2016

Collective cell motility in 3-dimensions: dynamics, adhesions, and emergence of heteroger

https://hdl.handle.net/2144/14625

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

COLLECTIVE CELL MOTILITY IN 3-DIMENSIONS: DYNAMICS, ADHESIONS, AND EMERGENCE OF HETEROGENEITY

by

YASHA SHARMA

B.S. University of California, San Diego, 2009
M.S. Boston University, 2013

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2016

Approved by

First Reader

Muhammad H. Zaman, PhD
Professor of Biomedical Engineering
Professor, Howard Hughes Medical Institute

Second Reader

Michael L. Smith, PhD
Associate Professor of Biomedical Engineering

Third Reader

Ahmad S. Khalil, PhD
Assistant Professor of Biomedical Engineering

Fourth Reader

Maria A. Kukuruzinska, PhD
Associate Dean of Research
Professor of Molecular and Cell Biology
Henry M. Goldman School of Dental Medicine

Research Assistant Professor of Biochemistry
School of Medicine

Fifth Reader

Roger D. Kamm, PhD
Cecil and Ida Green Distinguished Professor of Biological and Mechanical Engineering
Massachusetts Institute of Technology

DEDICATION

To my parents, Malay and Mridula.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Muhammad Zaman, for his incredibly kind and generous guidance during the development of this project, and for his confidence in me. I am grateful to my committee members, Dr. Michael Smith, Dr. Roger Kamm, Dr. Maria Kukuruzinska and Dr. Mo Khalil, for their guidance and mentorship. This work would not have taken any form without their support and belief. I thank my colleague and friend, Diego Vargas, who has been a partner-in-crime through almost all the endeavors presented here, from validating experimental tracking and analytical methods to computational modeling. I thank Dr. Kadin Tseng for advice, guidance and assistance with the BU shared computing cluster. I am especially appreciative of post-doctoral associates, Dr. Adrian Pegoraro and Dr. David Lepzelter, for their belief in this work and aid in the development of experimental and analytical methods. I thank Erika Fong for the opportunity to collaborate in early work of microrheology, and for her constant support and friendship through the years after. I am also indebted to many in the BME department at BU and in the Zaman Lab, their support and provocative conversation has led me down many paths: Dr. Dimitrije Stamenovic, Dr. Kamal Sen, Dr. Jaya Srivastava, Dr. Michael Mak, Dr. Dewi Harjanto, Dr. Oliver Bates, Dr. Brian Fallica, Dr. Joseph Maffei, Dr. Darash Desai, Dr. Grace Wu, Steve Wang, Nga Ho, Dan Reynolds, Ali Badreddine, Andrea Fernandes, and Alex Bloom.

Finally, I would like to acknowledge the unflinching support and steadfast encouragement of my family– my father, Malay, who first introduced me to scientific enquiry; my mother, Mridula, who made me love numbers a long time ago; my sister, Shreya, a constant source of joy; and my husband, Ali, who has held me steady through it all.

COLLECTIVE CELL MOTILITY IN 3-DIMENSIONS: DYNAMICS, ADHESIONS, AND EMERGENCE OF HETEROGENEITY

YASHA SHARMA

Boston University, College of Engineering, 2016

Major Professor: Muhammad H. Zaman, PhD, Professor of Biomedical Engineering, Professor, Howard Hughes Medical Institute

ABSTRACT

Collective cell migration is ubiquitous in biology, from development to cancer; it is influenced by heterogeneous cell types, signals and matrix properties, and requires large scale regulation in space and time. Understanding how cells achieve organized collective motility is crucial to addressing cellular and tissue function and disease progression. While current two-dimensional model systems recapitulate the dynamic properties of collective cell migration, quantitative three-dimensional equivalent model systems have proved elusive.

The overarching hypothesis of this work is that cell collectives are heterogeneous in nature; and that the influence of biochemical, physical, and mechanical factors combined leads to diverse physical behaviors. The central goal of this work is to establish standard tools for the understanding of 3D collective cell motility by treating individual cell-collectives as independent entities.

An experimental model studies cell collectives by tracking individual cells within cell cohorts embedded in three dimensional collagen scaffolding. A computational model of 3-dimensional multi-scale self-propelled particles recreates experimental data and accounts for intercellular adhesion dynamics. A custom algorithm identifies cellular cohorts from experimental and simulated data so these may be treated as independent entities. A second custom algorithm quantifies the temporal and spatial heterogeneity of motion in cell cohorts during ‘motility events’ observed in experiments and simulations.

The results show that cell-cohorts in 3D are dynamic with spatial and temporal heterogeneity; cohesive motility events can emerge without an external driving agent. Simulated cohorts are able to recreate experimental motility event signatures. Together these model systems and analytical techniques are some of the first to address collective motility of adhesive cellular cohorts in 3-dimensions.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
AJ	Adherens Junction
ATP	Adenosine Triphosphate
CCM	Collective Cell Motility
ECM	Extracellular Matrix
FBS	Fetal Bovine Serum
FOV	Field of View
GFP	Green Fluorescent Protein
MSD	Mean Squared Displacement
NHF	Neonatal Human dermal Fibroblasts
nM	nanoMolar
NLS	Nuclear Localization Signal
PBS	Phosphate Buffered Saline
PCA	Principal Component Analysis
RGD	Arginylglycylaspartic acid
SDM	Squared Displacement Median
SPP	Self Propelled Particle
TGF	Transforming Growth Factor

CHAPTER 1. INTRODUCTION

1.1. Problem Statement

Collective Cell Motility is fundamental to biological processes, and shares traits across multiple phenomenon such as development, morphogenesis, wound healing, and cancer metastasis (Rørth 2009a; Haeger et al. 2015). The dynamics of these processes are integral to their function, effectiveness and efficiency. However these dynamics are not easily amenable to *in vivo* or *in vitro* observations and they remain challenging to recreate and characterize in 3-dimensional systems.

Currently, studies of the inter-cellular dynamics of collective cell motion are limited to two dimensional (2D) monolayer experiments (Park et al. 2015). Experiments and computational models work in conjunction to describe collective motility, phases and forces in cellular monolayers (Méhes & Vicsek 2014). In 2D, epithelial cells and fish keratocytes have been used as model systems to study the dynamic aspects of collective cell migration (Rapanan et al. 2014). 3D *in vitro* studies of cell collectives typically involve invasion assays and immunohistochemistry to identify key proteins and mechanisms (Carey et al. 2013) and generally lack individual cell-based dynamic information. Conventional cell tracking research assumes that the behavior of cells in 2D and 3D is homogenous in time; information from all time points is averaged to deduce the timescales for various cellular behaviors (Zaman et al. 2006). Cell collective studies typically calculate the velocity correlation length of monolayers (Méhes & Vicsek 2014), which is not easily convertable to a 3D parallel since it depends on systems with uniform cellular density and does not account for matrix properties. Cells must balance adherence and motility to maintain organized coherent motion (Méhes & Vicsek 2014). While there are a large number of 2-dimensional experimental and computational models that recreate and predict experimental data, as well as analytical techniques for 2D collective motility and 3D cell migration, there are limited 3D counterparts for collective motility that account for the dynamics of individual cells.

1.2. Hypothesis and Aims

The overarching hypothesis of this work is that cell collectives are heterogeneous in nature, and that the influence of biochemical, physical, and mechanical factors combined leads to diverse physical behaviors. The central goal of this work is to establish standard tools for the understanding of collective cell motility by treating individual cell-collectives as independent entities. The following Specific Aims are proposed:

<u>Aim 1: To establish an experimental model for collective cell motility in 3D matrices</u>

Collective motility relies on cells cohorts establishing a heterogeneity within themselves to achieve supracellular polarity and organization. To this end, the material properties of different cells are tested for cancerous and non-cancerous cell populations under different biophysical and biochemical conditions. Studies of collective cell motility *in vivo* focus on development and morphogenesis. *In vitro* studies traditionally involve 2-dimensional monolayer dynamics, forces, and phases; or 3D static events that yield insight into biochemical triggers for collective cell motility but provide no dynamic information. This study provides an experimental model system that creates adhesive cellular clusters and grows and tracks these in a 3D culture system. In addition, a custom algorithm assigns each cell an identity for the cluster or cohort it belongs to, thus allowing analysis on a cohort-by-cohort basis. This work establishes a 3D experimental model that is able to create cellular cohorts in 3D, to track individuals within each cohort and also to treat each cohort as an independent entity for analysis purposes.

Aim 2: To establish a computational model for collective cell motility in 3D that recreates and predicts experimental data

For a large variety of systems, collective properties can be recreated using minimalist self-propelled particle models. This work adapts a self-propelled particle model that accounts for cellular attraction to 3D. The introduction of another dimension alone drives the need for a wider array of analytical techniques, which are then explored for ranges of input parameters. The model is then further complicated by allowing individual units or cells to have variable speed dependent on their adhesive properties. The clustering algorithm developed in Aim 1 is applied here as well, so that each simulated cohort may be analyzed on a case-by-case basis. Thus this computational model serves in conjunction with Aim 1, and creates simulated clusters that recreate experimental data, and can further co-evolve with experiments.

Aim 3: To establish analytical techniques for cell collectives that operate on a cohort-by-cohort basis

With a large number of cohorts and diverse observed behaviors, current cell analysis techniques failed to classify or quantify collective data. This work establishes a standard technique for analysis of such data on a cohort-by-cohort basis. Cohorts are characterized individidually using order parameters and displacement information, and motility events are identified for both experimental and simulated cohorts. Correlation functions serve to distinguish behavior of individual cells within the cohorts. These techniques classify and yield insight into the behaviors observed for three dimensional cell-collectives, and are also ameable for use in other systems of cells or collectively moving units.

1.3. Structure of Thesis

Chapter 2 introduces all the necessary background and literature. This chapter itself is divided into sections of cell motility, collective cell motility, and computational models for cell motility, which have traditionally been separate but connected fields. Chapter 3 covers studies into heterogeneity of material properties within cell populations, thus providing a motivation for studying cell collectives as evolving varied species. Chapter 4 covers the experimental model, validates experimental cell-tracking and explains the custom algorithm to categorize cells into cohorts or clusters. Chapter 5 describes the computational model for cell collectives and characterizes it for various input parameters. Chapter 6 describes analytical techniques that operate on a cohort-by-cohort basis, and analyzes the experimental and computational data presented to this point. Chapter 7 concludes the major findings of the work, and lays down potential future directions that can emerge from this work.

CHAPTER 2. BACKGROUND

2.1. Introduction

This study presents an experimental and computational model for collective cell motility in 3D as well as analytical techniques for such systems. The background literature is thus divided into three sections– Cell Motility and Migration in 3D, which is essential to understand the biological, biochemical and mechanistic origins of collective cell motility; Collective Cell Motility, a section which lays the foundation for typical collective patterns and motility observed *in vitro,* as well as significant factors *in vivo* and drives the motivation of this work; Computational Models for Cellular Systems, which are diverse in application and focus but essentially grounded in replicating specific aspects of observed phenomena with minimalistic rule-based tools, thus providing a reductionist approach to an increasingly complex system.

2.2. Cell Motility and Migration in 3D

In order to move inside a 3D system, a cell must participate in adhesion and invasion. Cell adhesion to the extra-cellular matrix (ECM) is accomplished when surface proteins on the cell attach to proteins in the ECM. Invasion refers to cells moving into a part of their environment that they did not previously occupy. Invasion is usually accomplished by one of two processes. Cells may break down ECM components in order to create a path in a process called *proteolysis.* Alternatively, cells may squeeze through available pores or holes in the ECM and find an already available path (Sharma & Zaman 2014). A combination of the two processes may also occur.

Conventionally, three characteristic steps have been used to describe cell motility:

1. Polarization and attachment to the front
2. Contraction of cell body forward
3. Detachment from the rear

Polarization refers to cells forming temporary differences in states across the cell body, generally to form a front and rear along the direction of motion. Cells in 3D do not have a dorsal-ventral polarity as observed in 2D and instead form an active ‘leading front’ or edge surface. Adhesion proteins in the cell membrane move towards this leading front and many active membrane protrusions reach out into the ECM. In the second step, adhesion proteins in membrane protrusions at the leading front attach to ECM components. These protrusions may also assist cells in breaking down local ECM proteins to carve a path, and are rich in cytoskeletal proteins which align along the direction of protrusions. The cytoskeleton must then exert force within the cell to contract the cell body forward in the third step. For the final step, adhesions in the rear are detached and either deposited in the matrix or relocated along the cell membrane (Friedl & Bröcker 2000).

Cells in 3D are highly dynamic. The three characteristic steps mentioned above can occur concurrently, a cell may polarize and adhere in one direction, and polarize and adhere in a different direction immediately after. In order to have a single successful step of cell motility, successive polarization, adhesion, contraction and detachment of a cell must occur along the same direction.

Cell motility is also crucial to cancer metastasis.(Friedl & Alexander 2011; Friedl

& Wolf 2008) Cancer cells acquire enhanced properties similar to those of stem cells and developmental cells, and the initial steps of cell motility allow cancer cells to escape from a primary tumor site and invade the surrounding environment. These cells may also then migrate to secondary sites on the host body and create secondary tumors. Understanding cell motility in 3D has long been considered a crucial aspect of understanding cancer dissemination and metastasis.

2.2.1. ExtraCellular Matrix

In vivo the extracellular matrix and environment of cells is dynamic, diverse, rich and complex and varies amongst different tissues in its composition, structure, and mechanical properties (Pedersen & Swartz 2005). It comprises proteins and polysaccharides that cells both respond to and modify, and therefore the ECM is constantly being 'remodeled' by cells. Transmembrane proteins known as integrins bind to ECM components to couple the cell cytoskeleton with the ECM. These cell-matrix adhesions are the fundamental drivers of ECM remodeling and cell fate, for they allow communication between the cell and ECM and the flow of signaling molecules in both directions (Baker & Zaman 2010; Berrier & Yamada 2007). The ECM provides structural integrity, assists in tissue organization, determines cell fate either directly or indirectly, and provides a substrate for motility. ECM proteins also regulate, distribute, and activate soluble growth factors as well as present them to cells (Griffith & Swartz 2006; Hynes 2009).

2.2.2. 3D Cell Culture

3D models of ECM allow us to attempt to bridge the gap between *in vitro* and *in vivo* cell studies. 3D models refer to 3D gel-like structures in which cells can be suspended, cultured, and observed. They are promising for therapeutic purposes as well from a tissue engineering perspective, allowing medical professionals to envision repairing and administering treatment for cell-ECM composites (Schmeichel & Bissell 2003). The types of 3D culture systems can be classified as natural or synthetic.

2.2.2.1. Natural Matrices

Natural 3D culture systems are usually harvested from cells or animals. When harvesting from animals, natural 3D models are derived by de-cellularizing tissues. Cell derived matrices are created by cells in culture systems and later de-cellularized (King & Parsons 2011; Ott et al. 2008). Matrices or gels can be made from reconsistituted collagen, fibrin, fibronectin, elastin, fibrinogen among others (Hooper et al. 2006). The advantage these gels offer over synthetic matrices is that they contain biochemical cues for the cells. Collagen gels can be tuned for pore size, stiffness, and ligand density by varying collagen concentration and pH (Harjanto et al. 2011). Ligand density refers to the number of available sites that the cells can cleave via proteolysis in a given volume of collagen. Denser and larger fibers of collagen result in stiffer matrices. Architecture control in collagen gels is achieved by aligning gels that are stretched. Tissue-transglutaminases are used to increase crosslinking, thus making collagen stiffer without increasing ligand density. However collagen gels are very heterogeneous on the micron scale, thus a cell may experience extremely different properties across its length or in different parts of a collagen gel.

2.2.2.2. Synthetic Matrices

Controlled or synthetic gels provide the option of molecular control and homogeneity in the matrix structure. However these gels are created at the expense of *in vivo* like biochemistry. Synthetic gels are often polymers such as poly(ethylene glycol), poly(vinyl alcohol), and polyacrylates such as poly(2-hydroxyethyl methacrylate). These are typically soft and elastic, thermodynamically compatible with water, and thus useful in many biomedical applications (Slaughter et al. 2009). Various properties such as network mesh size, fiber size, extent of crosslinking, gradients and known sequences for adhesion and degradation have been controlled in different synthetic matrices. Cells are viable in many such matrices but may not always be migrating. However the presence of degradable Arginylglycylaspartic acid (RGD) elements allows cells in many such matrices to cleave RGD and migrate through the hydrogel (Schwartz et al. 2010; Raeber et al. 2005; Guarnieri et al. 2010; Kenawy et al. 2003; Luo & Shoichet 2004).

2.2.3. Dimensionality

It is apparent that cells *in vivo* occupy a diverse environment in which they are surrounded by multiple ECM proteins, other cells, and cell-secreted factors from different cell types. Remarkable and numerous differences exist between these systems and traditional monolayer culture, and have been described by some as a ‘double-edged sword’ (Baker & Chen 2012). 2D monolayer systems allow for reductionist methods and simple and categorical classification of different cell types and properties, yet they do not resemble 3D counterparts.

