., and van Ooyen, A. (2009). NETMORPH: A Framework for the Stochastic Generation of Large Scale Neuronal Networks With Realistic Neuron Morphologies. *Neuroinformatics*, 7(3):195–210.

Kontomanolis, E. N., Koutras, A., Syllaios, A., Schizas, D., Mastoraki, A., Garmpis, N., Diakosavvas, M., Angelou, K., Tsatsaris, G., Pagkalos, A., Ntounis, T., and Fasoulakis, Z. (2020). Role of Oncogenes and Tumor-suppressor Genes in Carcinogenesis: A Review. *Anticancer Research*, 40(11):6009–6015.

Koppes, A. N., Seggio, A. M., and Thompson, D. M. (2011). Neurite outgrowth is significantly increased by the simultaneous presentation of Schwann cells and moderate exogenous electric fields. *Journal of Neural Engineering*, 8(4):046023.

Krakauer, D. C., Collins, J. P., Erwin, D., Flack, J. C., Fontana, W., Laubichler, M. D., Prohaska, S. J., West, G. B., and Stadler, P. F. (2011). The challenges and scope of theoretical biology. *Journal of Theoretical Biology*, 276(1):269–276.

Krakhmal, N. V., Zavyalova, M. V., Denisov, E. V., Vtorushin, S. V., and Perelmuter, V. M. (2015). Cancer Invasion: Patterns and Mechanisms. *Acta Naturae*, 7(2):17–28.

Krall, A. S. and Christofk, H. R. (2015). Rethinking glutamine addiction. *Nature Cell Biology*, 17(12):1515–1517.

Kramer, N., Walzl, A., Unger, C., Rosner, M., Krupitza, G., Hengstschläger, M., and Dolznig, H. (2013). In vitro cell migration and invasion assays. *Mutation Research/Reviews in Mutation Research*, 752(1):10–24.

Kunz-Schughart, L. A., Kreutz, M., and Knuechel, R. (1998). Multicellular spheroids: a three-dimensional *in vitro* culture system to study tumour biology. *International Journal of Experimental Pathology*, 79(1):1–23.

Laconi, E., Doratiotto, S., and Vineis, P. (2008). The microenvironments of multistage carcinogenesis. *Seminars in Cancer Biology*, 18(5):322–329.

Ladoux, B. and Mège, R.-M. (2017). Mechanobiology of collective cell behaviours. *Nature Reviews Molecular Cell Biology*, 18(12):743–757.

Lagziel, S., Lee, W. D., and Shlomi, T. (2019). Studying metabolic flux adaptations in cancer through integrated experimental-computational approaches. *BMC Biology*, 17(1):51.

Laird, A. K. (1964). Dynamics of Tumor Growth. *British Journal of Cancer*, 18(3):490–502.

Lam, W. A., Cao, L., Umesh, V., Keung, A. J., Sen, S., and Kumar, S. (2010). Extracellular matrix rigidity modulates neuroblastoma cell differentiation and N-myc expression. *Molecular Cancer*, 9(1):35.

Lardon, L. A., Merkey, B. V., Martins, S., Dötsch, A., Picioreanu, C., Kreft, J.-U., and Smets, B. F. (2011). iDynoMiCS: next-generation individual-based modelling of biofilms: iDynoMiCS for Biofilm Modelling. *Environmental Microbiology*, 13(9):2416–2434.

Lei, B., Kirk, T. Q., Bhattacharya, A., Pati, D., Qian, X., Arroyave, R., and Mallick, B. K. (2021). Bayesian optimization with adaptive surrogate models for automated experimental design. *npj Computational Materials*, 7(1):194.

LeSavage, B. L., Suhar, R. A., Broguiere, N., Lutolf, M. P., and Heilshorn, S. C. (2022). Next-generation cancer organoids. *Nature Materials*, 21(2):143–159.

Letort, G., Montagud, A., Stoll, G., Heiland, R., Barillot, E., Macklin, P., Zinovyev, A., and Calzone, L. (2019). PhysiBoSS: a multi-scale agent-based modelling framework integrating physical dimension and cell signalling. *Bioinformatics*, 35(7):1188–1196.

Li, B. and Wang, J. H.-C. (2011). Fibroblasts and myofibroblasts in wound healing: Force generation and measurement. *Journal of Tissue Viability*, 20(4):108–120.

Li, J. F. and Lowengrub, J. (2014). The effects of cell compressibility, motility and contact inhibition on the growth of tumor cell clusters using the Cellular Potts Model. *Journal of Theoretical Biology*, 343:79–91.

Li, X. J., Valadez, A. V., Zuo, P., and Nie, Z. (2012). Microfluidic 3D cell culture: potential application for tissue-based bioassays. *Bioanalysis*, 4(12):1509–1525.

Liberti, M. V. and Locasale, J. W. (2016). The Warburg Effect: How Does it Benefit Cancer Cells? *Trends in Biochemical Sciences*, 41(3):211–218.

Lima, E. A. B. F., Faghihi, D., Philley, R., Yang, J., Virostko, J., Phillips, C. M., and Yankeelov, T. E. (2021). Bayesian calibration of a stochastic, multiscale agent-based model for predicting in vitro tumor growth. *PLOS Computational Biology*, 17(11):e1008845.

Lintusaari, J., Vuollekoski, H., Kangasraasio, A., Skytén, K., Jarvenpaa, M., Marttinen, P., Gutmann, M. U., Vehtari, A., Corander, J., and Kaski, S. (2018). ELFI: Engine for likelihood-free inference. *Journal of Machine Learning Research*, 19(16):1–7.

Liu, J., Zhang, X., Cheng, Y., and Cao, X. (2021). Dendritic cell migration in inflammation and immunity. *Cellular & Molecular Immunology*, 18(11):2461–2471.

Liu, S., Tian, Y., Chlenski, A., Yang, Q., Salwen, H. R., and Cohn, S. L. (2005). 'Cross-talk' between Schwannian stroma and neuroblasts promotes neuroblastoma tumor differentiation and inhibits angiogenesis. *Cancer Letters*, 228(1-2):125–131.

Lo, C.-M., Wang, H.-B., Dembo, M., and Wang, Y.-l. (2000). Cell Movement Is Guided by the Rigidity of the Substrate. *Biophysical Journal*, 79(1):144–152.

Loessner, D., Flegg, J. A., Byrne, H. M., Clements, J. A., and Hutmacher, D. W. (2013). Growth of confined cancer spheroids: a combined experimental and mathematical modelling approach. *Integrative Biology*, 5(3):597.

Logan, J. A., Kelly, M. E., Ayers, D., Shipillis, N., Baier, G., and Day, P. J. (2010). Systems biology and modeling in neuroblastoma: practicalities and perspectives. *Expert Review of Molecular Diagnostics*, 10(2):131–145.

Lu, P., Weaver, V. M., and Werb, Z. (2012). The extracellular matrix: A dynamic niche in cancer progression. *Journal of Cell Biology*, 196(4):395–406.

Lugano, R., Ramachandran, M., and Dimberg, A. (2020). Tumor angiogenesis: causes, consequences, challenges and opportunities. *Cellular and Molecular Life Sciences*, 77(9):1745–1770.

Lunt, S. Y. and Vander Heiden, M. G. (2011). Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation. *Annual Review of Cell and Developmental Biology*, 27(1):441–464.