Dimensionality accounts for differences between monolayer, 3D *in vitro*, and *in vivo* systems of cells. Since cells sense, respond to, and modify their environment using a number of mechanisms, the effect of dimensionality cannot be understood as a single parameter. For example, while most cells move slower in 3D collagen gels than they do on 2D glass or collagen coated dishes (Harjanto et al. 2011), neutrophils that are almost still on glass dishes in 2D cultures display motility in 3D collagen gels (Brown 1982). Within 3D too, certain cells migrate proteolytically while others migrate by finding a path through squeeze through (Wolf et al. 2003). These observations, along with others, have led to the understanding that dimensionality must be explored while taking every factor of experimental conditions into account.

2.2.4. Initial Steps: Polarization and Adhesion

The polarization of cells is a decisive first step in all migration modes, regardless of dimension. Little is known about the exact initiator of polarization in 3D cell migration, but due to a combination of various cell and ECM properties, two co-dependent processes occur in order to initiate cell motility− actin polymerization to form membrane protrusions, and clustering of adhesive proteins at the leading edge. Cells are extremely dynamic structures on the scale of seconds; their membrane is constantly in flux, probing the environment all around as far as possible given the state of the cell. If the membrane, while exploring its environment, comes into contact with a cue that is able to differentiate the cell along a given direction, the polarization process can be initiated.

Cells form membrane extensions in 3D called protrusive pseudopodia or finger like filopodia at the leading edge. High actin polymerization is seen at the leading edge and generates a propulsive force in these bodies (Mogilner & Oster 1996). Actin monomers (G-actin) are globular proteins that spontaneously assemble to form filaments (F-actin) and have been extensively studied in the cytoskeletal context (Kim et al. 2007). These filaments have a barbed end, the site of addition of G-actin monomers and growth, as well as a pointed end, the site of removal of G-actin monomers. Many F-actin filaments are bundled together by cross-linking proteins such as Fascin, and the barbed end of collective fibers is oriented towards the membrane. Thus, continual actin polymerization at the barbed end results in the membrane protruding outwards to form filopodia (Vignjevic et al. 2006). Filopodia probe the environment for extracellular cues, and the active polymerization initiates clustering of integrins to the leading edge of the cell.

Cells in 3D form adhesions in all directions and more adhesions are recruited to the leading edge as the cell polarizes (Lundquist 2009; Chhabra & Higgs 2007). Integrins are transmembrane heterodimeric proteins containing two distinct subunits called alpha (α) and beta (β). So far eighteen α and eight β subunits have been characterized for mammals. Both subunits can enter the plasma membrane and both have small cytoplasmic portions as well (Humphries et al. 2006). Combinations of these subunits allow cells to bind to a number of proteins: collagen, laminin, fibronectin, vitronectin, fibrillin, fibrinogen, and more. When integrins transport to the leading edge after formation of filopodia and pseudopodia, their β units bind to ECM proteins. This results in integrin activation and anchorage proteins in the cytoplasmic portions of integrins bind to the actin cytoskeleton (Baker & Zaman 2010). Actin stabilizing proteins vinculin, α-actinin and paxillin assist in this binding, and recruit signaling proteins such as FAK and Src (Berrier & Yamada 2007). This starts a signaling cascade that results in recruiting additional integrins into the region, yielding a cluster of integrins or a focal complex.

2.2.5. Intermediate Steps: Force Generation and Contractility

The cellular skeleton or cytoskeleton is responsible for generating force inside of cells. Probing intracellular components in 3D systems presents a technical challenge, and there is still much to learn about the exact mechanism of force generation and contractility. The most important cytoskeletal proteins in the context of cell motility are actin, myosin, and microtubules. Disruption of any of these has been shown to decrease intracellular forces (Kraning-Rush et al. 2011). Live-imaging has shown that the actin protrusive force generated in filopodia oscillates with a retraction of actin in the rear to generate motion (Starke et al. 2013). In some cases, actomyosin inhibitors showed that forward protrusive force by actin polymerization could alone allow cells to move, but these forces are of much smaller magnitude than produced by the actomyosin system (Fournier et al. 2010). Actomyosin is formed when myosin motors are translocated along actin filaments using ATP, an extremely well characterized system that elucidates muscular contraction for us. The Rho/ROCK pathway is crucial to this process.

2.2.5.1. Rho/ROCK pathway

The Rho-associated protein kinsase (ROCK) is a downstream effector of a small GTPase, RhoA. In 3D this pathway is essential to cell motility, it causes collagen fibers in the ECM to align by remodeling in the direction of motion and induces phosphorylation of the myosin light chain to generate force (Provenzano et al. 2009). This process of aligning the matrix fibers along the direction of motion is called contact guidance. Reducing ROCK results in cells shaped dendritically with reduced motility compared to standard spindly shape cells. A dendritic shape implies that cells have plenty of protrusions in various directions with no preferred direction of movement. Active ROCK results in more polarity and matrix contraction (Gunzer et al. 2000; Roy et al. 1997; Kim et al. 2006).

2.2.5.2. Growth Factors

While an understanding of how growth factors and other small molecules assist force generation in 3D systems remains incomplete, some studies of 3D force generation have been conducted using 3D traction force microscopy (Franck et al. 2011). A biphasic response of force generated is seen after adding increasing amounts of the growth factor TGF-β to cells in collagen gels. This implies that increasing TGF-β initially increased traction force generated in cells, and after peaking as a bell shaped curve, decreased traction formation. Two forms of TGF-β, TGF-β1 and TGF-β3, together increased the biphasic response of force generated, suggesting a cooperative mechanism of action (Brown et al. 2002).

2.2.5.3. Adhesions

In both 2D and 3D systems, cell motility is biphasic with regard to adhesions. Adhesions allow cells to exert a frictional force opposite to the direction of motion and move forward. When there are no adhesions, cells are not able to push against any substrate, and do not move forward. With increasing adhesions, cells increase their frictional force. However this frictional force does not have the same linear relationship with motility as in classical mechanics. Too many adhesions imply that there are too many cell-matrix connections that the cells may not be able to completely detach in order to move forward. Thus increasing adhesions results in a bimodal motility behavior, with an initial increase in motility leading to a peak, and then a decrease in motility (Zaman et al. 2006).

2.2.6. Terminal Steps: Possible Detachment in the Rear

For 3D motility, the exact mechanism of detachment and retraction in the rear of the cell is unknown. Some research indicates that actomyosin contractility actively contracts and releases the rear (Clow & McNally 1999; Kirfel et al. 2004). Over 80% of the cells' integrins are left behind in the ECM during this process, while the rest will be used and recycled by the cell for a next motility step (Palecek et al. 1996). The ligands may be weakened simply by weakened kinetic affinity and a low binding state (Friedl & Bröcker 2000). Integrins can also be detached or broken down by proteases in the cell, and this debris can leave migration tracks that provide cues and guides for other cells so they follow the same path (Palecek et al. 1998). This is an example of the kind of natural biochemical signal that could be available to cells when using naturally derived matrices as opposed to synthetic and well controlled constructs. Rear retraction and detachment is considered to be the rate-limiting step in cell motility, implying that the speed of cell motion is primarily determined by how effective the cell is in detaching from the matrix, and not how protrusive it is, or how well it generates force.

2.2.7. Probing Intra-Cellular Material Properties in 3D

The thermal motion of embedded particles, intrinsic vesicles, or mitochondria within cells can yield information about material properties of the cytoplasm. Passive particle tracking microrheology is a method that allows extraction of viscoelastic moduli of complex fluids (Mason et al. 1997). The thermal fluctuations within the cytoplasm are related to the fluids viscoelastic creep compliance Γ(t) by a modified Stokes-Einstein relation as described in Equation 2-1 (Wirtz 2009). Here n is the degrees of freedom of the particle or bead; k_b is Boltzmann's constant, and T is temperature, so k_bT is the thermal energy available to the system; a is the radius of the assumed spherical tracer; and $<\Delta r^2(t)>$ is the time-lag dependent mean squared displacement of the tracer.

<u>Equation 2-1</u>

$$\Gamma(t) = \frac{3\pi a}{nk_bT}\langle \Delta r^2(t)\rangle$$

The viscoelastic creep compliance of 3D embedded cells is then calculated by imaging tracers at high frequencies for 10-100 seconds. Tracking of nanoparticle probes and fluorescent mitochondria has yielded viscoelastic moduli for cells in 3D (Panorchan et al. 2006; Mak et al. 2014). However this method only accounts for all the internal fluctuations within the cytoplasm arising from Brownian or thermal motion, and does not account for the diverse and varied effects of motor proteins, organelles, and cellular processes that may cause thermal seeming fluctuations over 10-100 seconds. To account for the forces within a cell, a combination of these methods with active particle tracking microrheology, in which an optical trap can exert a known force on the tracers and follow the response, allows for characterization of both thermal and active components of intracellular fluctuations (Guo et al. 2014).

2.2.8. Modes of Migration in 3D

When a cell undergoing motility becomes directed and moves towards a given direction with an average speed and velocity, the process is called cell migration. There are two primary modes of cell migration in 3D. The first is mesenchymal migration, in which cells invade into the matrix by creating a path in the matrix around them and is dependent on proteolysis and actin polymerization. The second is amoeboid migration in which cells squeeze through any path available to them using membrane protrusions called blebs.

2.2.8.1. Mesenchymal Migration

Mesenchymal migration can be said to have four characteristic steps. In addition to the 3-steps mentioned above, cells break down the matrix around them after forming protrusions and remodel it along their path. The proteolysis and matrix remodeling is followed by force generation and rear retraction. However these processes are not completely independent of each other, and all evidence to date indicates that they are correlated with each other and even positively regulating each other. For instance, the

Rho/ROCK pathway directly controls actin contractility, but seems to have a strong impact on matrix remodeling. Cells are observed to change the orientation of collagen fibers and align them along their direction of motion. However, when the same cells are ROCK-inhibited, they do not display such behavior. Interestingly, when the same cells are seeded in pre-aligned collagen, they do not reorient the fibers or display much motility (Provenzano et al. 2008).

Proteolysis is perhaps the most important step in 3D cell migration. Since there are so many steric hindrances in 3D, the ability of cells to break the matrix down around them and create a new path is critical for migration. It allows cells to create a path in extremely dense matrices, with pore sizes through which they would not be able to migrate otherwise. In fact, when proteases were inhibited in synthetic gels with a pore size of 0.025 μm, cells were not able to migrate at all (Raeber et al. 2005). Many cancer researchers are extensively focusing on protease activity in cancer cells to probe the critical contributors to cancer invasion and metastasis (Pietras & Ostman 2010; Friedl & Wolf 2008).

2.2.8.2. Amoeboid Migration

The other primary mode of cell migration and motility is amoeboid migration. It is characterized by rounder cells, with fewer protrusions, less structure in the cytoskeleton, no proteolysis, and no co-clustering of adhesion proteins (Guck et al. 2010). It can commonly be observed in cancer cells, leukocytes and embryonic cells. Cells form blebs, which are protruding portions of the cell membrane that are disconnected from the cortical cytoskeleton because of pressure gradients. These are distinct from filopodia, which protrude outwards because they are connected to a rapidly expanding cortical cytoskeleton. In this case, however, the cortical cytoskeleton may rupture or a high pressure may cause part of the cell membrane to protrude outwards. Once blebs are formed, actomyosin contractility allows cells to move forward and migrate (Charras & Paluch 2008; Narumiya & Watanabe 2009; Wolf et al. 2003). Some argue that amoeboid motility is not physiologically relevant because it has only been observed in collagen matrices that are not cross-linked, a scenario unlikely to be found *in vivo*. When cells were inhibited for MMPs and placed in collagen containing cross-linkers, cells reduced speed but did not migrate using an amoeboid mobility (Sabeh et al. 2009; Bloom et al. 2008). However leukocytes use amoeboid motility with higher actin contractility instead of bleb formation (Renkawitz et al. 2009). Similarly, dendritic cells can migrate with all integrin expression knocked out, indicating that their migration is occurring independent of adhesion and therefore likely by an amoeboid mechanism. However, this phenomenon is only observed in 3D matrices, suggesting perhaps that cells in 3D and *in vivo* have versatility and plasticity in migration modes (Lämmermann et al. 2008).

2.2.8.3. Determinants of Migration Mode

Amoeboid migration is often observed in cells that have undergone an Epithelial Mesenchymal Transition (EMT). HT-1080 fibrosarcoma cells and MDA-MB-231 carcinoma cells generally adopt a mesenchymal migration mechanism and have been shown to have proteolytic tracks, integrin-clustering and integrin colocalization with MMPs. However, when MMPs are inhibited, these cells are able to switch migration strategies (Wolf et al. 2003). Similarly MMP-inhibited cells in collagen or fibrin gels with large pore sizes of 0.5 and 7.4 μm respectively were unaffected in migration (Raeber et al. 2007; Raeber et al. 2005).

Cellular properties that determine the mode of migration are the volume and deformability of the cytoplasm, and the volume and deformability of the nucleus. If either the cytoplasm or the nucleus is not deformable to the appropriate degree, the cell will be unable to squeeze through pores in the matrix. ECM properties that determine the mode of migration are porosity, alignment, and stiffness or elasticity of the matrix. If the ECM does not have large enough pores, amoeboid migration is impossible. Stiffer gels may force cells to change the mode of migration as well. Material properties are also crucial to migration mode, if a synthetic and non-degradable material is used, proteolysis is impossible (Wolf & Friedl 2011).

2.3. Collective Cell Motility

From early development through morphogenesis, wound healing, and even through cancer pathologies, biological function hinges on collective cell motility. Collective cell motility is the phenomenon describing cells that move together while affecting one another, making contact either temporarily or permanently (Rørth 2009a). This includes cells that move together by being physically coupled to each other, as well as cells that are not connected but follow the same path because of guidance and cues placed into their environments by leading cells and magnified by other cells ahead of them (Gov 2007). In the former case, cells must balance adherence and motility to maintain organized coherent motion (Méhes & Vicsek 2014). Certain definitions of collective cell migration do not take the ECM-based coordination into account, and require that cell-cell contact is established permanently (Friedl & Gilmour 2009). Forms of collective motion where cells move locally in an ordered fashion in a globally disordered system are termed jamming, and have been reported in 2D and 3D (Haeger et al. 2014; Park et al. 2015). Three characteristic hallmarks of adhesion-based collective cell motility are as follows:

1. Increase of cell-cell contacts
2. Supra-cellular polarity
3. Joint guidance along migration tracks

Cell-cell contact or adhesion is cadherin-dependent, and in some cases is essential to collective motility (Gloushankova 2008). Cadherins are mechanically linked to the actin cytoskeleton, and are essential for actin organization at a supra-cellular scale in order to generate contractile force. Transfer of reagents between cells using vesicles has also been found to be essential for such forms of collective migration (Kriebel et al. 2008). In some cases cells will convert from collective to individual movement when either cell-cell or cell-matrix adhesions are disrupted (Hegerfeldt et al. 2002; Friedl et al. 2004).

Supra-cellular polarity can manifest due to shape or physical restriction, tensional polarity caused by organization of cellular acting cytoskeletal elements across a collective, as well as biochemical polarity due to varied expression of ECM markers across the front and back of the collective (Haeger et al. 2015).

Cells moving collectively can deposit a basement membrane or create or widen a path for other cells. This phenomenon has been termed tissue micropatterning (Gaggioli et al. 2007). Individual migration can transition to collective migration as small paths or tracks formed by leading cells are progressively widened to fit multiple cells at once (Friedl & Wolf 2008). Therefore, collective migration offers multiple physiological and energetic advantages. Clusters have increased sensitivity to chemical or mechanical gradients, reduced metabolic cost since mobile cells are able to carry other cells along, cooperative behavior since cells can influence each other and shape the tissue and collective decisions involving less wasted random motion.

Collective movement has historically been divided into two-dimensional migration in which a sheet of cells migrates across a tissue surface, or three-dimensional in which multi-cellular structures of cells move across a tissue scaffold (Friedl et al. 2004). Two dimensional or sheet migration has been extensively studied because it is convenient for imaging. Cells remain close together as the sheet moves forward. Cells at the front edge are more motile and protrusive and are defined as leaders. The leading front of cells moves along the direction of motion, either to heal a wound or to occupy unconstrained substrate space (Poujade et al. 2007). In three dimensions, collective migration has been observed in the form of vascular sprouting for angiogenesis, branching for morphogenesis in mammary glands, free groups of border cells, detached clusters, and multicellular 3D invasion strands in cancer (Friedl & Gilmour 2009).