Lv, D., Hu, Z., Lu, L., Lu, H., and Xu, X. (2017). Three-dimensional cell culture: A powerful tool in tumor research and drug discovery (Review). *Oncology Letters*.

Lämmermann, T., Bader, B. L., Monkley, S. J., Worbs, T., Wedlich-Söldner, R., Hirsch, K., Keller, M., Förster, R., Critchley, D. R., Fässler, R., and Sixt, M. (2008). Rapid leukocyte migration by integrin-independent flowing and squeezing. *Nature*, 453(7191):51–55.

Lämmermann, T. and Sixt, M. (2009). Mechanical modes of 'amoeboid' cell migration. *Current Opinion in Cell Biology*, 21(5):636–644.

Macklin, P., Edgerton, M. E., Thompson, A. M., and Cristini, V. (2012). Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression. *Journal of Theoretical Biology*, 301:122–140.

Macklin, P., Mumenthaler, S., and Lowengrub, J. (2013). Modeling Multiscale Necrotic and Calcified Tissue Biomechanics in Cancer Patients: Application to Ductal Carcinoma In Situ (DCIS). In Gefen, A., editor, *Multiscale Computer Modeling in Biomechanics and Biomedical Engineering*, volume 14, pages 349–380. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Studies in Mechanobiology, Tissue Engineering and Biomaterials.

Maclaren, O. J. and Nicholson, R. (2019). What can be estimated? Identifiability, estimability, causal inference and ill-posed inverse problems. Publisher: arXiv Version Number: 4.

Macnamara, C. K. (2021). Biomechanical modelling of cancer: Agent-based force-based models of solid tumours within the context of the tumour microenvironment. *Computational and Systems Oncology*, 1(2).

Macnamara, C. K., Caiazzo, A., Ramis-Conde, I., and Chaplain, M. A. (2020). Computational modelling and simulation of cancer growth and migration within a 3D heterogeneous tissue: The effects of fibre and vascular structure. *Journal of Computational Science*, 40:101067.

Maris, J. M. (2010). Recent Advances in Neuroblastoma. *New England Journal of Medicine*, 362(23):2202–2211.

Mark, C., Grundy, T. J., Strissel, P. L., Böhringer, D., Grummel, N., Gerum, R., Steinwachs, J., Hack, C. C., Beckmann, M. W., Eckstein, M., Strick, R., O'Neill, G. M., and Fabry, B. (2020). Collective forces of tumor spheroids in three-dimensional biopolymer networks. *eLife*, 9:e51912.

Mark, C., Metzner, C., Lautscham, L., Strissel, P. L., Strick, R., and Fabry, B. (2018). Bayesian model selection for complex dynamic systems. *Nature Communications*, 9(1):1803.

Martinez-Outschoorn, U. E., Peiris-Pagés, M., Pestell, R. G., Sotgia, F., and Lisanti, M. P. (2017). Cancer metabolism: a therapeutic perspective. *Nature Reviews Clinical Oncology*, 14(1):11–31.

Martins, M., Warren, S., Kimberley, C., Margineanu, A., Peschard, P., McCarthy, A., Yeo, M., Marshall, C. J., Dunsby, C., French, P. M. W., and Katan, M. (2012). Activity of PLCε contributes to chemotaxis of fibroblasts towards PDGF. *Journal of Cell Science*, 125(23):5758–5769.

Marusyk, A. and Polyak, K. (2010). Tumor heterogeneity: Causes and consequences. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1805(1):105–117.

Mascheroni, P., Savvopoulos, S., Alfonso, J. C. L., Meyer-Hermann, M., and Hatzikirou, H. (2021). Improving personalized tumor growth predictions using a Bayesian combination of mechanistic modeling and machine learning. *Communications Medicine*, 1(1):19.

Massagué, J. and Obenauf, A. C. (2016). Metastatic colonization by circulating tumour cells. *Nature*, 529(7586):298–306.

Masserdotti, C. (2006). Architectural patterns in cytology: correlation with histology. *Veterinary Clinical Pathology*, 35(4):388–396.

Mathias, S., Coulier, A., Bouchnita, A., and Hellander, A. (2020). Impact of Force Function Formulations on the Numerical Simulation of Centre-Based Models. *Bulletin of Mathematical Biology*, 82(10):132.

Mathias, S., Coulier, A., and Hellander, A. (2022). CBMOS: a GPU-enabled Python framework for the numerical study of center-based models. *BMC Bioinformatics*, 23(1):55.

Matsumura, N., Wang, Y., and Nakazato, Y. (2014). Coexpression of glial and neuronal markers in the neurocytic rosettes of rosette-forming glioneuronal tumors. *Brain Tumor Pathology*, 31(1):17–22.

Mayor, R. and Etienne-Manneville, S. (2016). The front and rear of collective cell migration. *Nature Reviews Molecular Cell Biology*, 17(2):97–109.

McEvoy, E., Han, Y. L., Guo, M., and Shenoy, V. B. (2020). Gap junctions amplify spatial variations in cell volume in proliferating tumor spheroids. *Nature Communications*, 11(1):6148.

McEvoy, E., Sneh, T., Moeendarbary, E., Javanmardi, Y., Efimova, N., Yang, C., Marino-Bravante, G. E., Chen, X., Escribano, J., Spill, F., Garcia-Aznar, J. M., Weeraratna, A. T., Svitkina, T. M., Kamm, R. D., and Shenoy, V. B. (2022). Feedback between mechanosensitive signaling and active forces governs endothelial junction integrity. *Nature Communications*, 13(1):7089.

McIntosh, J. R., Molodtsov, M. I., and Ataullakhanov, F. I. (2012). Biophysics of mitosis. *Quarterly Reviews of Biophysics*, 45(2):147–207.

Mehta, P., Rahman, Z., ten Dijke, P., and Boukany, P. E. (2022). Microfluidics meets 3D cancer cell migration. *Trends in Cancer*, 8(8):683–697.

Menzel, A. M. and Ohta, T. (2012). Soft deformable self-propelled particles. *EPL (Europhysics Letters)*, 99(5):58001.

Merino-Casallo, F., Gomez-Benito, M. J., Hervas-Raluy, S., and Garcia-Aznar, J. M. (2022a). Unravelling cell migration: defining movement from the cell surface. *Cell Adhesion & Migration*, 16(1):25–64.

Merino-Casallo, F., Gomez-Benito, M. J., Juste-Lanas, Y., Martinez-Cantin, R., and Garcia-Aznar, J. M. (2018). Integration of in vitro and in silico Models Using Bayesian Optimization With an Application to Stochastic Modeling of Mesenchymal 3D Cell Migration. *Frontiers in Physiology*, 9:1246.

Merino-Casallo, F., Gomez-Benito, M. J., Martinez-Cantin, R., and Garcia-Aznar, J. M. (2022b). A mechanistic protrusive-based model for 3D cell migration. *European Journal of Cell Biology*, 101(3):151255.

Metzcar, J., Duggan, B. S., Fischer, B., Murphy, M., Heiland, R., and Macklin, P. (2022). A model of multicellular communication mediated through extracellular matrix microstructure. preprint, Systems Biology.

Metzcar, J., Wang, Y., Heiland, R., and Macklin, P. (2019). A Review of Cell-Based Computational Modeling in Cancer Biology. *JCO Clinical Cancer Informatics*, (3):1–13.

Metzner, C., Hörsch, F., Mark, C., Czerwinski, T., Winterl, A., Voskens, C., and Fabry, B. (2021). Detecting long-range interactions between migrating cells. *Scientific Reports*, 11(1):15031.

Miller, K. E. and Suter, D. M. (2018). An Integrated Cytoskeletal Model of Neurite Outgrowth. *Frontiers in Cellular Neuroscience*, 12:447.

Mirams, G. R., Arthurs, C. J., Bernabeu, M. O., Bordas, R., Cooper, J., Corrias, A., Davit, Y., Dunn, S.-J., Fletcher, A. G., Harvey, D. G., Marsh, M. E., Osborne, J. M., Pathmanathan, P., Pitt-Francis, J., Southern, J., Zemzemi, N., and Gavaghan, D. J. (2013). Chaste: An Open Source C++ Library for Computational Physiology and Biology. *PLoS Computational Biology*, 9(3):e1002970.

Miron-Mendoza, M., Lin, X., Ma, L., Ririe, P., and Petroll, W. M. (2012). Individual versus collective fibroblast spreading and migration: Regulation by matrix composition in 3D culture. *Experimental Eye Research*, 99:36–44.

Montagud, A., Ponce-de Leon, M., and Valencia, A. (2021). Systems biology at the giga-scale: Large multiscale models of complex, heterogeneous multicellular systems. *Current Opinion in Systems Biology*, 28:100385.

Mora, J. and Gerald, W. L. (2004). Origin of neuroblastic tumors: clues for future therapeutics. *Expert Review of Molecular Diagnostics*, 4(3):293–302.

Moreno-Arotzena, O., Borau, C., Movilla, N., Vicente-Manzanares, M., and García-Aznar, J. M. (2015). Fibroblast Migration in 3D is Controlled by Haptotaxis in a Non-muscle Myosin II-Dependent Manner. *Annals of Biomedical Engineering*, 43(12):3025–3039.

Moreno-Arotzena, O., Mendoza, G., Cóndor, M., Rüberg, T., and García-Aznar, J. M. (2014). Inducing chemotactic and haptotactic cues in microfluidic devices for three-dimensional *in vitro* assays. *Biomicrofluidics*, 8(6):064122.

Morgan-Parkes, J. H. (1995). Metastases: mechanisms, pathways, and cascades. *American Journal of Roentgenology*, 164(5):1075–1082.

Mortimer, D., Feldner, J., Vaughan, T., Vetter, I., Pujic, Z., Rosoff, W. J., Burrage, K., Dayan, P., Richards, L. J., and Goodhill, G. J. (2009). A Bayesian model predicts the response of axons to molecular gradients. *Proceedings of the National Academy of Sciences*, 106(25):10296–10301.

Mortimer, D., Pujic, Z., Vaughan, T., Thompson, A. W., Feldner, J., Vetter, I., and Goodhill, G. J. (2010). Axon guidance by growth-rate modulation. *Proceedings of the National Academy of Sciences*, 107(11):5202–5207.

Moscona, A. A. (1976). Cell Recognition in Embryonic Morphogenesis and the Problem of Neuronal Specificities. In Barondes, S. H., editor, *Neuronal Recognition*, pages 205–226. Springer US, Boston, MA.

Moss, T. (1983). Evidence for differentiation in medulloblastomas appearing primitive on light microscopy: an ultrastructural study. *Histopathology*, 7(6):919–930.

Movilla, N., Borau, C., Valero, C., and García-Aznar, J. (2018). Degradation of extracellular matrix regulates osteoblast migration: A microfluidic-based study. *Bone*, 107:10–17.

Movilla, N., Gonçalves, I. G., Borau, C., and García-Aznar, J. M. (2023). A novel integrated experimental and computational approach to unravel fibroblast motility in response to chemical gradients in 3D collagen matrices. *Integrative Biology*, page zyad002.

Movilla, N., Valero, C., Borau, C., and García-Aznar, J. M. (2019). Matrix degradation regulates osteoblast protrusion dynamics and individual migration. *Integrative Biology*, 11(11):404–413.

Moya-Garcia, C. R., Okuyama, H., Sadeghi, N., Li, J., Tabrizian, M., and Li-Jessen, N. Y. K. (2022). In vitro models for head and neck cancer: Current status and future perspective. *Frontiers in Oncology*, 12:960340.

Mueller-Klieser, W. (1987). Multicellular spheroids: A review on cellular aggregates in cancer research. *Journal of Cancer Research and Clinical Oncology*, 113(2):101–122.

Muller, E., Bednar, J. A., Diesmann, M., Gewaltig, M.-O., Hines, M., and Davison, A. P. (2015). Python in neuroscience. *Frontiers in Neuroinformatics*, 9.

Murphy, H., Jaafari, H., and Dobrovolny, H. M. (2016). Differences in predictions of ODE models of tumor growth: a cautionary example. *BMC Cancer*, 16(1):163.

Murphy, R. J., Browning, A. P., Gunasingh, G., Haass, N. K., and Simpson, M. J. (2022). Designing and interpreting 4D tumour spheroid experiments. *Communications Biology*, 5(1):91.

Musy, M., Jacquenot, G., Dalmasso, G., De Bruin, R., Neoglez, Müller, J., Pollack, A., Claudi, F., Badger, C., Sol, A., Zhou, Z.-Q., Sullivan, B., Lerner, B., Hrisca, D., Volpatto, D., Evan, Mkerrinrapid, Schlömer, N., RichardScottOZ, RobinEnjalbert, Lu, X., and Schneider, O. (2022). marcomusy/vedo: 2022.4.1.

Nagrath, S., Sequist, L. V., Maheswaran, S., Bell, D. W., Irimia, D., Ulkus, L., Smith, M. R., Kwak, E. L., Digumarthy, S., Muzikansky, A., Ryan, P., Balis, U. J., Tompkins, R. G., Haber, D. A., and Toner, M. (2007). Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature*, 450(7173):1235–1239.

Nelson, C. M. and Bissell, M. J. (2006). Of Extracellular Matrix, Scaffolds, and Signaling: Tissue Architecture Regulates Development, Homeostasis, and Cancer. *Annual Review of Cell and Developmental Biology*, 22(1):287–309.

Newland, C. P., Maier, H. R., Zecchin, A. C., Newman, J. P., and van Delden, H. (2018). Multi-objective optimisation framework for calibration of Cellular Automata land-use models. *Environmental Modelling & Software*, 100:175–200.

Ngo, D. C., Ververis, K., Tortorella, S. M., and Karagiannis, T. C. (2015). Introduction to the molecular basis of cancer metabolism and the Warburg effect. *Molecular Biology Reports*, 42(4):819–823.

Noble, R., Burri, D., Le Sueur, C., Lemant, J., Viossat, Y., Kather, J. N., and Beerenwinkel, N. (2021). Spatial structure governs the mode of tumour evolution. *Nature Ecology & Evolution*, 6(2):207–217.

Nogueira, F. (2014). Bayesian Optimization: Open source constrained global optimization tool for Python.

Norton, L. (1988). A Gompertzian model of human breast cancer growth. *Cancer Research*, 48(24 Pt 1):7067–7071.