2.3.1. 2-dimensional Collective Motility

2-dimensional assays of collective motility typically involve tracking individual units on a plated sheet or monolayer. Tracking assays on monolayers reveal that collective cell motility in 2D-sheets undergoes density dependent phase transitions (Szabó et al. 2009). Phases are identified by and order parameter $\overline{V}$ according to Equation 2-2 and range from disordered to ordered, here N is the number of cells, p is the cell in question, and v_p is the velocity of this cell at time t.

Equation 2-2

$$\bar{V}(t) = \langle \frac{1}{N} \left| \sum_{p=1}^{N} \frac{\vec{v}_p\,(t)}{\left|\vec{v}_p\,(t)\right|} \right| \rangle$$

Density dependent behavior has also been observed in epithelial Madin Darby Canine Kidney (MDCK) cells which behave in a correlated manner up to specific correlation lengths (Doxzen et al. 2013). Correlation length for monolayers is typically obtained by calculating the radius r at which a velocity-correlation function C(r) as depicted in Equation 2-3 is 0. Here $u^*(r)$ is the velocity obtained after subtracting mean velocity u of the monolayer form a particular position r.

Equation 2-3

$$C(r) = \langle \frac{u^*(\vec{r}\,' + \vec{r}, t\,) \times u^*(\vec{r}\,', t\,)}{[\langle u^*(\vec{r}\,' + \vec{r}, t\,)^2 \rangle \langle u^*(\vec{r}\,', t\,)^2 \rangle]^{\frac{1}{2}}} \rangle_t$$

Monolayer stress microscopy identifies forces exerted by the monolayer by tracking fluorescent probes in the cell monolayers as well as in the substrate beneath these (Brugués et al. 2014; Tambe et al. 2011). Force measurements on specialized substrates have found that cell sheets exert significantly more force than a single cell (Haga et al. 2005; du Roure et al. 2005; Tambe et al. 2011). A comprehensive study of the mechanical properties of epithelial monolayers identified E-cadherin and P-cadherin as key proteins contributing to intercellular forces (Bazellières et al. 2015). Madin Darby Canine Kidney (MDCK) epithelial cell sheets migrating collectively in two dimensions on top of a collagen gel are disrupted by the removal of a leader cell (Yamaguchi et al. 2015). Further live imaging and immunohistochemistry revealed Rac, Integrin β1, and PI3K as key molecules upregulated in leader cells in these sheets. In the case of keratinocytes grown on a 2D monolayer, locally correlated rows of cells are observed in an adhesion dependent phenomenon dubbed streaming (Czirók et al. 2013).

2.3.2. 3D Collective Motility

In development and morphogenesis, 2D time-lapse imaging and immunohistochemistry of drosophila border cells (Cetera & Horne-Badovinac 2015; Vogler & Bodmer 2015), frog embryos (Theveneau & Mayor 2012), and zebrafish (Dalle Nogare et al. 2014), reveal relevant biochemical markers and mechanical characteristics of collective cell migration (CCM) (Weijer 2009).

For example, in *Drosophila*, E-cadherin is essential for collective direction sensing (Cai et al. 2014), and tissue rotation is essential for building an extracellular-matrix (ECM) to control egg shape (Haigo & Bilder 2011). During frog or amphibian gastrulation, the F-actin network is responsible for changing patterns of strain rate (Feroze et al. 2015). In the zebrafish primordium, communication between leading and trailing cells has been attributed to FGF signalling (Dalle Nogare et al. 2014).

In sprouting and branching, a moving outgrowth comprising many cells arises from an existing structure. The tip has automatic polarity because it is free whereas the back end is attached. The vascular endothelial growth factor (VEGF) has been proven to be critical in such forms of migration in the case of angiogenesis (Ferrara et al. 2003). Neural Crest cells migrate in loosely connected streams and may reflect typical Epithelial to Mesenchymal Transition (EMT). The cell shapes are elongated and polarized, migrating cells secrete matrix themselves and significantly alter their microenvironment. The noncanonical Wnt pathway, a reaction pathway in cells that is independent of transcriptional function of β-catenin, elicited by cell-cell contacts is believed to inhibit intermingling of streams (Rørth 2009a).

Drosophilia border cells are the only known category of collective cell migration that do not have an inherent polarity within the cluster. A small cluster of eight cells migrates by squeezing between surrounding cells. Cells often exchange position within the cluster and are associated with surrounding cells by E-cadherins. However it has been shown that polarized cell behavior is important for starting the first phase of migration even in these cells, and the polarity disappears as dynamic collective behavior emerges. Platelet Derived Growth Factor (PDGF) and VEGF are important for both the initial and later stage, but employ different pathways in each. For these cells, multicellular level information processing has been proven to guide the direction of collective behavior (Bianco et al. 2007).

2.3.3. Cancer Metastasis and Motility

Cancer metastasis refers to the ability of cancer cells to escape from their native environment, travel to a secondary site, and create a new secondary tumor. Cell motility and plasticity is at the crux of cancer metastasis. Cancer pathologies are not amenable to direct observations of coherent translation for mammalian cells; however, indirect evidence from *in vivo* measurements demonstrates that cancer metastases can migrate through tissue layers as a collective mass (Deisboeck & Couzin 2009; Friedl & Gilmour 2009). For such pathologies, immunohistochemistry has elucidated crucial biochemical markers (Khalil & Friedl 2010); time-lapse brightfield microscopy highlights cell dynamics and cell jamming (Haeger et al. 2014); together these data suggest that cancer cells have inherent plasticity of migration modes and the ability to transition between these modes (Friedl & Alexander 2011).

Collective decision-making and sensing of environmental signals represent adaptive behaviors believed to offer advantage to tumors (Deisboeck & Couzin 2009). Collective motion in tumors has many behaviors similar to morphogenesis and development. This has been observed to occur by two different mechanisms. On one hand, protruding sheets and strands can maintain contact with the initial tumor and then generate local invasion. On the other hand, a cluster of cells may detach from the original tumor, and extend across gaps in tissue along the path of least resistance (Mierke 2013).

Physical observations have led to the belief that collective migration is the principal method of invasion in epithelial cancers. This is especially the case for epithelial cancer cells that do not undergo EMT, which used to be considered the only form of tumor dissemination and invasion (Thiery & Lim 2013; Thiery 2002; Thiery 2009). Cells that retain their epithelial characteristics are more inclined to retain cell-cell contacts and migrate in a cluster or collective, as has been observed in Squamous Carcinoma Cells (SCC). When only a few of these cells acquire invasive properties, they become leader cells and can allow an otherwise non-invasive cluster to become invasive, as has been shown in some *in vitro* studies (Gaggioli 2008; Gaggioli et al. 2007; Carey et al. 2013). In fact, it is believed that the larger a group or cluster, the fewer leader cells are required to guide it. Thus, EMT is no longer considered a requirement for a cancer to be invasive (Rørth 2009b; Kabla 2012; Khalil & Friedl 2010).

On a molecular level, cells at the leading edge depend on increased MMP (Friedl et al. 2004) and integrin expression. In fact, collectively moving melanoma cells in 3D with inhibited integrin activity switched to an individual, amoeboid, proteolysis-independent mode of motion (Hegerfeldt et al. 2002). The E-cadherin-adhesion stabilizing protein p120ctn in tumor cells is advantageous to 3D migration and disadvantageous to 2D migration (Macpherson et al. 2007). Pancreatic tumor progression has also been shown to be invasive without the presence of EMT. Leader cells rich in the actin regulator Fascin have been observed in colorectal carcinomas which are also invasive in the absence of EMT (Vignjevic et al. 2007). Podoplanin, an integral membrane glycoprotein, is found *in vivo* at the invasive leading edge on human breast cancer cells (Wicki et al. 2006). Amplification of chemical signals is an important contributor to the invasive potential of gliomas. They use an amino acid transporter system to import cysteine in exchange for the release of glutamate. Glutamate appears to induce local collective migration of these cells. There are hypotheses that positive feedback also occurs with glutamate production (Sontheimer 2003; Takano et al. 2001). The higher presence of glutamate could also be contributing to the process of tissue micropatterning, allowing these cells to create a path through which to migrate (Domercq et al. 2007).

2.3.4. Collective Motion in Other Species

Recent studies by statisticians, biologists, physicists and scientists alike have come to believe that the same set of physical laws can be used to describe all collective behavior. Therefore, it is advantageous to understand collective motility in other systems and the set of laws satisfied by these.

Collective motion or flocking has been observed in bacterial colonies, cells, insects, fish schools, bird flocks, and mammals (Tunstrøm et al. 2013; Hayakawa & Furuhashi 2012; Buhl et al. 2006; Czirók et al. 1996; Zhang et al. 2010; Tkačik et al. 2014). Despite the vast differences in individual units in these systems, similarities have been observed in the various patterns formed: coherently moving clusters, mills, stripes or toruses. Similar statistics and modeling techniques have been successful at replicating behavior for many of these systems, leading to the belief that common physical laws can be used to describe them. The important difference between these systems and inactive systems behaving as a collective is that conservation of momentum laws no longer apply. Most systems have a specific range of interaction both spatially and temporally, within which a single unit influences the actions of others.

In bacterial colonies, various patterns such as rotating dense aggregates, vortices, cluster migrations, turbulent and rotating states as well as fractal patterns of growth have been observed (Czirók et al. 1996; Fujikawa & Matsushita 1989; Wu et al. 2009). While under standard conditions these colonies do not exhibit a large scale of organization, under certain hostile conditions organization is possible. The bacterium *Bacillus subtilis* has cell-differentiation and long range information transfer under conditions with limited nutrients or harder agar surfaces. Colonies of *Myxococcus xanthus* periodically reverse their direction in order to generate a more orderly swarm and reduce collisions with each other (Wu et al. 2009). The effect of cell-density on collective behavior of bacterial colonies suggests that denser colonies become more organized (Zhang et al. 2010). Studies also suggest that local alignment among *Escherichia coli* cells occurs because of collisions and hydrodynamic interaction (Darnton et al. 2010).

2.4. Computational Models

The scales and complexities involved in cell motility provide a challenge to computational modeling, especially in 3D systems. Recently though, new data continues to increase and expand our approaches to mathematical models. Computational models are able to replicate behavior from real experiments and have predictive power for trends and dependence on various parameters. A few categories of computational models for 3D motility that are able to recreate and predict experimental data are described below.

2.4.1. Force-Based Models

Force-based models utilize the characteristic steps of mesenchymal migration to calculate migration speeds. They account for traction forces, matrix density, cell matrix adhesivity, and drag force from the matrix, and have been further extrapolated to include proteolysis and small cell-clusters. These models are thus able to calculate cell migration speed dependence on stiffness, adhesion, ligand characteristics among other variables (Zaman et al. 2005; Zaman et al. 2006).

The first force-based 3D migration model uses a force balance equation, as shown in Equation 2-4, to calculate instantaneous speed of the cell for each time-step (Zaman et al. 2006). The traction force ($F_{traction}$) comprises at least two components: one in the direction of the movement and the other opposing the movement, represented as F_{trac-f} and F_{trac-r} forward and rearward, respectively and calculated using Equation 2-5 and Equation 2-6. In these equations, the traction force in either direction is dependent upon the force per receptor ligand complex (F_{R-L}) which has a linear dependence (with a constant $c_1$representing area) on Young's modulus (E_{mod}) of surrounding microenvironment up to a certain maximum value, after which the traction force is a constant maximum (c_2). The traction force also scales with a time-dependent adhesivity term $\beta(t)$ that can account for differences in the front and rear of the cell. The protrusive force (F_{prot}), is a function of actin polymerization; and drag force (F_{drag}) that is dependent on medium viscosity. Direction of protrusion is randomly chosen after every time-step, equivalent to 600s which is the time required for cells to form stable protrusions that result in cell motility. Smaller protrusions and edge-dynamics generally occur at time scales of a few seconds and can be excluded from consideration.

Equation 2-4

$$F_{total} = F_{prot} + F_{drag} + F_{traction} = 0$$

<u>Equation 2-5</u>

$$\begin{bmatrix} F_{trac-f} = F_{R-L} * \ \beta_f(t) \\ F_{trac-b} = F_{R-L} * \ \beta_r(t) \end{bmatrix}$$

<u>Equation 2-6</u>

$$F_{R-L} = \begin{Bmatrix} c_1 \times \ E_{mod} & E_{mod} < 1MPa \\ c_2 & E_{mod} \geq 1MPa \end{Bmatrix}$$

The adhesivity factor $\beta(t)$ is a dimensionless parameter calculated according to Equation 2-7. $\beta_f(t)$ is a multiple of n which is the number of receptors in the front or rear, $[L]$ is the concentration of ligands and k the binding constant in the front or rear (subscripts f and r respectively). Depending on the cell type and microenvironment being considered, these values can be varied to result in either equal or varying adhesivity in the front and rear. The ligand concentration is assumed to be constant for the first model. The drag force (F_{drag}) in Equation 2-8 can now be used to calculate a speed (v) of the cell assuming the cell is a sphere with known radius (r) in a viscoelastic medium with a calculated effective viscosity (η) .

<u>Equation 2-7</u>

$$\begin{bmatrix} \beta_f(t) = \ k_f \times n_f \times [L_f] \\ \beta_r(t) = \ k_r \times n_r \times [L_r] \end{bmatrix}$$

<u>Equation 2-8</u>

$$F_{drag} = 6\pi \times r \times \eta v$$

This model was able to successfully replicate the bimodal dependence of cell speed on ligand density, adhesivity, and detachment force (traction force at the rear.) This model also considers asymmetry of cells and shows that the higher degree of asymmetry, the lower cell speed is likely to be. However this initial model is simplistic in assuming uniform ligand density in the environment of the cell, and successful improvements on the model involve the inclusion of proteolysis to modify ligand density (Harjanto & Zaman 2010). MMPs in proteolysis degrade the matrix around them, and under the assumption that ligand concentration is linearly related to MMP concentrations, two possible profiles of ligand concentration in relation to MMP concentration ($[M]$) were investigated: linear and log-linear, as shown in Equation 2-9. In both of these cases only the ligand concentration in the front of the cell was investigated since MMPs are recruited to the front or leading edge of the cell in the first step of migration. Since cell adhesion receptors or integrins also regulate MMP expression, the number of receptors available to the cell is a function of MMP concentration ($[M]$) and total receptors available (r) as shown in Equation 2-10. Of these, 95% of receptors are assigned to the front of the cell while the remaining are assigned to the rear according to Equation 2-11. With these modifications the model predicts a bimodal relationship of cell speed with MMP concentration as well as ligand concentration.

<u>Equation 2-9</u>

$$\begin{bmatrix} [L_f] = [L_{f1}] - [M] \\ [L_f] = [L_{f1}] \times (1 - [M]) \end{bmatrix}$$

<u>Equation 2-10</u>

$$n \propto r \times [M]$$

<u>Equation 2-11</u>

$$\begin{bmatrix} n_f = 0.95 \times n \\ n_r = n - n_f \end{bmatrix}$$

This model was also successfully modified to track a single cell cluster speed in 3D (Vargas et al. 2011). Cell clusters comprising five cells were investigated in 7 different geometries with clusters, 6 of which are more planar geometries and one of which has a 3D-arrangement of cells. Protrusion forces and traction forces are calculated for all cells in the cluster and added to obtain the force balance in Equation 2-4.

In order to account for collective motion, a leader cell and a trailing cell are assigned to the cluster in accordance with theoretical and experimental observations. A single cell was assigned as the trailing cell for all geometries. Any cell, based on the randomized direction of the protrusion force, can be assigned as the leading cell for each time step. In any given time step, there is a small probability of cells that are not either leading or trailing to exchange positions; this accounts for cellular translocation within the cluster. All non-trailing cells that do not switch positions contribute traction forces to the cluster. The leader cell contributes the major protrusive force, and all non-trailing cells contribute 25% of the protrusive force of the leader cell.

Since cells are represented with spherical geometry and various cluster structures are studied, the drag force is modified to account for this change as shown in Equation 2-12. In this equation, ρ is the medium density, v is the speed of the cluster, C_D is a dimensionless drag coefficient and A is projected area of cluster in the direction of motion.

<u>Equation 2-12</u>

$$F_{drag} = \frac{1}{2}\rho v^2 \times C_D \times A$$

The incorporation of different cluster geometries into the simulation resulted in noisy data for average cluster speed over hundred runs of the simulation. While bimodal trend of data previously observed were preserved in some cases of tetrahedral geometry alone with no changes, incorporating transitions in geometry gives more complicated and realistic data for cluster behavior.

Force based models have also been modified in clusters to account for cell-cell adhesion and for cell-pair interactions in clusters as big as 50 cells (Frascoli et al. 2013). Recently, a stress-based model in 3D has also been adapted to account for actin flow, adhesion and morphology (Shao et al. 2012). This model accounts for protrusions into the ECM as flat lamellopodia and has a diffusion gradient across the membrane. Actin is treated as a viscous fluid, and contractile stresses are generated for both actin and myosin protrusions. Focal adhesions are created at random across the cell interior in proportional concentrations to local actin density. This model was effective at predicting morphology for stationary and moving cells, and biphasic dependence on adhesion among other factors.

2.4.2. Other Single-Cell models

2.4.2.1. Stochastic Models of persistent random walks

Stochastic random walk models treat each cell as having a persistent random walk (Dickinson & Tranquillo 1993) and are effective at describing and predicting the behavior of a population. They describe cells within a cubic volume element and find the path taken by each cell according to Equation 2-13 (Dickinson & Tranquillo 1993) where S_n is the root mean square speed, $\langle D^2(t)\rangle$ is the mean square displacement of the cell as a function of time, a is the dimensionality of the system, and P is the persistence time, which is defined as the time span in which cells do not change direction. These models have been successfully used to predict neutrophil motility in 3D matrices (Parkhurst & Saltzman 1992). Other models utilizing this method have incorporated anisotropy and stimulus gradients (Dickinson 2000). These models have been successfully used to predict neutrophil motility in 3D matrices and have been adapted to incorporate anisotropy and stimulus gradients (Parkhurst & Saltzman 1992; Dickinson 2000).

Equation 2-13

$$\langle D^2(t)\rangle = a(S_n)^2 P\left(t - P + Pe^{-\frac{t}{P}}\right)$$

2.4.2.2. Monte-Carlo Models

Monte Carlo methods are based on lattice site occupation of cells, and when applied to model cell motility and migration, are useful for adding ECM factors and accounting for other parameters qualitatively. Some lattice based models assign space to lattice sites that represent the cell and the ECM, where the cell can be proliferative, quiescent or necrotic depending on nutrient levels and local conditions, and the ECM can be in fibrous or non-fibrous states (Rubenstein & Kaufman 2008). Cells can change between the three conditions with necrosis being a terminal state. Proliferating cells result in more cells replacing ECM sites. A differential adhesion hypothesis that states that cells move until they reach an arrangement with lowest adhesion energy is applied. Cell-cell and cell-ECM adhesion sites are given different energies based on experimental data, and the lattice evolves until the lowest energy state is obtained. Although this is a 2D simulation, it captures cell contractility in cases where cell mass goes through ECM channels smaller than the cell body, and thus can be considered a 2D slice of a 3D experiment. Such models have also been expanded to 3D lattices with varying pore size that allow cells to choose between forms of mesenchymal or amoeboid migration, effectively capturing features of early tumor metastasis for a single cell (Zaman 2007).

2.4.2.3. Reaction Diffusion Models

Reaction diffusion models treat a cluster of cells as spheroids and account for effects of nutrient and oxygen availability to drive cell movement, and cell-cell and cell-matrix adhesions to drive stability. Such models have been successfully used to model cancer invasion by obtaining model parameters from tumor growth and invasion in cell culture plates (Frieboes et al. 2006).

2.4.3. Cellular Potts Models

Cellular Potts are lattice based models in which a cell is defined by the connecting boundaries of a number of pixels of a square or hexagonal lattice (Szabó & Merks 2013). Each pixel retains a value σ that identifies the cell it belongs to, and cells are distributed by minimizing the energy of the system pixel by pixel– each pixels energy is transferred to a neighboring pixel to see if the overall system energy lowers. Adhesion and cortical tension are incorporated into a membrane tension parameter J, which is in terms of energy cost/unit length. Cell volume is constrained to reference volume and compressibility. The goal function of distinguishing between favorable (low u) and unfavorable (high u) configurations is described in Equation 2-14 and the function J for combinations of cells is given in Equation 2-15.

Equation 2-14

$$u = \sum_{\langle x,x' \rangle} J_{\sigma(x)\sigma(x')} + \lambda \sum_{p=1}^{N} \delta A_p{}^2$$

Equation 2-15

$$J_{pq=}\begin{cases} 0, for\ p = q \\ \alpha, for\ pq > 0\ and\ p \neq q\ )(intercellular\ boundary) \\ \beta, for\ pq = 0\ and\ p\ \neq q\ (free\ cell\ boundary) \end{cases}$$

Cellular Potts models have been successfully adapted to recreate phenomenon such as cell streaming (Szabó et al. 2010) and the leading front or ‘fingers’ of MDCK cells (Lee & Wolgemuth 2011). The latter of these suggests that wound healing can arise without substantial biochemical cues on the basis of force production at cellular boundaries. Cellular Potts models have also been expanded to include motile force and cohesion energy of individual cells and allow some cells to be biased “leader cells” with directional bias (Kabla 2012). This work has suggested that a 1% concentration is sufficient to coordinate a population migrating as a sheet. It also suggests that a criticial motile force or decrease of energy J is sufficient for coordinated sheet migration without specific signaling cues.

2.4.4. Self-Propelled Particle Models

Collectively moving entities such as locust swarms, bird flocks, fish schools, bacterial swarms and cells appear to share distinct characteristic traits and motion patterns. Computational modeling techniques employ Self Propelled Particle (SPP) models to recreate and interpret such motion patterns (Vicsek & Zafeiris 2012). These systems self-organize to behave as ‘larger than the sum of their parts’, displaying behavior in which the dominating influence on a unit within a collective is of the properties of the units around it (Meschede & Hallatschek 2013). These collectives are composed of similar individual units that move with nearly the same velocity, are subject to a varying magnitude of noise, and are capable of changing direction and interacting with each other to improve alignment. The simplest rule for each particle p ∈ {1,N} with a direction θ updating at time t+1 is represented in Equation 2-16. Here each particle has a constant speed, its direction at time t+1 is influence by an average of its neighbors within a certain interaction radius, and by a noise term η.

Equation 2-16

$$\theta_p^{t+1} = arg\left[\sum_{p\sim q} \overrightarrow{v_q}\right] + \eta$$

Simplistic SPP particle models have been expanded further to be more cell like (Grégoire et al. 2003) by scaling a direction dependent term and a force dependent term, but also maintaining constant speed according to Equation 2-17. Here α and β are scaling factors that contribute to the effect of neighbor directionality and forces, respectively. The force term between two cells is described in Equation 2-18. This updated model is able to account for attraction and repulsion between cells on a monolayer and recreate different phases of interaction.

Equation 2-17

$$\theta_p^{t+1} = arg\left[\alpha \sum_{q\sim p} \overrightarrow{v_q} + \beta \sum_{q\sim p} \overrightarrow{f_{pq}}\right] + \eta$$

Equation 2-18

$$\left[\vec{f}_{pq} = \vec{e}_{pq} \begin{cases} -\infty & if\ r_{pq} < r_c \\ \frac{1}{4}\frac{r_{pq} - r_e}{r_a - r_e} & if\ r_c < r_{pq} < r_a \\ 1 & if\ r_a < r_{pq} < r_0 \end{cases}\right]$$

A third adaptation of SPP models was used to recreate the densitiy dependent phase transition observed in fish keratocytes (Szabó et al. 2006). Here each particle p ∈ {1,N} moves with a constant velocity v_0 in the direction of a unit vector $\vec{n}(t)$. In addition cell pairs also experience an intercellular force $\vec{F}$, similar to Equation 2-18 but with a known repulsion and adhesion force scaled by distance, and a motility μ as described by Equation 2-19. This model is able to recreate a continuous phase transition from disordered to ordered by changing cell density alone.

Equation 2-19

$$\frac{d\vec{r}_p(t)}{dt} = v_0\vec{n}(t) + \mu \sum_{q=1}^{N} \vec{F}(\vec{r}_p, \vec{r}_q)$$

2.5. Conclusion

Research on cells in 3D is progressing faster than ever before with newer biomaterials and imaging modalities being explored and heightened interest in understanding fundamental questions and their relevance to disease. 3D *in vitro* studies of cancer cells typically involve invasion assays and immunohistochemistry to isolate key proteins and mechanisms (Carey et al. 2013). In 2D, epithelial cells and fish keratocytes (Rapanan et al. 2014) have been used as model systems to study the dynamic aspects of collective cell migration. In 3D, however, there is currently no model system to recapitulate the dynamics of collective cell migration. 3D tracking of human breast cancer cells *in vitro* has shown up to four cells undergoing multiple rotations, a process proved essential to acini formation (Tanner et al. 2012). While these findings are critical to the biomechanics of development, the experiment does not provide a model system for 3D cell migration. In development, morphogenesis, and wound healing, coordinated motion is a helpful aspect of a collective phenomenon, resulting in increasingly efficient processes. While cancer CCM has similarities to morphogenetic events, it is dysregulated; efficiency is difficult to define. The plasticity of migration modes in cancer also suggests that collective migration and invasion can emerge under critical conditions. A 3D model system for collective migration pertinent to cancer would ideally recreate some of the dysregulation and plasticity of migration modes. 2-dimensional computational models of collective cell motility are varied and specific in their descriptions of various phenomenon. However, adaptations to the third dimension for collective cell motility have to date been challenging. In addition, while 2D computational models have evolved in conjunction with analytical techniques for phenomenon observed in 2D monolayers, there is currently no established method to classify or quantify dynamics of 3D collective cell phenomenon.

CHAPTER 3. HETEROGENEITY WITHIN CELL POPULATIONS

3.1. Introduction

The intracellular milieu is a complex system comprising of many force transmitting components including active motor proteins such as actin and myosin as well as larger organelles. Placing inert particles within cells and tracking their fluctuations with nanometer resolution at high frequencies yields information of the timescales of intracellular activity (Wirtz 2009). Fluorescent beads have been successfully used as intracellular tracers in multiple two dimensional (2D) microrheology experiments; this study explores the efficacy of fluorescent beads as a tracer for three dimensional (3D) intracellular microrheology.

For an inert particle traced in a given medium, a Mean Squared Displacement (MSD) can be calculated as a function of time lag τ according to Equation 3-1 where T^{max} is the maximum timespan of observations.

<u>Equation 3-1</u>

$$\Delta r^2(\tau)_x = \frac{\sum_{t=1}^{t=T^{max}-\tau}(x_{t+\tau} - x_t)^2}{T^{max} - \tau}$$

This MSD, $\Delta r^2(\tau)$ scales linearly with time lag τ in a purely viscous fluid as expected by the Stokes Einstein relation. This is because the particle is purely diffusive and has no memory of its previous position at a given point. Conversely, in a purely elastic system, $\Delta r^2(\tau)$ is constant and does not scale with time lag, since an inert particle vibrates about the same point and always has a memory of where it existed before. The logarithmic slope of $\Delta r^2(\tau)$ is thus an indication of the material properties of the intracellular milieu. For viscoelastic media, the logarithmic slope is expected to have a value between 0 (purely elastic) and 1 (viscous). For cells with active ATP driven motor proteins, this slope values greater than 1. Thus even within active systems such as cells, the logarithmic slope of $\Delta r^2(\tau)$ contains information about timescales of intracellular motor activity, providing insight into intracellular protein fluctuations and turnover rates.

3.2. Methods

3.2.1. Cell Culture

Cells are cultured in media supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin and propagated in 2D monolayers at 37°C, 5% $C0_2$, and ~ 70% humidity. Fibrosarcoma cells from the line HT1080 are cultured in EMEM media; Neonatal Human dermal Fibroblasts (NHF) are cultured in DMEM media; and osteosarcoma cell line U20S is cultured in RPMI media. All cell culture components are purchased from ATCC, Manassas, VA. When fluorescence is desired, media are supplemented with CellTracker^{TM} (Invitrogen, Carlsbad, CA) of the desired wavelength. HT1080 and NHF cell lines are a model system for non-cancerous and cancerous fibroblasts with clear differences in morphology and physicality (Schwartz et al. 2013), and U20S cells are known to alter cellular response to PI3K drug (Fallica et al. 2012); thus these cell lines and conditions are explored and probed for differences in intracellular material properties.

3.2.2. Particle Delivery

Cell lines are cultured to 70% confluency in either a 3-cm or a 10-cm dish and then embedded with 1.0-µm carboxylate coated, fluorescent polystyrene tracer beads (Invitrogen, Carlsbad, CA). Beads are embedded using a ballistic particle delivery system (BioRad Laboratories, Hercules, CA) with grade Ultra High Purity Helium (Airgas) at 1900-2200 psi. Immediately after injection of beads, cell surfaces are rinsed with phosphate buffered saline (PBS). After a 30 minute rest and recovery period, cells are detached using 0.25% Trypsin/0.53 mM EDTA (ATCC) before re-suspending as desired.

3.2.3. 3D Collagen Culture

In order to seed cells in collagen, Type I collagen (BD Biosciences, San Jose, CA) is mixed with an equal volume of neutralizing solution (100 mM HEPES buffer (Fisher Scientific, Pittsburgh, PA) in 2× PBS, pH 7.3) (Wozniak & Keely 2005) before adding cells and supplemented growth medium to dilute collagen to the desired concentration (2-4 mg/ml). Collagen gels are allowed to form in the incubator at 37˚C, 5% CO_2, 70% humidity before supplementing the entire gel with additional growth medium. Cells are then imaged after ~24 hours of additional incubation. An image of a typical bead embedded inside a cell with 3D morphology is shown in Figure 3-1.

3.2.4. Drug and Serum-Starve Treatment

For HT1080 cells, drug exposure experiments are conducted with 10µM Y27632, a Rho-associated kinase inhibitor (Ishizaki et al. 2000). For the U20S cells, drug exposure experiments are conducted by adding 250 nM PI-103 drug, a PI3K kinase inhibitor (Fan et al. 2006). Serum-starved cells are incubated in media supplemented without FBS. To simulate dead cells, cells are exposed to 4% paraformaldehyde (PFA). All treatments are done for ~24 hours at at 37˚C, 5% CO_2, 70% humidity except for PFA which is conducted for 2 hours at room temperature.

3.2.5. Particle Imaging and Tracking

XY fields of view of identified beads within the 3D-embedded cells are imaged at 10-30 Hz for durations of 10-100 seconds. Imaging is performed using a 63× magnification oil immersion lens in a spinning disk confocal microscope (Leica CTR 6000). For the 3D cases, cells close to the glass surface are not considered. Total number of beads in any given condition ranges from 15-52, with ~1-2 beads from each cell.

Spot-tracking is conducted to find individual bead trajectories using either Imaris image analysis software (Bitplane, St. Paul, MN) or a custom Matlab (MathWorks, Natick, MA) algorithm provided by the Kilfoil group (Gao & Kilfoil 2009; Pelletier et al. 2009).

3.2.6. Mean Squared Displacement and Curve Fitting

Individual particle trajectories are first subject to Principal Component Analysis (PCA) to find the maximum and minimum axis of motion, henceforth referred to as Major axis and Minor axis respectively. For each individual axis of motion and each bead traced MSD is calculated according to Equation 3-1. Minor axis MSDs are fit to a one term power law described by Equation 3-2 using Matlab curve fitting toolbox. Only fits with an R^2 value >0.95 are considered for analysis.

Equation 3-2

$$\Delta r^2(\tau)_x = \alpha\tau^\beta + C$$

3.2.7. *Data Noise Reduction*

To address noise introduced into data from external sources (fan above microscope, slanted floor, freeway behind building etc.), sand is placed inside a double layered Ziploc freezer bag and secured under the camera while ensuring no camera fan vents are blocked.

3.2.8. *Statistical Analysis*

Populations of Major and Minor Axis MSD at $\tau = 1s$ and at $\tau = 100ms$ are compared between pairs of datasets from conditions of similar cell types (HT1080 with NHF, U20S with and without drug at different collagen concentrations, HT1080 multiple conditions). Unequal variances t-tests are conducted between pairs of samples with a p-value of 0.05. The null hypothesis is that the datasets come from normal populations with unequal means and unequal variances.

3.3. **Results**

3.3.1. *Sand absorbs some external noise*

Initial data shows an obvious noise frequency (2-4 Hz, determined empirically) superimposing data. Since these experiments are dependent on detecting Brownian motion, post-processing the data to filter the signal is not considered an adequate solution. Balancing the microscope table, ensuring no detachment of collagen gels, and conducting the experiment when the vent directly above the microscope did not significantly improved the observed noise. Signal processing methods are also not able to remove this noise, because the noise frequency itself fluctuates between 2-4 Hz. The camera essentially operates as a heavy weight held by a long beam that attaches to the microscope; thus relieving the tension with a vibration absorber proved an effective solution, with sand being the most cost-efficient and abundant of vibration absorbers. As Figure 3-2 illustrates, the observable waveform of noise is no longer visibly apparent when a bag of sand is placed under the microscope camera.

3.3.2. Higher variation within cell populations than between

MSD data of beads from U20S cells embedded in 2 mg/ml, 3 mg/ml and 4 mg/ml concentration collagen gels with drug PI103, is displayed in Figure 3-3, and MSD distributions at τ values of 1s and 0.01s are depicted in Figure 3-4. Similarly, MSD data of beads from HT1080 and NHF cells in 2D, 2 mg/ml and 3.5 mg/ml concentration collagen gels is displayed in Figure 3-5, and MSD distributions at τ values of 1s and 0.01s are depicted in Figure 3-6.