Nyga, A., Cheema, U., and Loizidou, M. (2011). 3D tumour models: novel in vitro approaches to cancer studies. *Journal of Cell Communication and Signaling*, 5(3):239–248.

Odell, G., Oster, G., Alberch, P., and Burnside, B. (1981). The mechanical basis of morphogenesis. *Developmental Biology*, 85(2):446–462.

Oraiopoulou, M.-E., Tzamali, E., Tzedakis, G., Liapis, E., Zacharakis, G., Vakis, A., Papamatheakis, J., and Sakkalis, V. (2018). Integrating in vitro experiments with in silico approaches for Glioblastoma invasion: the role of cell-to-cell adhesion heterogeneity. *Scientific Reports*, 8(1):16200.

Orgaz, J. L., Pandya, P., Dalmeida, R., Karagiannis, P., Sanchez-Laorden, B., Viros, A., Albrengues, J., Nestle, F. O., Ridley, A. J., Gaggioli, C., Marais, R., Karagiannis, S. N., and Sanz-Moreno, V. (2014). Diverse matrix metalloproteinase functions regulate cancer amoeboid migration. *Nature Communications*, 5(1):4255.

Otte, J., Dyberg, C., Pepich, A., and Johnsen, J. I. (2021). MYCN Function in Neuroblastoma Development. *Frontiers in Oncology*, 10:624079.

Ozcelikkale, A., Moon, H., Linnes, M., and Han, B. (2017). *In vitro* microfluidic models of tumor microenvironment to screen transport of drugs and nanoparticles. *WIREs Nanomedicine and Nanobiotechnology*, 9(5).

Ozik, J., Collier, N., Wozniak, J. M., Macal, C., Cockrell, C., Friedman, S. H., Ghaffarizadeh, A., Heiland, R., An, G., and Macklin, P. (2018). High-throughput cancer hypothesis testing with an integrated PhysiCell-EMEWS workflow. *BMC Bioinformatics*, 19(S18):483.

P., B. (2013). Wound healing and the role of fibroblasts. *Journal of Wound Care*, 22(8):407–412.

Pahlman, S. and Hoehner, J. C. (1996). Neurotrophin receptors, tumor progression and tumor maturation. *Molecular Medicine Today*, 2(10):432–438.

Palm, W. (2021). Metabolic plasticity allows cancer cells to thrive under nutrient starvation. *Proceedings of the National Academy of Sciences*, 118(14):e2102057118.

Palsson, E. and Othmer, H. G. (2000). A model for individual and collective cell movement in *Dictyostelium discoideum*. *Proceedings of the National Academy of Sciences*, 97(19):10448–10453.

Pampaloni, F., Reynaud, E. G., and Stelzer, E. H. K. (2007). The third dimension bridges the gap between cell culture and live tissue. *Nature Reviews Molecular Cell Biology*, 8(10):839–845.

Paszek, M. J., Zahir, N., Johnson, K. R., Lakins, J. N., Rozenberg, G. I., Gefen, A., Reinhart-King, C. A., Margulies, S. S., Dembo, M., Boettiger, D., Hammer, D. A., and Weaver, V. M. (2005). Tensional homeostasis and the malignant phenotype. *Cancer Cell*, 8(3):241–254.

Patel, A. A., Gawlinski, E. T., Lemieux, S. K., and Gatenby, R. A. (2001). A Cellular Automaton Model of Early Tumor Growth and Invasion: The Effects of Native Tissue Vascularity and Increased Anaerobic Tumor Metabolism. *Journal of Theoretical Biology*, 213(3):315–331.

Pathak, A. and Kumar, S. (2011). Biophysical regulation of tumor cell invasion: moving beyond matrix stiffness. *Integrative Biology*, 3(4):267.

Pavel, M., Renna, M., Park, S. J., Menzies, F. M., Ricketts, T., Füllgrabe, J., Ashkenazi, A., Frake, R. A., Lombarte, A. C., Bento, C. F., Franze, K., and Rubinsztein, D. C. (2018). Contact inhibition controls cell survival and proliferation via YAP/TAZ-autophagy axis. *Nature Communications*, 9(1):2961.

Pavlides, S., Whitaker-Menezes, D., Castello-Cros, R., Flomenberg, N., Witkiewicz, A. K., Frank, P. G., Casimiro, M. C., Wang, C., Fortina, P., Addya, S., Pestell, R. G., Martinez-Outschoorn, U. E., Sotgia, F., and Lisanti, M. P. (2009). The reverse Warburg effect: Aerobic glycolysis in cancer associated fibroblasts and the tumor stroma. *Cell Cycle*, 8(23):3984–4001.

Payne, S. R. and Kemp, C. J. (2005). Tumor suppressor genetics. *Carcinogenesis*, 26(12):2031–2045.

Paňková, K., Rösel, D., Novotný, M., and Brábek, J. (2010). The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cellular and Molecular Life Sciences*, 67(1):63–71.

Pedersen, J. A. and Swartz, M. A. (2005). Mechanobiology in the Third Dimension. *Annals of Biomedical Engineering*, 33(11):1469–1490.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, (2012). Scikit-learn: Machine Learning in Python. Publisher: arXiv Version Number: 4.

Persi, E., Duran-Frigola, M., Damaghi, M., Roush, W. R., Aloy, P., Cleveland, J. L., Gillies, R. J., and Ruppin, E. (2018). Systems analysis of intracellular pH vulnerabilities for cancer therapy. *Nature Communications*, 9(1):2997.

Persson, C. U., von Stedingk, K., Bexell, D., Merselius, M., Braekeveldt, N., Gisselsson, D., Arsenian-Henriksson, M., Påhlman, S., and Wigerup, C. (2017). Neuroblastoma patient-derived xenograft cells cultured in stem-cell promoting medium retain tumorigenic and metastatic capacities but differentiate in serum. *Scientific Reports*, 7(1):10274.

Petrie, R. J. and Yamada, K. M. (2015). Fibroblasts Lead the Way: A Unified View of 3D Cell Motility. *Trends in Cell Biology*, 25(11):666–674.

Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001). Cooperation and Competition in the Evolution of ATP-Producing Pathways. *Science*, 292(5516):504–507.

Pickup, M. W., Mouw, J. K., and Weaver, V. M. (2014). The extracellular matrix modulates the hallmarks of cancer. *EMBO reports*, 15(12):1243–1253.

Pijuan, J., Barceló, C., Moreno, D. F., Maiques, O., Sisó, P., Marti, R. M., Macià, A., and Panosa, A. (2019). In vitro Cell Migration, Invasion, and Adhesion Assays: From Cell Imaging to Data Analysis. *Frontiers in Cell and Developmental Biology*, 7:107.

Piotrowska, M. J. and Angus, S. D. (2009). A quantitative cellular automaton model of in vitro multicellular spheroid tumour growth. *Journal of Theoretical Biology*, 258(2):165–178.

Pizzo, A. M., Kokini, K., Vaughn, L. C., Waisner, B. Z., and Voytik-Harbin, S. L. (2005). Extracellular matrix (ECM) microstructural composition regulates local cell-ECM biomechanics and fundamental fibroblast behavior: a multidimensional perspective. *Journal of Applied Physiology*, 98(5):1909–1921.

Pleyer, J. and Fleck, C. (2023). Agent-based models in cellular systems. *Frontiers in Physics*, 10:968409.

Plou, J., Juste-Lanas, Y., Olivares, V., del Amo, C., Borau, C., and García-Aznar, J. M. (2018). From individual to collective 3D cancer dissemination: roles of collagen concentration and TGF-β. *Scientific Reports*, 8(1):12723.

Polacheck, W. J., Zervantonakis, I. K., and Kamm, R. D. (2013). Tumor cell migration in complex microenvironments. *Cellular and Molecular Life Sciences*, 70(8):1335–1356.

Ponzoni, M., Bachetti, T., Corrias, M. V., Brignole, C., Pastorino, F., Calarco, E., Bensa, V., Giusto, E., Ceccherini, I., and Perri, P. (2022). Recent advances in the developmental origin of neuroblastoma: an overview. *Journal of Experimental & Clinical Cancer Research*, 41(1):92.

Preziosi, L. and Tosin, A. (2009). Multiphase modelling of tumour growth and extracellular matrix interaction: mathematical tools and applications. *Journal of Mathematical Biology*, 58(4-5):625–656.

Purcell, E. M. (1977). Life at low Reynolds number. *American Journal of Physics*, 45(1):3–11.

Pérez-García, V. M., Fitzpatrick, S., Pérez-Romasanta, L. A., Pesic, M., Schucht, P., Arana, E., and Sánchez-Gómez, P. (2016). Applied mathematics and nonlinear sciences in the war on cancer. *Applied Mathematics and Nonlinear Sciences*, 1(2):423–436.

Pérez-Rodríguez, S., Tomás-González, E., and García-Aznar, J. (2018). 3D Cell Migration Studies for Chemotaxis on Microfluidic-Based Chips: A Comparison between Cardiac and Dermal Fibroblasts. *Bioengineering*, 5(2):45.

Quaranta, V., Tyson, D. R., Garbett, S. P., Weidow, B., Harris, M. P., and Georgescu, W. (2009). Trait Variability of Cancer Cells Quantified by High-Content Automated Microscopy of Single Cells. In *Methods in Enzymology*, volume 467, pages 23–57. Elsevier.

Ram, A., Jalal, S., Jalal, A. S., and Kumar, M. (2010). A Density Based Algorithm for Discovering Density Varied Clusters in Large Spatial Databases. *International Journal of Computer Applications*, 3(6):1–4.

Ramis-Conde, I., Drasdo, D., Anderson, A. R., and Chaplain, M. A. (2008). Modeling the Influence of the E-Cadherin-β-Catenin Pathway in Cancer Cell Invasion: A Multiscale Approach. *Biophysical Journal*, 95(1):155–165.

Rangarajan, R. and Zaman, M. H. (2008). Modeling cell migration in 3D: Status and challenges. *Cell Adhesion & Migration*, 2(2):106–109.

Read, M. N., Alden, K., Timmis, J., and Andrews, P. S. (2018). Strategies for calibrating models of biology. *Briefings in Bioinformatics*.

Rejniak, K. A. and Anderson, A. R. A. (2011). Hybrid models of tumor growth. *WIREs Systems Biology and Medicine*, 3(1):115–125.

Renardy, M., Yi, T.-M., Xiu, D., and Chou, C.-S. (2018). Parameter uncertainty quantification using surrogate models applied to a spatial model of yeast mating polarization. *PLOS Computational Biology*, 14(5):e1006181.

Resendis-Antonio, O., González-Torres, C., Jaime-Muñoz, G., Hernandez-Patiño, C. E., and Salgado-Muñoz, C. F. (2015). Modeling metabolism: A window toward a comprehensive interpretation of networks in cancer. *Seminars in Cancer Biology*, 30:79–87.

Rey, R. and García-Aznar, J. M. (2013). A phenomenological approach to modelling collective cell movement in 2D. *Biomechanics and Modeling in Mechanobiology*, 12(6):1089–1100.

Reynolds, B. A. and Weiss, S. (1992). Generation of Neurons and Astrocytes from Isolated Cells of the Adult Mammalian Central Nervous System. *Science*, 255(5052):1707–1710.

Ribatti, D., Tamma, R., and Annese, T. (2020). Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. *Translational Oncology*, 13(6):100773.

Richmond, P., Walker, D., Coakley, S., and Romano, D. (2010). High performance cellular level agent-based simulation with FLAME for the GPU. *Briefings in Bioinformatics*, 11(3):334–347.

Robertson-Tessi, M., Gillies, R. J., Gatenby, R. A., and Anderson, A. R. (2015). Impact of Metabolic Heterogeneity on Tumor Growth, Invasion, and Treatment Outcomes. *Cancer Research*, 75(8):1567–1579.

Rocha, H. L., Godet, I., Kurtoglu, F., Metzcar, J., Konstantinopoulos, K., Bhoyar, S., Gilkes, D. M., and Macklin, P. (2021). A persistent invasive phenotype in post-hypoxic tumor cells is revealed by fate mapping and computational modeling. *iScience*, 24(9):102935.

Rockne, R. C., Hawkins-Daarud, A., Swanson, K. R., Sluka, J. P., Glazier, J. A., Macklin, P., Hormuth, D. A., Jarrett, A. M., Lima, E. A. B. F., Tinsley Oden, J., Biros, G., Yankeelov, T. E., Curtius, K., Al Bakir, I., Wodarz, D., Komarova, N., Aparicio, L., Bordyuh, M., Rabadan, R., Finley, S. D., Enderling, H., Caudell, J., Moros, E. G., Anderson, A. R. A., Gatenby, R. A., Kaznatcheev, A., Jeavons, P., Krishnan, N., Pelesko, J., Wadhwa, R. R., Yoon, N., Nichol, D., Marusyk, A., Hinczewski, M., and Scott, J. G. (2019). The 2019 mathematical oncology roadmap. *Physical Biology*, 16(4):041005.

Rockne, R. C. and Scott, J. G. (2019). Introduction to Mathematical Oncology. *JCO Clinical Cancer Informatics*, (3):1–4.

Roeder, B. A., Kokini, K., Sturgis, J. E., Robinson, J. P., and Voytik-Harbin, S. L. (2002). Tensile Mechanical Properties of Three-Dimensional Type I Collagen Extracellular Matrices With Varied Microstructure. *Journal of Biomechanical Engineering*, 124(2):214–222.

Rousset, N., Sandoval, R. L., Modena, M. M., Hierlemann, A., and Misun, P. M. (2022). Modeling and measuring glucose diffusion and consumption by colorectal cancer spheroids in hanging drops using integrated biosensors. *Microsystems & Nanoengineering*, 8(1):14.

Roy, M. and Finley, S. D. (2019). Metabolic reprogramming dynamics in tumor spheroids: Insights from a multicellular, multiscale model. *PLOS Computational Biology*, 15(6):e1007053.

Rubin, H. (2002). The disparity between human cell senescence in vitro and lifelong replication in vivo. *Nature Biotechnology*, 20(7):675–681.

Sackmann, E. K., Fulton, A. L., and Beebe, D. J. (2014). The present and future role of microfluidics in biomedical research. *Nature*, 507(7491):181–189.

Sahu, A., Blätke, M.-A., Szymański, J. J., and Töpfer, N. (2021). Advances in flux balance analysis by integrating machine learning and mechanism-based models. *Computational and Structural Biotechnology Journal*, 19:4626–4640.