Pairwise samples that passed the t-test for these conditions are listed in Table 3-1 for $\tau = 100$ ms and Table 3-2 for $\tau = 1$s. Although some conditions reject the null hypothesis with a p-value <0.05, the Pr(H) value indicates the probability of the observation if the hypothesis is true. These extremely low Pr(H) values imply that there is likely insufficient data to test for significant difference between samples.

Since the results indicate that there is a larger variance within individual populations than between them, a single cell line is tested in multiple conditions. HT1080 cells are then tested with a sampling time of ~100 seconds, although data is only retained up to τ values of 10 seconds, thus increasing the sampling by tenfold. The data for HT1080 cells in collagen, with serum-starved media, drug Y27632 and PFA is in Figure 3-7 and MSD distributions at τ values of 1s and 0.01s are depicted in Figure 3-8. Despite smoother lines and larger sample size, Pr(H) values of pairs of conditions that rejected the null hypothesis do not increase, as illustrated in Table 3-1 and Table 3-2**.**

One term power fits according to Equation 3-1 are attempted for all datasets presented, but only retained if R^2 values > 0.95. With this restriction, datasets in which beads are imaged for ~100 seconds are the only ones in which individual bead data is retained. A plot of coefficient α compared with logarithmic slope β for each bead in three conditions of HT1080 cells is shown in Figure 3-9. Here α is a qualitative measure of relative thermal energy in any given bead, and β ranges from 0-2, with 0 indicating an elastic medium, 0-1 indicating a viscoelastic medium, and 1-2 indicating an active medium. It is apparent that HT1080 cells are generally very active, but serum starving and Rho-kinase inhibition decreases level of motor activity. However these methods do not achieve complete inhibition of motors. Thus this study shows that the material properties of cells vary between two orders of magnitude within individual populations in 3D, a larger variance than between select populations of cells with clear established differences.

3.4. Discussion

These results show that the intracellular milleu of cells is highly heterogenous in 3D within individual cell populations. This heterogeneity makes different cell-populations indistinguishable when the data is acquired with fluorescent tracers, since fluorescent tracers or beads offer a limited number of data points. This can be overcome by studying motion of inherent mitochondria (Mak et al. 2014) in 3D, thus obtaining much larger datasets than possible with embedded fluorescent beads.These results also show that cells in 3D are not purely viscoelastic and their material properties are influenced by active matter. This can be overcome by combining active and passive microrheology to decouple the effects of inherent fluctuations and active motors (Guo et al. 2014).

The results of this study imply that the variation within each individual cell type is far greater than between different cell types, especially when considering human cancerous cells. This is especially observable in the diversity of all MSD plots, and further proved by the inability of statistical tests to resolve data populations, and by the overlapping of α and β correlation plot populations. The heterogeneity in material properties within a single cell population possibly contributes to heterogeneity of biochemical and physical properties as well and may be a factor driving the supracellular polarity of cell collectives.

3.5. Tables

Condition 1	**Condition 2**	**Pr(H)**
U20S, 2 mg/ml	U20S, 3 mg/ml, + PI103	0.001>
U20S, 2 mg/ml	U20S, 4 mg/ml	0.001>
U20S, 2 mg/ml	U20S, 4 mg/ml, + PI103	0.001>
U20S, 2 mg/ml, + PI103	U20S, 3 mg/ml, + PI103	0.001>
U20S, 2 mg/ml, + PI103	U20S, 4 mg/ml	0.001>
U20S, 2 mg/ml, + PI103	U20S, 4 mg/ml, + PI103	0.001>
U20S, 3 mg/ml	U20S, 3 mg/ml, + PI103	0.001>
U20S, 3 mg/ml	U20S, 4 mg/ml	0.008
U20S, 3 mg/ml	U20S, 4 mg/ml, + PI103	0.005
HT1080, 2D	HT1080, 3.5 mg/ml	0.017
HT1080, 2D	NHF, 2 mg/ml	0.001>
HT1080, 2D	NHF, 3.5 mg/ml	0.001>
HT1080, 2mg/ml	HT1080, 3.5 mg/ml	0.026
HT1080, 2mg/ml	NHF, 2 mg/ml	0.001>
HT1080, 2mg/ml	NHF, 3.5 mg/ml	0.034
HT1080, 3.5 mg/ml	NHF, 2D	0.023
HT1080, 3.5 mg/ml	NHF, 3.5 mg/ml	0.002
NHF, 2D	NHF, 2mg/ml	0.001>
NHF, 2D	NHF, 3.5 mg/ml	0.001>
NHF, 2mg/ml	NHF, 3.5 mg/ml	0.001>
HT1080, 3.5 mg/ml, 100 s sample	HT1080, 3 mg/ml, ss, 100 s sample	0.002
HT1080, 3.5 mg/ml, 100 s sample	HT1080, 3 mg/ml, Y27632, 100 s sample	0.001>
HT1080, 3.5 mg/ml, 100 s sample	HT1080, 3 mg/ml, PFA, , 100 s sample	0.001>

Table 3-1: Results of T-Test for $\tau = 100$ ms

Pairwise data that reject the null hypothesis with 95% certainty as indicated by the unequal variances t-test for $\tau = 100$ms. Pr(H) for these cases is the likelihood of the observation if the hypothesis is correct.

Condition 1	**Condition 2**	**Pr(H)**
U20S, 2 mg/ml	U20S, 3 mg/ml	0.032
U20S, 2 mg/ml	U20S, 3 mg/ml, + PI103	0.001>
U20S, 2 mg/ml	U20S, 4 mg/ml	0.003
U20S, 2 mg/ml	U20S, 4 mg/ml, + PI103	0.001>
U20S, 2 mg/ml, + PI103	U20S, 3 mg/ml	0.001
U20S, 2 mg/ml, + PI103	U20S, 3 mg/ml, + PI103	0.001>
U20S, 2 mg/ml, + PI103	U20S, 4 mg/ml	0.001>
U20S, 2 mg/ml, + PI103	U20S, 4 mg/ml, + PI103	0.001>
U20S, 3 mg/ml	U20S, 3 mg/ml, + PI103	0.001
U20S, 3 mg/ml	U20S, 4 mg/ml, + PI103	0.004
HT1080, 2D	HT1080, 3.5 mg/ml	0.006
HT1080, 2D	NHF, 2 mg/ml	0.001>
HT1080, 2mg/ml	HT1080, 3.5 mg/ml	0.008
HT1080, 2mg/ml	NHF, 2 mg/ml	0.001>
HT1080, 3.5 mg/ml	NHF, 2D	0.003
HT1080, 3.5 mg/ml	NHF, 3.5 mg/ml	0.005
NHF, 2D	NHF, 2mg/ml	0.001>
NHF, 2mg/ml	NHF, 3.5 mg/ml	0.001>
HT1080, 3.5 mg/ml, 100 s sample	HT1080, 3 mg/ml, ss, 100 s sample	0.03
HT1080, 3.5 mg/ml, 100 s sample	HT1080, 3 mg/ml, PFA, , 100 s sample	0.001>
HT1080, 3 mg/ml, ss, 100 s sample	HT1080, 3 mg/ml, PFA, , 100 s sample	0.028
HT1080, 3 mg/ml, Y27632, 100 s sample	HT1080, 3 mg/ml, PFA, , 100 s sample	0.011

Table 3-2: Results of T-test for τ = 1s

Pairwise data that reject the null hypothesis with 95% certainty as indicated by the unequal variances t-test for τ = 100ms. Pr(H) for these cases is the likelihood of the observation if the hypothesis is correct.

3.6. Figures

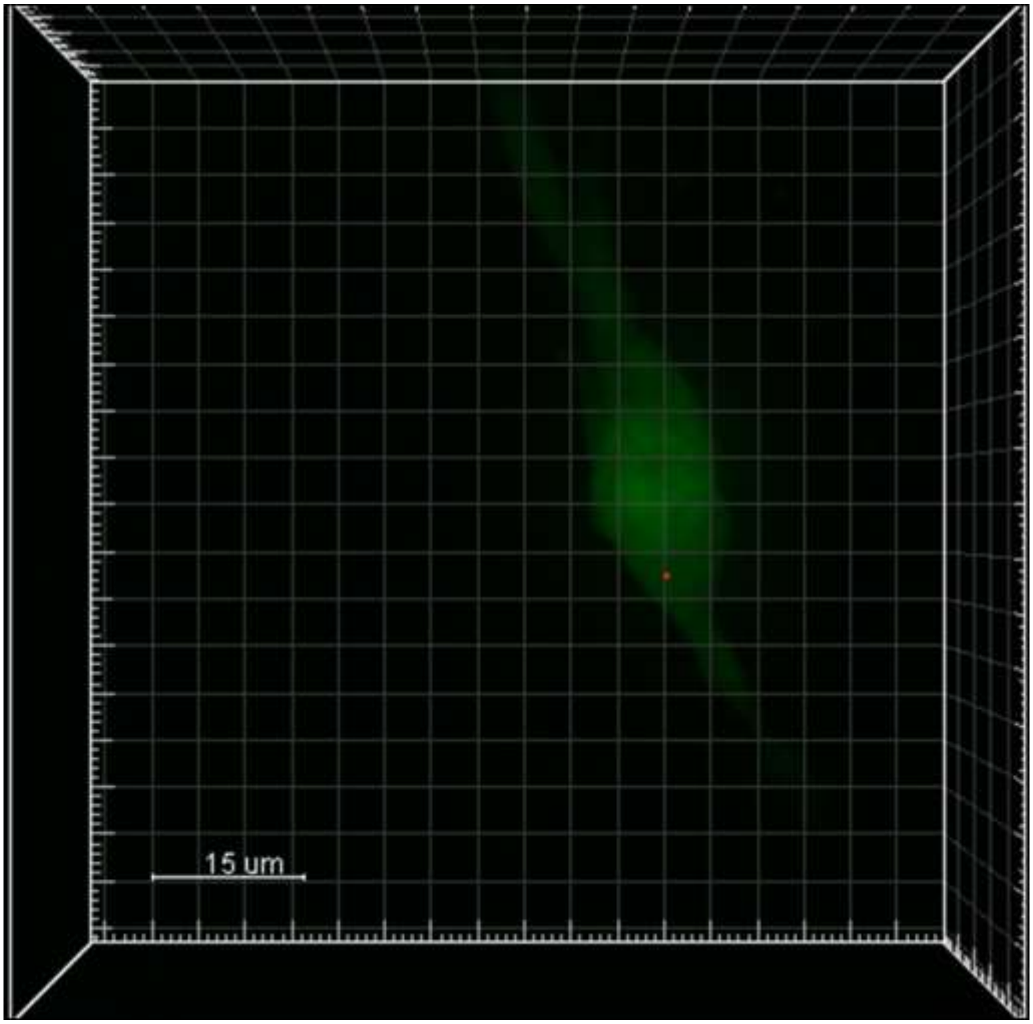
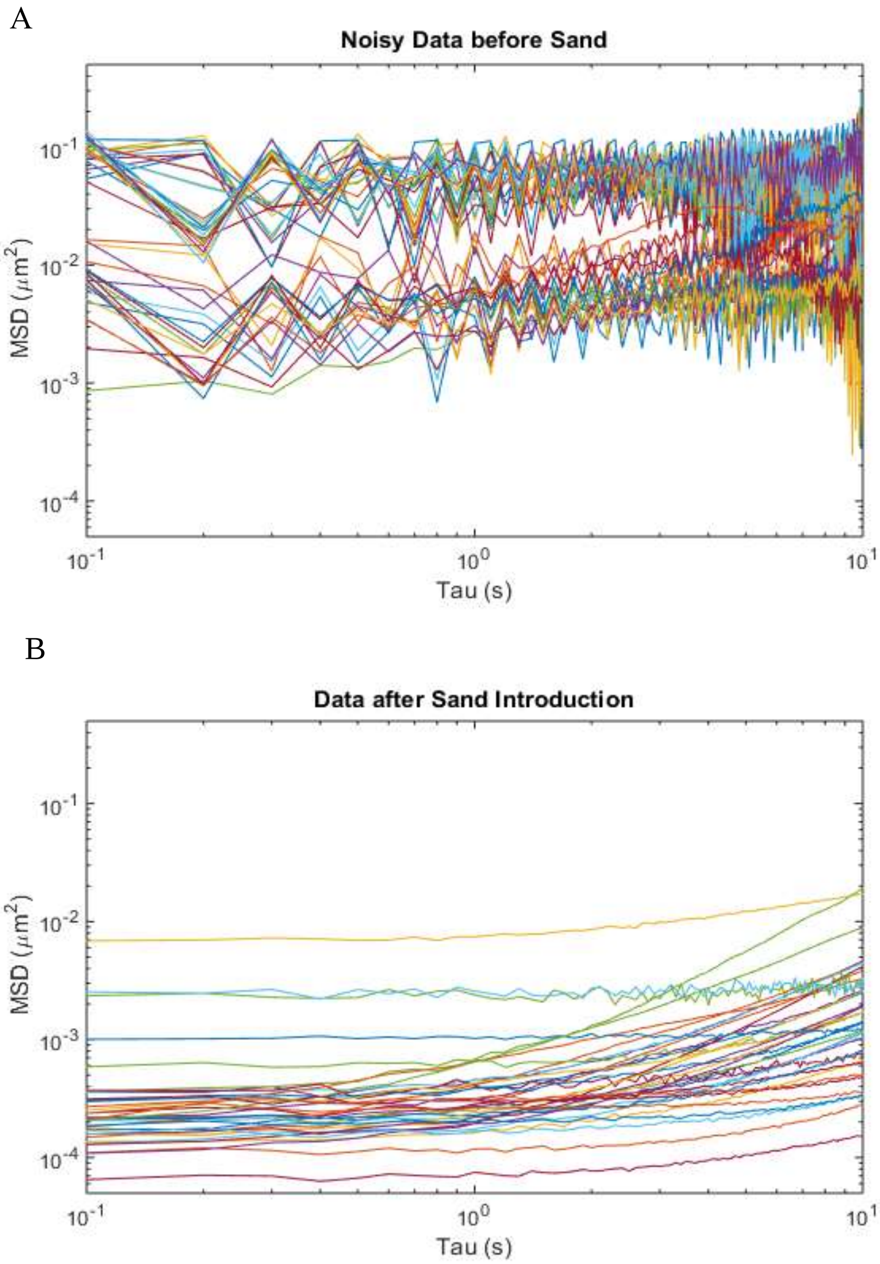


Figure 3-1: Image of a Single Bead inside a 3D cell

This image shows an HT1080 cell embedded within a 3 mg/ml collagen matrix with a single 1 μm diameter red fluorescent bead embedded inside.



<u>Figure 3-2: Effect of Sand as a Vibration Absorber</u>

A: Major axis MSD data acquired from beads within cells. B: typical Major axis MSD data acquired from beads upon adding a bag of sand under the camera.