Salvi, A. M. and DeMali, K. A. (2018). Mechanisms linking mechanotransduction and cell metabolism. *Current Opinion in Cell Biology*, 54:114–120.

Sapudom, J. and Pompe, T. (2018). Biomimetic tumor microenvironments based on collagen matrices. *Biomaterials Science*, 6(8):2009–2024.

Sapudom, J., Rubner, S., Martin, S., Kurth, T., Riedel, S., Mierke, C. T., and Pompe, T. (2015). The phenotype of cancer cell invasion controlled by fibril diameter and pore size of 3D collagen networks. *Biomaterials*, 52:367–375.

Sarkar, S., Horn, G., Moulton, K., Oza, A., Byler, S., Kokolus, S., and Longacre, M. (2013). Cancer Development, Progression, and Therapy: An Epigenetic Overview. *International Journal of Molecular Sciences*, 14(10):21087–21113.

Schaller, G. and Meyer-Hermann, M. (2006). Continuum versus discrete model: a comparison for multicellular tumour spheroids. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1843):1443–1464.

Schmerler, S. (2022). elcorto/psweep: 0.9.0.

Schuster, S., Boley, D., Möller, P., Stark, H., and Kaleta, C. (2015). Mathematical models for explaining the Warburg effect: a review focussed on ATP and biomass production. *Biochemical Society Transactions*, 43(6):1187–1194.

Seewaldt, V. (2014). ECM stiffness paves the way for tumor cells. *Nature Medicine*, 20(4):332–333.

Senthilkumar, I., Howley, E., and McEvoy, E. (2022). Thermodynamically-motivated chemo-mechanical models and multicellular simulation to provide new insight into active cell and tumour remodelling. *Experimental Cell Research*, 419(2):113317.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and de Freitas, N. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148–175.

Shamsi, M., Saghafian, M., Dejam, M., and Sanati-Nezhad, A. (2018). Mathematical Modeling of the Function of Warburg Effect in Tumor Microenvironment. *Scientific Reports*, 8(1):8903.

Shan, M., Dai, D., Vudem, A., Varner, J. D., and Stroock, A. D. (2018). Multi-scale computational study of the Warburg effect, reverse Warburg effect and glutamine addiction in solid tumors. *PLOS Computational Biology*, 14(12):e1006584.

Shatkin, G., Yeoman, B., Birmingham, K., Katira, P., and Engler, A. J. (2020). Computational models of migration modes improve our understanding of metastasis. *APL Bioengineering*, 4(4):041505.

Shimada, H., Ambros, I. M., Dehner, L. P., Hata, J., Joshi, V. V., and Roald, B. (1999a). Terminology and morphologic criteria of neuroblastic tumors: recommendations by the International Neuroblastoma Pathology Committee. *Cancer*, 86(2):349–363.

Shimada, H., Ambros, I. M., Dehner, L. P., Hata, J., Joshi, V. V., Roald, B., Stram, D. O., Gerbing, R. B., Lukens, J. N., Matthay, K. K., and Castleberry, R. P. (1999b). The International Neuroblastoma Pathology Classification (the Shimada system). *Cancer*, 86(2):364–372.

Shin, Y., Han, S., Jeon, J. S., Yamamoto, K., Zervantonakis, I. K., Sudo, R., Kamm, R. D., and Chung, S. (2012). Microfluidic assay for simultaneous culture of multiple cell types on surfaces or within hydrogels. *Nature Protocols*, 7(7):1247–1259.

Shortt, J. and Johnstone, R. W. (2012). Oncogenes in Cell Survival and Cell Death. *Cold Spring Harbor Perspectives in Biology*, 4(12):a009829–a009829.

Shuttleworth, R. and Trucu, D. (2020). Cell-Scale Degradation of Peritumoural Extracellular Matrix Fibre Network and Its Role Within Tissue-Scale Cancer Invasion. *Bulletin of Mathematical Biology*, 82(6):65.

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, 72(1):7–33.

Smallbone, K., Gatenby, R. A., Gillies, R. J., Maini, P. K., and Gavaghan, D. J. (2007). Metabolic changes during carcinogenesis: Potential impact on invasiveness. *Journal of Theoretical Biology*, 244(4):703–713.

Somogyi, E. T., Bouteiller, J.-M., Glazier, J. A., König, M., Medley, J. K., Swat, M. H., and Sauro, H. M. (2015). libRoadRunner: a high performance SBML simulation and analysis library. *Bioinformatics*, 31(20):3315–3321.

Stalidzans, E., Seiman, A., Peebo, K., Komasilovs, V., and Pentjuss, A. (2018). Model-based metabolism design: constraints for kinetic and stoichiometric models. *Biochemical Society Transactions*, 46(2):261–267.

Stander, N. and Craig, K. (2002). On the robustness of a simple domain reduction scheme for simulation-based optimization. *Engineering Computations*, 19(4):431–450.

Starruß, J., de Back, W., Brusch, L., and Deutsch, A. (2014). Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics*, 30(9):1331–1332.

Stéphanou, A., Ballet, P., and Powathil, G. (2020). Hybrid data-based modelling in oncology: successes, challenges and hopes. *Mathematical Modelling of Natural Phenomena*, 15:21.

Sullivan, C. and Kaszynski, A. (2019). PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *Journal of Open Source Software*, 4(37):1450.

Sullivan, W. J., Mullen, P. J., Schmid, E. W., Flores, A., Momcilovic, M., Sharpley, M. S., Jelinek, D., Whiteley, A. E., Maxwell, M. B., Wilde, B. R., Banerjee, U., Coller, H. A., Shackelford, D. B., Braas, D., Ayer, D. E., de Aguiar Vallim, T. Q., Lowry, W. E., and Christofk, H. R. (2018). Extracellular Matrix Remodeling Regulates Glucose Metabolism through TXNIP Destabilization. *Cell*, 175(1):117–132.e21.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3):209–249.

Sunyer, R., Conte, V., Escribano, J., Elosegui-Artola, A., Labernadie, A., Valon, L., Navajas, D., García-Aznar, J. M., Muñoz, J. J., Roca-Cusachs, P., and Trepat, X. (2016). Collective cell durotaxis emerges from long-range intercellular force transmission. *Science*, 353(6304):1157–1161.

Sutherland, R. M., McCredie, J. A., and Inch, W. R. (1971). Growth of multicell spheroids in tissue culture as a model of nodular carcinomas. *Journal of the National Cancer Institute*, 46(1):113–120.

Swat, M. H., Thomas, G. L., Belmonte, J. M., Shirinifard, A., Hmeljak, D., and Glazier, J. A. (2012). Multi-Scale Modeling of Tissues Using CompuCell3D. In *Methods in Cell Biology*, volume 110, pages 325–366. Elsevier.

Swat, M. H., Thomas, G. L., Shirinifard, A., Clendenon, S. G., and Glazier, J. A. (2015). Emergent Stratification in Solid Tumors Selects for Reduced Cohesion of Tumor Cells: A Multi-Cell, Virtual-Tissue Model of Tumor Evolution Using CompuCell3D. *PLOS ONE*, 10(6):e0127972.

Talkenberger, K., Cavalcanti-Adam, E. A., Voss-Böhme, A., and Deutsch, A. (2017). Amoeboid-mesenchymal migration plasticity promotes invasion only in complex heterogeneous microenvironments. *Scientific Reports*, 7(1):9237.

Tambe, D. T., Corey Hardin, C., Angelini, T. E., Rajendran, K., Park, C. Y., Serra-Picamal, X., Zhou, E. H., Zaman, M. H., Butler, J. P., Weitz, D. A., Fredberg, J. J., and Trepat, X. (2011). Collective cell guidance by cooperative intercellular forces. *Nature Materials*, 10(6):469–475.

te Boekhorst, V., Preziosi, L., and Friedl, P. (2016). Plasticity of Cell Migration In Vivo and In Silico. *Annual Review of Cell and Developmental Biology*, 32(1):491–526.

Thomas, P. C., Raghavan, S. R., and Forry, S. P. (2011). Regulating Oxygen Levels in a Microfluidic Device. *Analytical Chemistry*, 83(22):8821–8824.

Thoumine, O., Cardoso, O., and Meister, J.-J. (1999). Changes in the mechanical properties of fibroblasts during spreading: a micromanipulation study. *European Biophysics Journal*, 28(3):222–234.

Tomezak, M., Abbadie, C., Lartigau, E., and Cleri, F. (2016). A biophysical model of cell evolution after cytotoxic treatments: Damage, repair and cell response. *Journal of Theoretical Biology*, 389:146–158.

Transtrum, M. K. and Qiu, P. (2016). Bridging Mechanistic and Phenomenological Models of Complex Biological Systems. *PLOS Computational Biology*, 12(5):e1004915.

Trepat, X., Chen, Z., and Jacobson, K. (2012). Cell Migration. In Terjung, R., editor, *Comprehensive Physiology*, pages 2369–2392. Wiley, 1 edition.

Tripathi, S., Park, J. H., Pudakalakatti, S., Bhattacharya, P. K., Kaipparettu, B. A., and Levine, H. (2022). A mechanistic modeling framework reveals the key principles underlying tumor metabolism. *PLOS Computational Biology*, 18(2):e1009841.

Tschumperlin, D. J. (2013). Fibroblasts and the Ground They Walk On. *Physiology*, 28(6):380–390.

Tuveson, D. and Clevers, H. (2019). Cancer modeling meets human organoid technology. *Science*, 364(6444):952–955.

Tyson, J. J., Chen, K. C., and Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221–231.

Vaghi, C., Rodallec, A., Fanciullino, R., Ciccolini, J., Mochel, J. P., Mastri, M., Poignard, C., Ebos, J. M. L., and Benzekry, S. (2020). Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors. *PLOS Computational Biology*, 16(2):e1007178.

Valastyan, S. and Weinberg, R. (2011). Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell*, 147(2):275–292.

Valero, C., Amaveda, H., Mora, M., and García-Aznar, J. M. (2018). Combined experimental and computational characterization of crosslinked collagen-based hydrogels. *PLOS ONE*, 13(4):e0195820.

van Duinen, V., Trietsch, S. J., Joore, J., Vulto, P., and Hankemeier, T. (2015). Microfluidic 3D cell culture: from tools to tissue models. *Current Opinion in Biotechnology*, 35:118–126.

Van Liedekerke, P., Buttenschön, A., and Drasdo, D. (2018). Off-Lattice Agent-Based Models for Cell and Tumor Growth. In *Numerical Methods and Advanced Simulation in Biomechanics and Biological Processes*, pages 245–267. Elsevier.

Van Liedekerke, P., Palm, M. M., Jagiella, N., and Drasdo, D. (2015). Simulating tissue mechanics with agent-based models: concepts, perspectives and some novel results. *Computational Particle Mechanics*, 2(4):401–444.

van Zijl, F., Krupitza, G., and Mikulits, W. (2011). Initial steps of metastasis: Cell invasion and endothelial transmigration. *Mutation Research/Reviews in Mutation Research*, 728(1-2):23–34.

Vander Heiden, M. G., Cantley, L. C., and Thompson, C. B. (2009). Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science*, 324(5930):1029–1033.

Varella, V., Quintela, B. d. M., Kasztelnik, M., and Viceconti, M. (2022). Effect of particularisation size on the accuracy and efficiency of a multiscale tumours' growth model. *International Journal for Numerical Methods in Biomedical Engineering*.

Vargas, D. A., Gonçalves, I. G., Heck, T., Smeets, B., Lafuente-Gracia, L., Ramon, H., and Van Oosterwyck, H. (2020). Modeling of Mechanosensing Mechanisms Reveals Distinct Cell Migration Modes to Emerge From Combinations of Substrate Stiffness and Adhesion Receptor–Ligand Affinity. *Frontiers in Bioengineering and Biotechnology*, 8:459.

Vendramin, R., Litchfield, K., and Swanton, C. (2021). Cancer evolution: Darwin and beyond. *The EMBO Journal*, 40(18).

Venkatasubramanian, R., Henson, M. A., and Forbes, N. S. (2006). Incorporating energy metabolism into a growth model of multicellular tumor spheroids. *Journal of Theoretical Biology*, 242(2):440–453.

Vernon, I., Liu, J., Goldstein, M., Rowe, J., Topping, J., and Lindsey, K. (2018). Bayesian uncertainty analysis for complex systems biology models: emulation, global parameter searches and evaluation of gene functions. *BMC Systems Biology*, 12(1):1.

Vicente-Manzanares, M. and Horwitz, A. R. (2011). Cell Migration: An Overview. In Wells, C. M. and Parsons, M., editors, *Cell Migration*, volume 769, pages 1–24. Humana Press, Totowa, NJ. Series Title: Methods in Molecular Biology.

Victori, P. and Buffa, F. M. (2018). The many faces of mathematical modelling in oncology. *The British Journal of Radiology*, page 20180856.

Vogelstein, B. and Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine*, 10(8):789–799.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., and Kinzler, K. W. (2013). Cancer Genome Landscapes. *Science*, 339(6127):1546–1558.

Walpole, J., Papin, J. A., and Peirce, S. M. (2013). Multiscale Computational Models of Complex Biological Systems. *Annual Review of Biomedical Engineering*, 15(1):137–154.

Wang, L. L., Suganuma, R., Ikegaki, N., Tang, X., Naranjo, A., McGrady, P., London, W. B., Hogarty, M. D., Gastier-Foster, J. M., Look, A. T., Park, J. R., Maris, J. M., Cohn, S. L., Seeger, R. C., and Shimada, H. (2013). Neuroblastoma of undifferentiated subtype, prognostic significance of prominent nucleolar formation, and MYC/MYCN protein expression: A report from the Children's Oncology Group: Undifferentiated Neuroblastoma. *Cancer*, 119(20):3718–3726.

Wang, Z., Butner, J. D., Kerketta, R., Cristini, V., and Deisboeck, T. S. (2015). Simulating cancer growth with multiscale agent-based modeling. *Seminars in Cancer Biology*, 30:70–78.

Warburg, O. (1956a). On respiratory impairment in cancer cells. *Science (New York, N.Y.)*, 124(3215):269–270.

Warburg, O. (1956b). On the Origin of Cancer Cells. *Science*, 123(3191):309–314.

Warth, A., Cortis, J., Soltermann, A., Meister, M., Budczies, J., Stenzinger, A., Goeppert, B., Thomas, M., Herth, F. J. F., Schirmacher, P., Schnabel, P. A., Hoffmann, H., Dienemann, H., Muley, T., and Weichert, W. (2014). Tumour cell proliferation (Ki-67) in non-small cell lung cancer: a critical reappraisal of its prognostic role. *British Journal of Cancer*, 111(6):1222–1229.