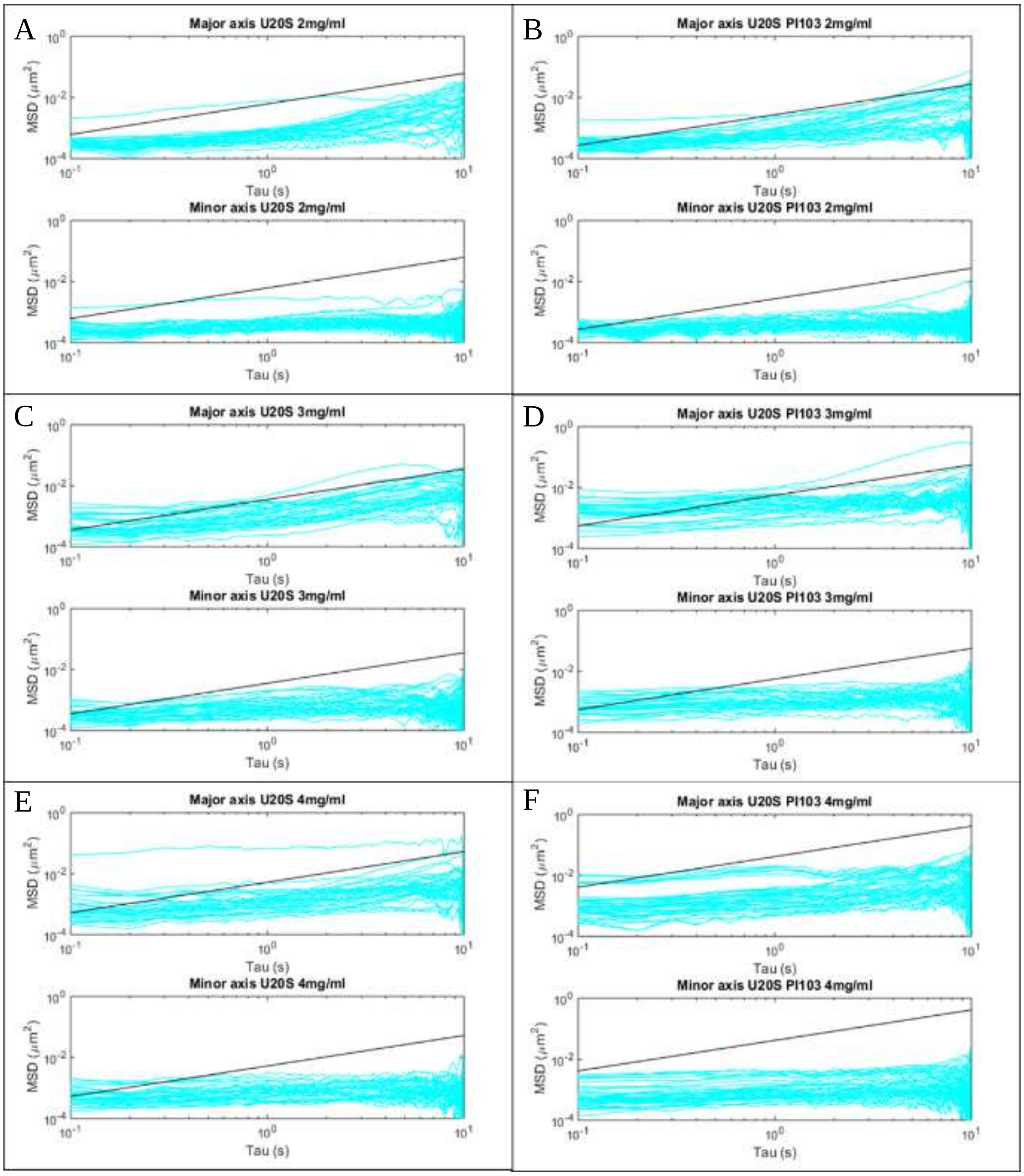
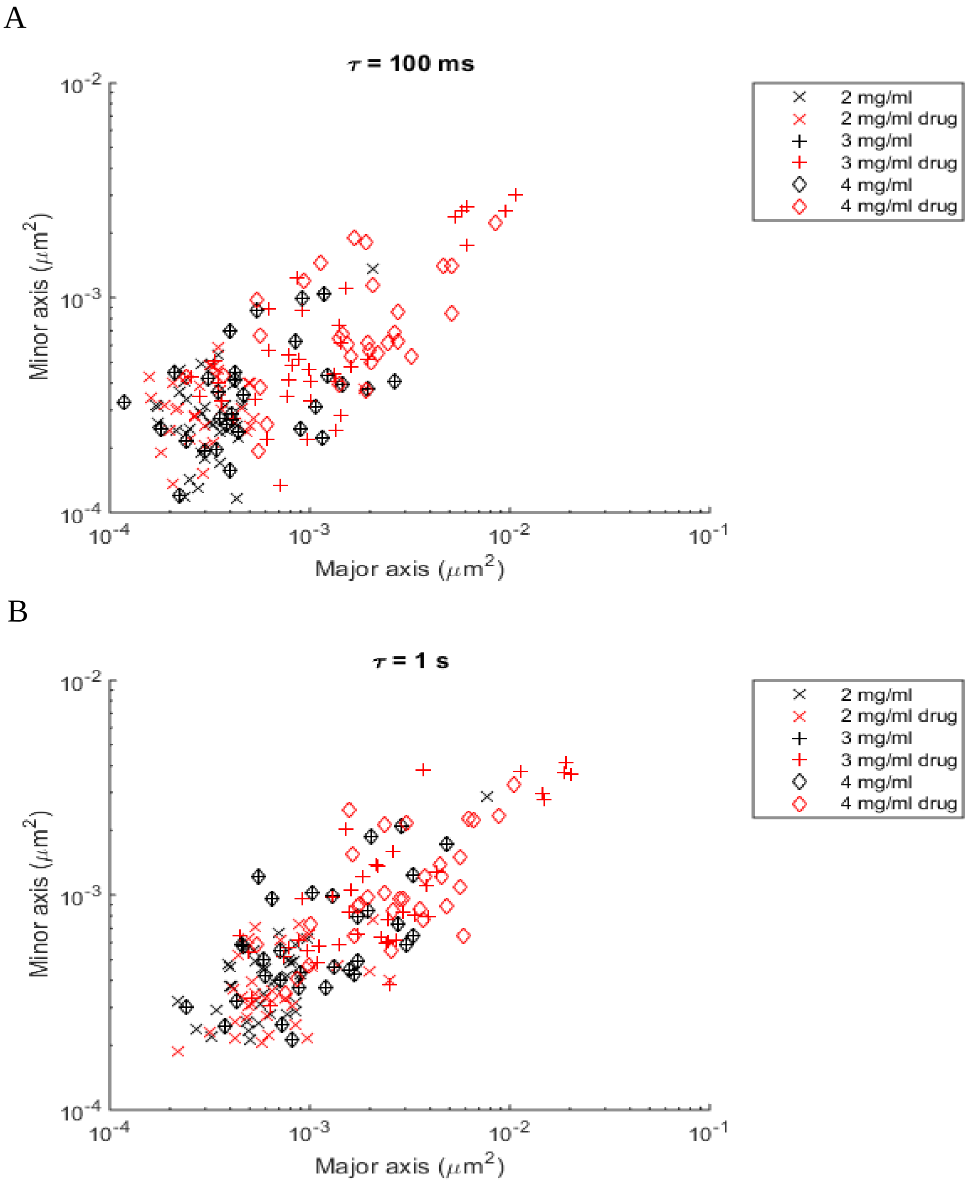


Figure 3-3: Mean Squared Displacement of U20S Cells

MSD of Major and Minor axes for U20S cells A. 2 mg/ ml collagen, B. 2 mg/ml collagen + PI103 drug, C. 2 mg/ ml collagen, D. 3 mg/ml collagen + PI103 drug, E. 4 mg/ ml collagen, F. 4 mg/ml collagen + PI103 drug. Black lines depict a logarithmic slope of 1. A slope of 0 indicates an elastic material, a slope of 1 indicates a viscous material, and the region in between indicates viscoelasticity. A slope >1 indicates the presence of active motor proteins.



<u>Figure 3-4: MSD of U20S cells at τ values of 1s and 100ms</u>

Major Minor MSD distributions at τ values of A. 100 ms and B. 1 s along the major and minor axis of motion of all existing beads for the conditions of U20S cells in 2(× symbol), 3(+), and 4(◊ symbol) mg/ml collagen concentration. The presence of drug PI103 is indicated by red colored symbol. The data are not able to resolve into significantly different populations.

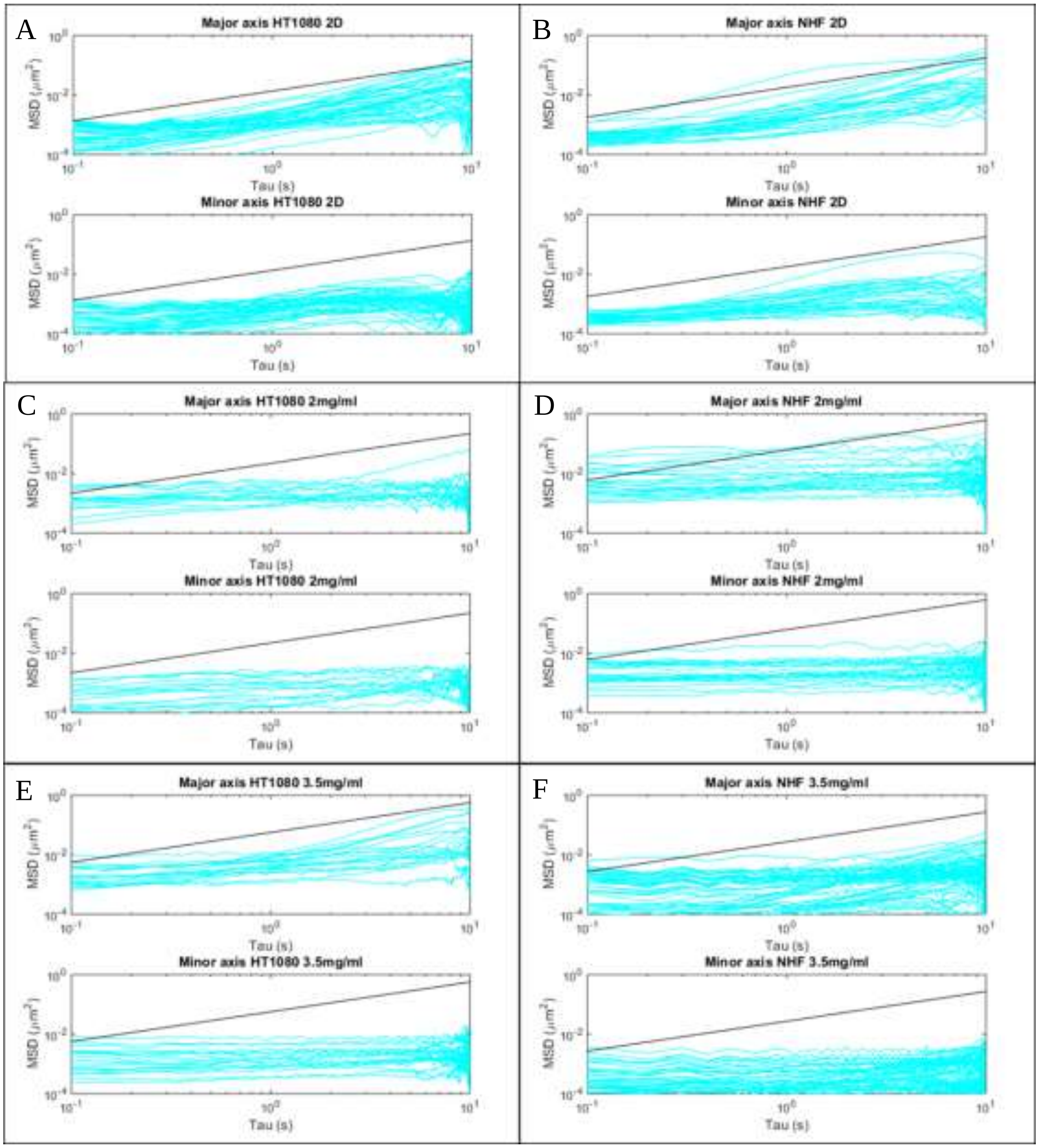


Figure 3-5: Mean Squared Displacement of HT1080 and NHF Cells

MSD of Major and Minor axes for A. HT1080 in 2D, B. NHF in 2D, C. HT1080 in 2 mg/ml, D. NHF in 2 mg/ml, E. HT1080 in 3.5 mg/ml, and F. NHF in 3.5 mg/ml collagen. Black lines depict a logarithmic slope of 1. A slope of 0 indicates an elastic material, a slope of 1 indicates a viscous material, and the region in between indicates viscoelasticity. A slope >1 indicates the presence of active motor proteins.

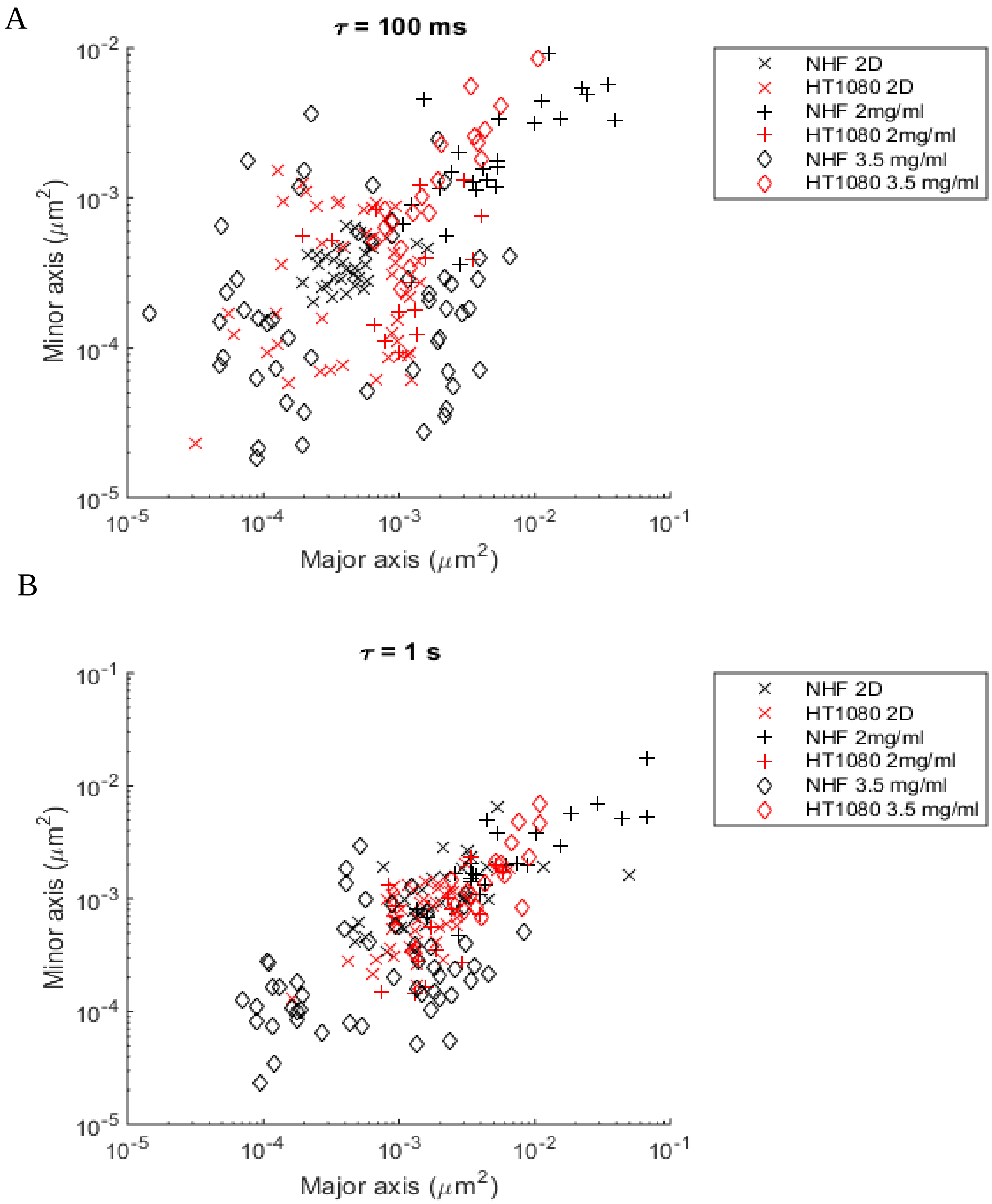


Figure 3-6: MSD of HT1080 and NHF cells at τ values of 1s and 100ms

Major Minor MSD distributions at τ values of A. 100 ms and B. 1 s along the major and minor axis of motion of all existing beads for the conditions of HT1080(red) and NHF (black) cells in 2D(× symbol), 2 mg/ml (+ symbol) and 3.5(◊ symbol) mg/ml collagen concentration. The data are not able to resolve into significantly different populations.