Weinberg, R. A. (2013). *The Biology of Cancer*. W.W. Norton & Company, 0 edition.

Weinhouse, S. (1955). Oxidative Metabolism of Neoplastic Tissues. In *Advances in Cancer Research*, volume 3, pages 269–325. Elsevier.

Weiss, T., Taschner-Mandl, S., Janker, L., Bileck, A., Rifatbegovic, F., Kromp, F., Sorger, H., Kauer, M. O., Frech, C., Windhager, R., Gerner, C., Ambros, P. F., and Ambros, I. M. (2021). Schwann cell plasticity regulates neuroblastic tumor cell differentiation via epidermal growth factor-like protein 8. *Nature Communications*, 12(1):1624.

Westermark, U. K., Wilhelm, M., Frenzel, A., and Henriksson, M. A. (2011). The MYCN oncogene and differentiation in neuroblastoma. *Seminars in Cancer Biology*, 21(4):256–266.

Weston, A. and Harris, C. (2003). Multistage Carcinogenesis. In *Holland-Frei Cancer Medicine*. Hamilton (ON): BC Decker, 6th edition edition.

Wilson, P. G. and Stice, S. S. (2006). Development and differentiation of neural rosettes derived from human embryonic stem cells. *Stem Cell Reviews*, 2(1):67–77.

Wippold, F. J. and Perry, A. (2006). Neuropathology for the neuroradiologist: rosettes and pseudorosettes. *AJNR. American journal of neuroradiology*, 27(3):488–492.

Wise, D. R. and Thompson, C. B. (2010). Glutamine addiction: a new therapeutic target in cancer. *Trends in Biochemical Sciences*, 35(8):427–433.

Wolf, K., Mazo, I., Leung, H., Engelke, K., von Andrian, U. H., Deryugina, E. I., Strongin, A. Y., Bröcker, E.-B., and Friedl, P. (2003). Compensation mechanism in tumor cell migration. *Journal of Cell Biology*, 160(2):267–277.

Wolf, K., te Lindert, M., Krause, M., Alexander, S., te Riet, J., Willis, A. L., Hoffman, R. M., Figdor, C. G., Weiss, S. J., and Friedl, P. (2013). Physical limits of cell migration: Control by ECM space and nuclear deformation and tuning by proteolysis and traction force. *Journal of Cell Biology*, 201(7):1069–1084.

Wright, J. H. (1910). NEUROCYTOMA OR NEUROBLASTOMA, A KIND OF TUMOR NOT GENERALLY RECOGNIZED. *Journal of Experimental Medicine*, 12(4):556–561.

Wu, P.-H., Giri, A., Sun, S. X., and Wirtz, D. (2014). Three-dimensional cell migration does not follow a random walk. *Proceedings of the National Academy of Sciences*, 111(11):3949–3954.

Wu, P.-H., Giri, A., and Wirtz, D. (2015). Statistical analysis of cell migration in 3D using the anisotropic persistent random walk model. *Nature Protocols*, 10(3):517–527.

Yamada, K. M. and Sixt, M. (2019). Mechanisms of 3D cell migration. *Nature Reviews Molecular Cell Biology*, 20(12):738–752.

Yamaguchi, H., Wyckoff, J., and Condeelis, J. (2005). Cell migration in tumors. *Current Opinion in Cell Biology*, 17(5):559–564.

Yang, J., Antin, P., Berx, G., Blanpain, C., Brabletz, T., Bronner, M., Campbell, K., Cano, A., Casanova, J., Christofori, G., Dedhar, S., Derynck, R., Ford, H. L., Fuxe, J., García de Herreros, A., Goodall, G. J., Hadjantonakis, A.-K., Huang, R. Y. J., Kalcheim, C., Kalluri, R., Kang, Y., Khew-Goodall, Y., Levine, H., Liu, J., Longmore, G. D., Mani, S. A., Massagué, J., Mayor, R., McClay, D., Mostov, K. E., Newgreen, D. F., Nieto, M. A., Puisieux, A., Runyan, R., Savagner, P., Stanger, B., Stemmler, M. P., Takahashi, Y., Takeichi, M., Theveneau, E., Thiery, J. P., Thompson, E. W., Weinberg, R. A., Williams, E. D., Xing, J., Zhou, B. P., Sheng, G., and On behalf of the EMT International Association (TEMTIA) (2020). Guidelines and definitions for research on epithelial–mesenchymal transition. *Nature Reviews Molecular Cell Biology*, 21(6):341–352.

Yang, J., Virostko, J., Hormuth, D. A., Liu, J., Brock, A., Kowalski, J., and Yankeelov, T. E. (2021). An experimental-mathematical approach to predict tumor cell growth as a function of glucose availability in breast cancer cell lines. *PLOS ONE*, 16(7):e0240765.

Yang, Y., Jolly, M. K., and Levine, H. (2019). Computational Modeling of Collective Cell Migration: Mechanical and Biochemical Aspects. In La Porta, C. A. M. and Zapperi, S., editors, *Cell Migrations: Causes and Functions*, volume 1146, pages 1–11. Springer International Publishing, Cham. Series Title: Advances in Experimental Medicine and Biology.

Yao, D., Dai, C., and Peng, S. (2011). Mechanism of the Mesenchymal–Epithelial Transition and Its Relationship with Metastatic Tumor Formation. *Molecular Cancer Research*, 9(12):1608–1620.

Yasemi, M. and Jolicoeur, M. (2021). Modelling Cell Metabolism: A Review on Constraint-Based Steady-State and Kinetic Approaches. *Processes*, 9(2):322.

Yoo, H. C., Yu, Y. C., Sung, Y., and Han, J. M. (2020). Glutamine reliance in cell metabolism. *Experimental & Molecular Medicine*, 52(9):1496–1516.

Yu, J. S. and Bagheri, N. (2020). Agent-Based Models Predict Emergent Behavior of Heterogeneous Cell Populations in Dynamic Microenvironments. *Frontiers in Bioengineering and Biotechnology*, 8:249.

Yuan, S., Norgard, R. J., and Stanger, B. Z. (2019). Cellular Plasticity in Cancer. *Cancer Discovery*, 9(7):837–851.

Zaman, M. H., Kamm, R. D., Matsudaira, P., and Lauffenburger, D. A. (2005). Computational Model for Cell Migration in Three-Dimensional Matrices. *Biophysical Journal*, 89(2):1389–1397.

Zanotelli, M. R., Zhang, J., Ortiz, I., Wang, W., Chada, N. C., and Reinhart-King, C. A. (2022). Highly motile cells are metabolically responsive to collagen density. *Proceedings of the National Academy of Sciences*, 119(18):e2114672119.

Zhang, L., Athale, C. A., and Deisboeck, T. S. (2007). Development of a three-dimensional multiscale agent-based tumor model: Simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer. *Journal of Theoretical Biology*, 244(1):96–107.

Zhang, Z. and Nagrath, S. (2013). Microfluidics and cancer: are we there yet? *Biomedical Microdevices*, 15(4):595–609.

Zubler, F. (2009). A framework for modeling the growth and development of neurons and networks. *Frontiers in Computational Neuroscience*, 3.