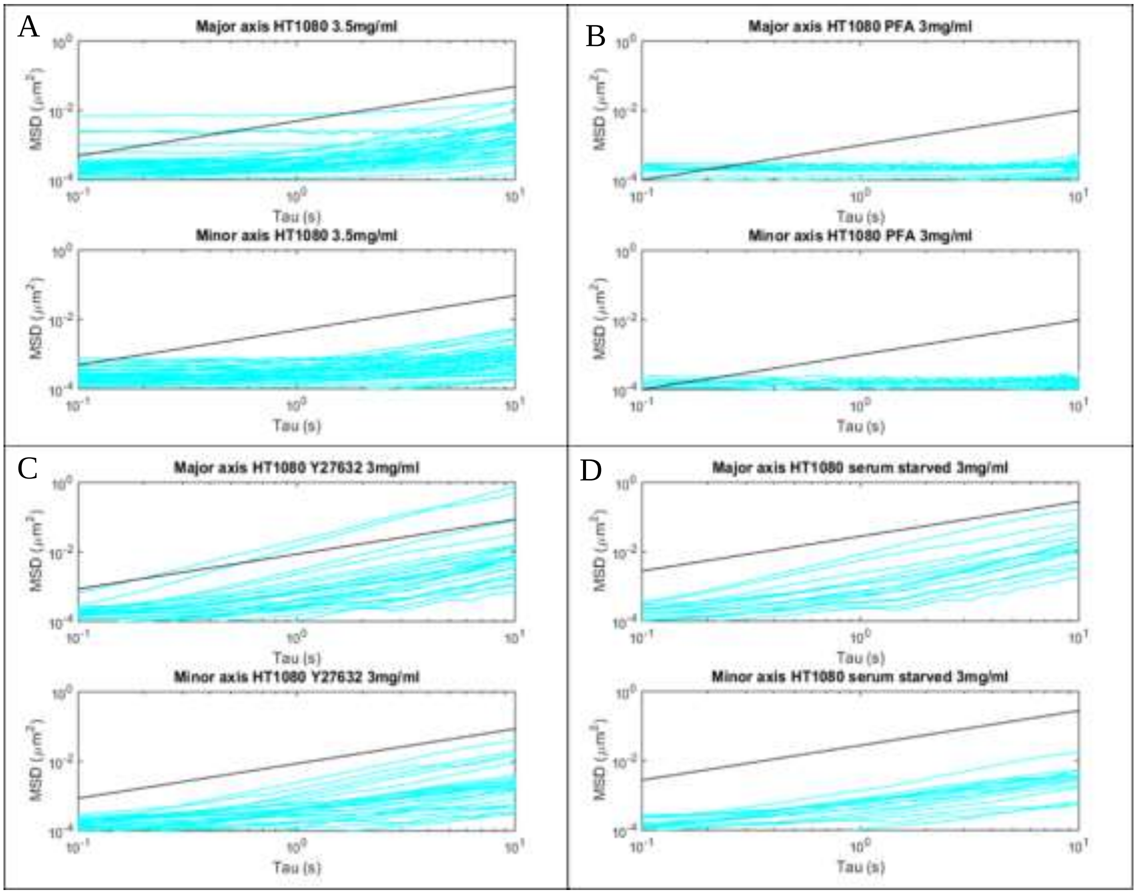


Figure 3-7: MSD for HT1080 Cells in various conditions

MSD of Major and Minor axes for HT1080 cells in 3-3.5 mg collagen with the addition of serum starving, drug Y27632, and PFA. Images were acquired over ~100 seconds and the last decade was eliminated to obtain clear slopes. Black lines depict a logarithmic slope of 1.A slope of 0 indicates an elastic material, a slope of 1 indicates a viscous material, and the region in between indicates viscoelasticity. A slope >1 indicates the presence of active motor proteins

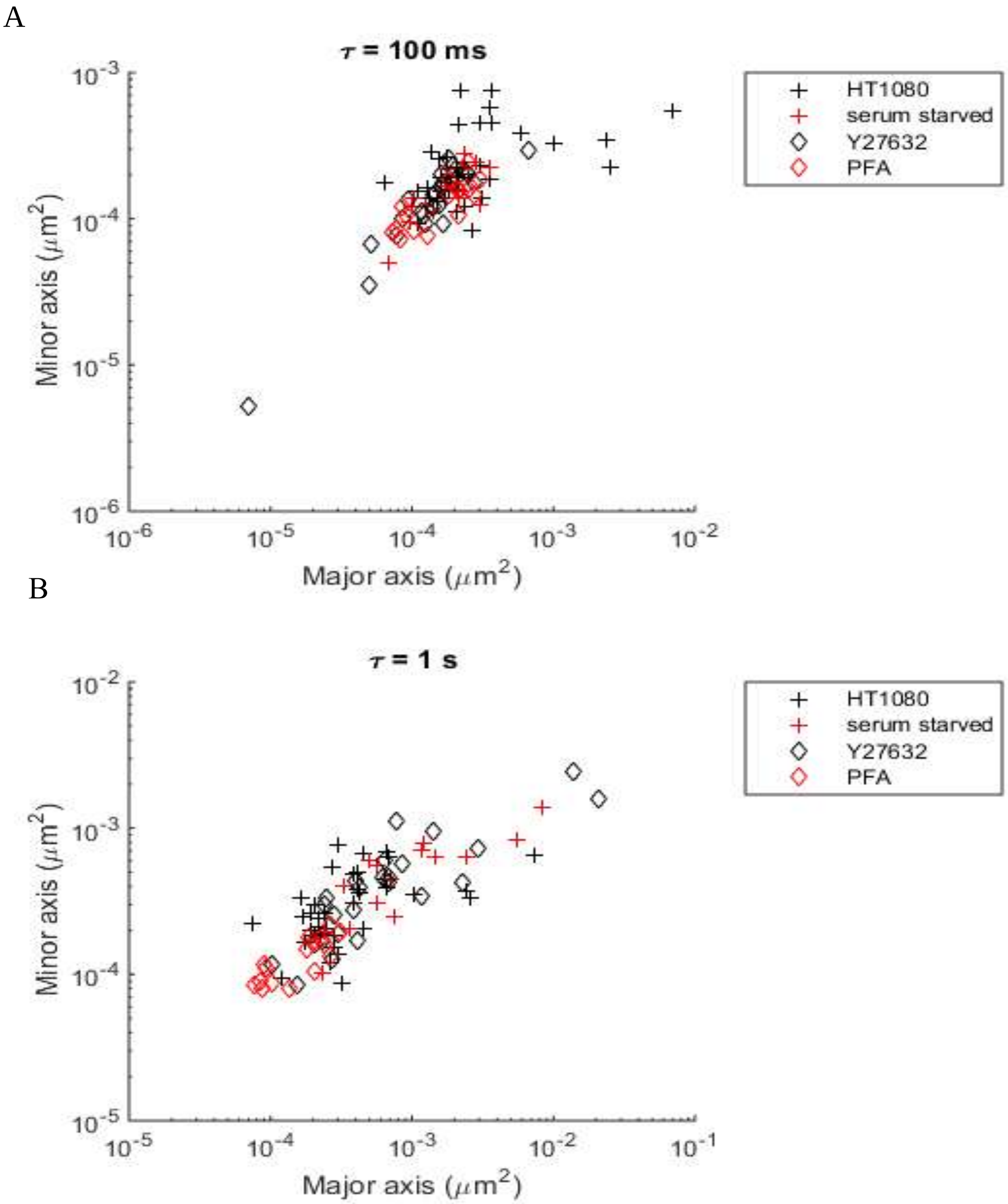


Figure 3-8: MSD of HT1080 cells at τ values of 1s and 100 ms

Major Minor MSD distributions at τ values of A. 100 ms and B. 1 s along the major and minor axis of motion of all existing beads for the conditions of HT1080 cells(black + for no addition) in 3-3.5 mg/ml collagen with serum starving (red + symbol), drug Y27632 (black ◊ symbol) and PFA (red ◊ symbol). These datasets were all acquired over ~100 ms. The data are not able to resolve into significantly different populations.

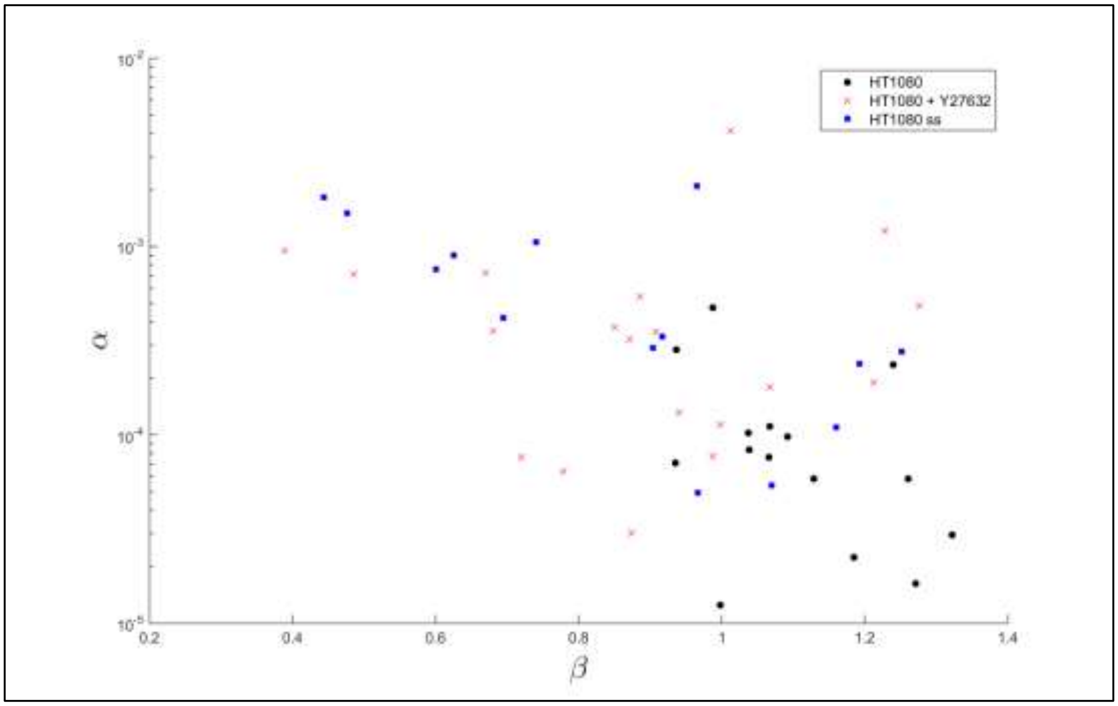


Figure 3-9: Coefficient vs. Logarithmic Slope for HT1080 cells

For the three data sets of HT1080, HT1080 + Y27632 drug, and serum starved (ss) cells, each point shows curve fitting parameters α, the scaling coefficient and β, the logarithmic slope. The black dots are from beads in HT1080 cells in a 3.5 mg/ml collagen matrix. The red crosses are from beads in HT1080 cells subjected to Y27632 treatment in a 3 mg/ml collagen matrix The blue squares are from beads in serum-stared HT1080 cells in a 3 mg/ml collagen matrix. β values between 0-1 indicate a viscoelastic regime. β values >1 indicate presence of active motor proteins and motile components.

CHAPTER 4. EXPERIMENTAL MODEL FOR COLLECTIVE CELL MOTILITY

4.1. Introduction

Currently, studies of the inter-cellular dynamics of collective cell motion are limited to two dimensional (2D) monolayer experiments (Park et al. 2015). Cells must balance adherence and motility to maintain organized coherent motion (Méhes & Vicsek 2014). Monolayer based cell-tracking assays have revealed density dependent phase transitions (Szabó et al. 2009), substrate dependence (Doxzen et al. 2013), and the forces driving Collective Cell Motility (CCM) (Brugués et al. 2014; Shaw & Martin 2009). Heterogeneity emerges within groups of cells exhibiting collective behavior– key molecules are upregulated to form leader cells that are in the front of sheets and essential to driving collective motility; removal of these leader cells disrupts collective sheet migration (Yamaguchi et al. 2015).

In 2D, epithelial cells and fish keratocytes have been used as model systems to study the dynamic aspects of collective cell migration (Rapanan et al. 2014). 3D *in vitro* studies of cancer cells typically involve invasion assays and immunohistochemistry to identify key proteins and mechanisms (Carey et al. 2013). Establishing robust models of collective cell migration in 3D has proved difficult. Tracking of human breast cancer cells *in vitro* has shown up to four cells in 3D undergoing multiple rotations, a process essential to acini formation (Tanner et al. 2012). However, compared to 2D counterparts, no systematic experimental model to the study of 3D dynamics of cell migration has been introduced.

This study presents an experimental model system for 3D collective migration

(Sharma et al. 2015) using mammalian cell cohorts and collagen matrices that provides dynamic information of individual cells by visualizing and tracking individual nuclei within distinct cellular clusters.

4.2. Experimental Methods

4.2.1. Cell Culture

MDCK Type II epithelial cells are propagated in monolayers at 37°C, 5% CO_2, and 70% humidity; monolayers are cultured in DMEM media supplemented with 10% FBS and 1% penicillin/streptomycin. Cells are stably transfected to express Nuclear Localization Signal (NLS) fused to Green Fluorescent Protein (GFP). Transfection is accomplished by a GFP-NLS plasmid (Clontech, Takara Bio, Japan) of Lipofectamine 2000 (Invitrogen, Life Technologies, Grand Island, NY). To maintain fluorescence, 0.5 mg/ml G418 (Fisher Scientific, Pittsburgh, PA) is added to the media; Fluorescence Activated Cell Sorting (FACS) selects for the brightest 1% of cells.

4.2.2. 3D Cluster protocol

Single cell suspensions are formed by immersing cells in media after detaching from monolayers; cells are passed through a 40 µm cell strainer (BD Biosciences, San Jose, CA); cells are then seeded onto a 10 cm diameter Ultra Low Attachment Dish (Corning, Corning NY) with 10 ml media. After ~48 hours, clusters are extracted by passing the solution through either a 100 µm or a 70 µm cell strainer followed by a 40 µm cell strainer, retaining clusters of 10-20 cells. These are re-suspended and centrifuged at 800 rpm, and then immersed in a collagen solution for 3D culture and imaging.

4.2.3. 3D collagen protocol

A 2 mg/ml collagen dilution is obtained by mixing equal volumes of collagen Type 1 stock (BD Biosciences, San Jose, CA) solution and neutralizing buffer (100mM Hepes in 2× PBS, pH 7.3) with PBS. Cell clusters are added to the 2 mg/ml collagen solution; 500 µl of this cluster-collagen suspension is seeded onto several wells of a 24-well plate (MatTek, Ashland MA). The plates are incubated at 37°C, 5% CO_2 and ~ 70% humidity for 2 hours until the collagen has polymerized, after which ~1-2 ml of growth media is added to each well. When necessary, FluoSpheres® Carboxylate-Modified Microspheres in 1.0 µm (Invitrogen) with red fluorescence (580/605) are diluted to ~ 10^8 beads/ml collagen.

4.2.4. Image Acquisition

Images are acquired with a DMI600B Microscope (Leica, Solms, Germany) and ImagEM EM-CCD Camera (Hamamatsu Photonics, Hamamatsu, Japan) using a Spinning Disk Confocal setup (Yokogawa, Tokyo, Japan). Micro-Manager 1.4 Software (Edelstein et al. 2014) employs a 10× 0.3 NA objective lens to image multiple ~560 × 560 × 100 μm^3 fields of view. 3D stacks are acquired in the XY plane with a Z-step of 4 µm, every 10 minutes, for 48-72 hours. Since cells can sense the substrate beneath the 3D matrix from inside the collagen gel (Leong et al. 2010), acquisition and analysis is restricted to cells clusters located >100 µm from the glass bottom. 3D morphology is further verified before tracking and analysis.

4.2.5. Trajectory Data for Individual Cells within Cohorts

Experimental results acquired from confocal microscopy yield XY images across multiple Z-frames and time frames. Figure 4-1 illustrates XY fields of view with Z-slices superimposed for typical experiments. This XYZT data is essentially a 4D matrix comprising of brightness values of the size 512 x 512 x N_z x N_T where N_Z is the number of images in the Z-stack and N_T is the number of time-points imaged. In order to have a model capable of quantitative analysis, this matrix needs to be converted to values of individual cells

4.2.6. Feature Finding Optimization

Most commercially available software is unable to track individual nuclei unless the cytoplasm of the cell is also fluorescent. Since the experimental setup only allows for visualization of cellular nuclei, a 3D feature finding algorithm designed for concentrated fluorescent spheres in colloidal systems is optimized for nucleus tracking (Gao & Kilfoil 2009). For lack of applicable statistical methods, the feature finding algorithm is optimized within reasonable limits to range in which varying the input parameters does not have a significant impact on output number of features (N_F) as described by Equation 4-1.

<u>Equation 4-1</u>

$$\frac{dN_F}{dParameter} \cong 0$$

4.2.6.1. Diameter and Mask

One pixel in the X or Y dimension of the image equals about 1.13 µm, whereas the z-dimension of the corresponding voxel is 4 µm. Even with a confocal microscope, nuclei features appear distorted in the Z dimension, resulting in a larger apparent Z diameter. The feature finding algorithm operates by integrating voxels of a volume as indicated by an input *mask* to find features of an approximate volume indicated by an input *diameter*. The integrating volume *mask* is therefore always larger than the *diameter*. The feature finding algorithm is able to distinguish between sizes of *mask* and *diameter* in the X, Y and Z dimensions. Since the imaging technique is equivalent in the X and Y axis, the X and Y values of the *mask* and *diameter* are always the same, whereas the Z value is usually larger.

To inform a starting guess for the algorithm, the nucleus diameter is estimated to be ~8µm by measuring individual nuclei in ImageJ (Schneider et al. 2012). N_F identified in a given XYZ frame are plotted against time-points for various *diameter* and *mask* combinations as shown in Figure 4-2. These values impact the output N_F by ~80-100%, and therefore XY and Z values of the *diameter* and *mask* are explored independently to identify a region that satisfies Equation 4-1.

The effect of varying the *mask* without changing the *diameter* is explored in Figure 4-3 and suggests that *mask* values have negligible impact on output N_F, in fact it is likely that changing the Z-value of the *mask* is what affects N_F in Field of View (FOV) 1 and 3. The effect of changing only the Z-value of the *diameter* is shown in Figure 4-4. The effect of changing the XY-value of the *diameter* is shown in Figure 4-5. Within the range of values explored, changing either the XY-value or Z-value of the *diameter* does not alter the output N_F value >10% for an XY *diameter* of 8 and Z *diameter* of 5. The *mask* value is significantly larger and set at 13 for XY and 7 for Z. The algorithm automatically finds features separated by a minimum distance of half a *diameter.*

4.2.6.2. Threshold

The threshold value is a scaling factor applied to peaks of features, and accounts for biasing caused by brighter features. Testing multiple threshold values at multiple diameters shows that there is no significant impact of the threshold on output data as apparent in Figure 4-6. The threshold is henceforth conservatively set at 0.3 for all data analyses.

4.3. Tracking Algorithm Optimization

The tracking algorithm searches for nearest neighbors to find obvious candidates before making tracks for features that have undergone more motility. Two input parameters facilitate tracking: the minimum track duration min(T_{track}) and the maximum distance between consecutive time-points max($d_{consecutive}$). The percentage and error of features found retained in tracks for six datasets is displayed in Table 4-1. It is highly unlikely that even with drift a nucleus would move >5 pixels within 10 minutes, which is the span between consecutive time-steps in this case. For a max($d_{consecutive}$) of 5, a min(T_{track}) of 6 has the most feature retention with the least error across 6 datasets. Therefore these values are established for the tracking algorithm. Features are retained in a track if they are missing for up to 3 timeframes. This value is chosen because it is half the minimal track duration of 6 timeframes.

4.3.1. De-drifting and Filling Missing Frames

A de-drifting algorithm is employed to remove motion common to all features in the frame (Pelletier et al. 2009). Drift is smoothed and eliminated over 1-hour since externally induced stage drift occurs over the period of a few hours. After eliminating drift, any missing cells that skipped up to 3 timeframes are filled by employing a linear interpolation of their path along the missing frames. The drift eliminated for the six datasets presented here is shown in Figure 4-7.

4.3.1.1. Validating Lack of Intra-Cluster Flows

To verify that cluster motion is not occurring because of local flows, a separate experiment is conducted with fluorescent tracers or beads embedded in the collagen matrix. These show no sign of local flows as shown in Figure 4-8, in fact the matrix only moves in the vicinity of motile clusters.

4.3.2. Identifying Clusters from Trajectory Data

The tracking algorithm successfully converts Pixel-Matrix data to a list of variables of the form XYZTN where N is the individual ID of each cell. However this study aims to study cell collectives, and even within the same FOV different cell cohorts have different behavior. Thus a Clustering Algorithm is proposed to add an identifier value or cluster ID to the list of variables for each cell at each time-point. Cluster IDs are first assigned in a single frame, and the same cluster is then correlated across consecutive frames.

4.3.3. Assigning Clusters IDs for Each Time-point

An agglomerative hierarchical clustering algorithm is used to identify clusters in each field of view. A cluster is defined as a minimum of 2 cells positioned within a radius of interaction r_i of each other. The first cell in the frame is assigned a Cluster ID of 1. Then a search is performed to find all cells within distance r_i. These cells are also assigned a Cluster ID of 1. For all these newly identified cells, a similar search is performed to find all cells within distance r_i excluding cells that have already been assigned a Cluster ID. These cells are also assigned a Cluster ID of 1. The process is repeated until there is no cell within a distance r_i for the last cluster assigned an ID of 1. At this point the next cell is assigned a Cluster ID of 2, and so on until all cell clusters in the field of view have been assigned a unique Cluster ID. At the end of assigning Cluster IDs a search is performed for any Clusters with only 1 cells, and these are reassigned a Cluster ID of 0, indicating that they are single cells not in a cluster. A schematic for this process is depicted in the top panel of Figure 4-9.

4.3.4. Correlating Clusters Across Consecutive Time-points

Cluster ID's are assigned by the clustering algorithm based on positions of cells linearly along each axis. Therefore it is likely that as clusters moved, a cluster identified as Cluster 1 in one time-point is assigned as Cluster 2 or 3 in the next. Thus clusters are correlated across consecutive timepoints T_i and T_{i+1} according to the following rules:

1. For all clusters in T_i, the cluster in the T_{i+1} which has maximal overlap of same cells is found and assigned the corresponding identity.

2. If two or more clusters from T_i have maximal overlap with the same cluster in T_{i+1}, the identity of largest of these clusters in T_i is assigned to the merged cluster in T_{i+1}.
3. After all the clusters in T_i have been accounted for or merged, unassigned clusters in T_{i+1} are assigned a new identity.
4. If T_i has no clusters and T_{i+1} has clusters, all clusters in T_{i+1} are assigned new identities. This is to say once a cluster breaks and reforms, it is assigned a new identity.

These rules allow identification and retention of clusters that merge and break apart, and are not disrupted by a single cell detaching and disseminating from a cluster. A schematic of this cluster correlation algorithm is presented in Figure 4-9. This algorithm is applicable in any collective system with similar units, and only depends on the definition of interaction radius r_i. Once r_i is defined, this algorithm can convert data of the form XYZTN to XYZTNC where C is the Cluster Identifier.

4.3.5. Identifying interaction radius for Cellular Cohorts

For experimental cellular cohorts, the correct interaction radius r_i needs to be defined in order to form cohorts. Cells in 3D are smaller than their 2D counterparts- our average cohort diameter of 40-70 μm is on the order of 2D single cell lengths (Sepúlveda et al. 2013), but in 3D, it spans ~6 -15 nuclei. On the assumption that a 40 μm cluster is a sphere comprising six spherical cells, each cell would have a radius of 11 μm. However cells and clusters are hardly spheres, the cytoplasm may be extending far beyond the cells.

As a starting point the clustering algorithm is run for the 6 FOV's in Gel I and II respectively with r_i values from 20 to 50 μm with 5 μm spacing. The clustering algorithm is unstable and finds too many clusters at 20 μm. The distance distributions of 5 closest neighboring cells for all clusters identified for r_c values 25, 30, 35 and 40 μm are depicted in Figure 4-10. The median distance of 5 nearest neighbors is ~18 μm for all r_i values shown as expected. The outlier values likely arise from cells in the same cluster that are within 2 r_i values and included because the cell does not have 5 neighbors within a distance of r_i. Setting r_i to 25 μm results in some clusters identified uniquely that are otherwise absorbed into larger clusters. In order to obtain data with the most fidelity without grouping cells too far apart, r_i is set to 30 μm for these MDCK cellular cohorts.

4.4. Discussion

This study has presented a robust system that serves as an experimental model system for 3D collective cell motility. It provides a method to visualize and study the motion of individuals within cellular clusters in three-dimensions. These methods can be applied onto various cell-matrix combinations to extract cellular clusters of 4-30 cells spaced ~100-400 μm apart in any 3D-matrix with adhesive cells. Clusters close to the bottom (100 μm) that can sense glass are more proliferative and sink to the glass to form a 2D-morphology, thus the use of synthetic matrices with coatings that prevent sensing the glass bottom by pulling on the collagen will improve the experimental model system.

The data with embedded beads shown in Figure 4-8 verifies that local flows are not driving the motion of clusters observed, thus the clusters move without being pushed by external matrix deformations. This experimental set-up can also be used for 3D traction force microscopy (Franck et al. 2011) to evaluate the forces exerted by moving cell collectives.

The clustering algorithm presented here allows easy identification of collectives from a large population. Each cluster is assigned an identity and this identity is retained regardless of merging events or individual cells leaving the cluster. This identity can be employed to analyze a single cluster and compare various clusters to see any temporal or spatial dependence. It can also distinguish cells from different clusters from each other and enable analysis of cellular patters within clusters undergoing varoius behaviors. The clustering algorithm also has potential applications at any scale at which collective emerge, since the only variable needed is an interaction radius.

This work is a first step at multiplexing and expanding 3D cell motility assays to 3D collective motility assays. This research can be expanded to ask many questions- such as what are the traction forces generated by cell collectives, and what the effects of biochemical and mechanical perturbations on the system are on collective cell behavior. The methods presented can be expanded to collective systems of any scale including those with large populations and 3D freedom.

4.5. Tables

$\max(d_{consecutive})$	$\min(T_{track})$	Average Feature Retention	Error across 6 datasets
3	6	83.26%	2.68%
3	12	66.79%	4.54%
3	18	51.09%	7.84%
4	6	90.29%	1.43%
4	12	79.27%	2.81%
4	18	66.86%	5.89%
5	6	92.92%	1.13%
5	12	85.41%	1.80%
5	18	76.30%	4.23%
6	6	94.66%	0.57%
6	12	88.43%	1.93%
6	18	80.98%	3.03%
7	6	95.80%	0.48%
7	12	90.08%	1.70%
7	18	83.28%	3.47%

Table 4-1: Features Retained in Tracks for Parameter Combinations

The number of features found that are retained by the tracking algorithm for settings of $\max(d_{consecutive})$ and $\min(T_{track})$ are displayed here. $\max(d_{consecutve})$ is the maximum displacement in pixels a feature can move between consecutive timeframes. Similarly, $\min(T_{track})$ is the minimum track time of a feature.

4.6. Figures

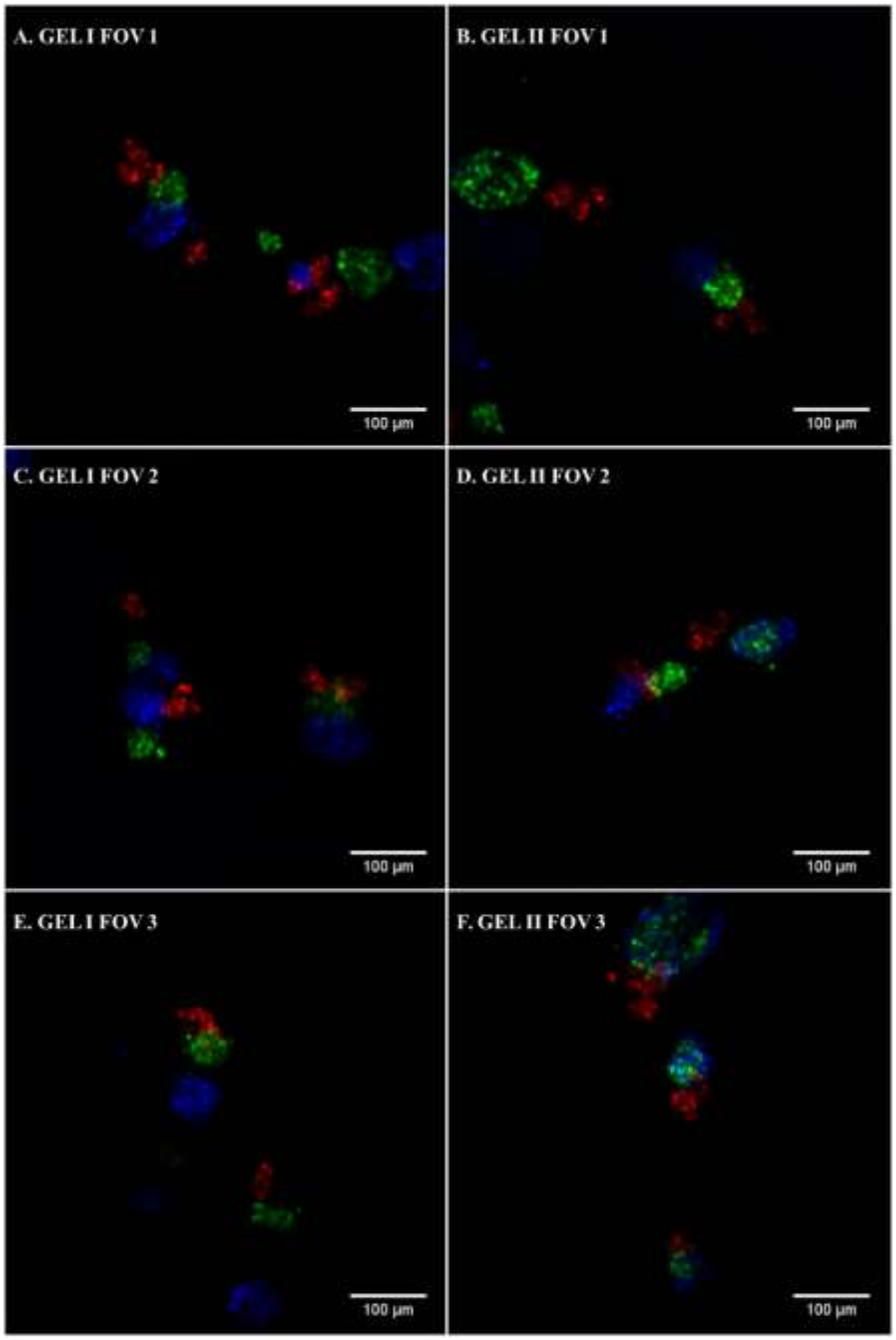


Figure 4-1: Sum of Z-slices of experimental data

Sum of all z-slices of experimental data at t = 0 (red), t = 24h (green) and t = 48h (blue) for 2 gels and 6 fields of view are superimposed. A. Gel I Field of View(FOV) 1. B. Gel II FOV 1 C. Gel I FOV 2 D. Gel II FOV 2 E. Gel 1 FOV 3 F. Gel II FOV 3

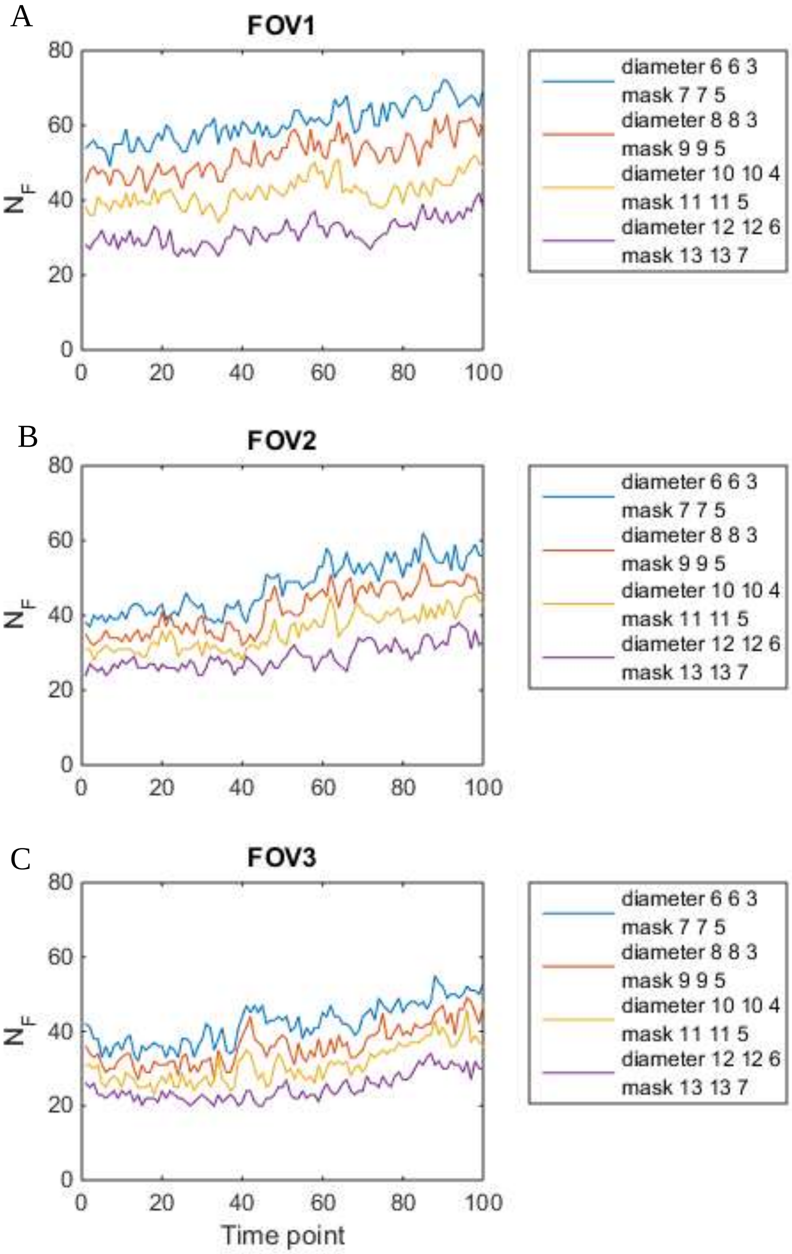


Figure 4-2: Dependence of Number of Features on Diameter and Mask Size

Changing values of diameter and mask together changes output number of features drastically. The number of features for each time point for A. FOV1, B. FOV2, and C. FOV3. Blue lines represent diameter (6 6 3) mask (7 7 5). Red lines represent diameter (8 8 3) mask (9 9 5). Yellow lines represent diameter (12 12 6) mask (11 11 5). Purple lines represent diameter (12 12 6) mask (13 13 7).

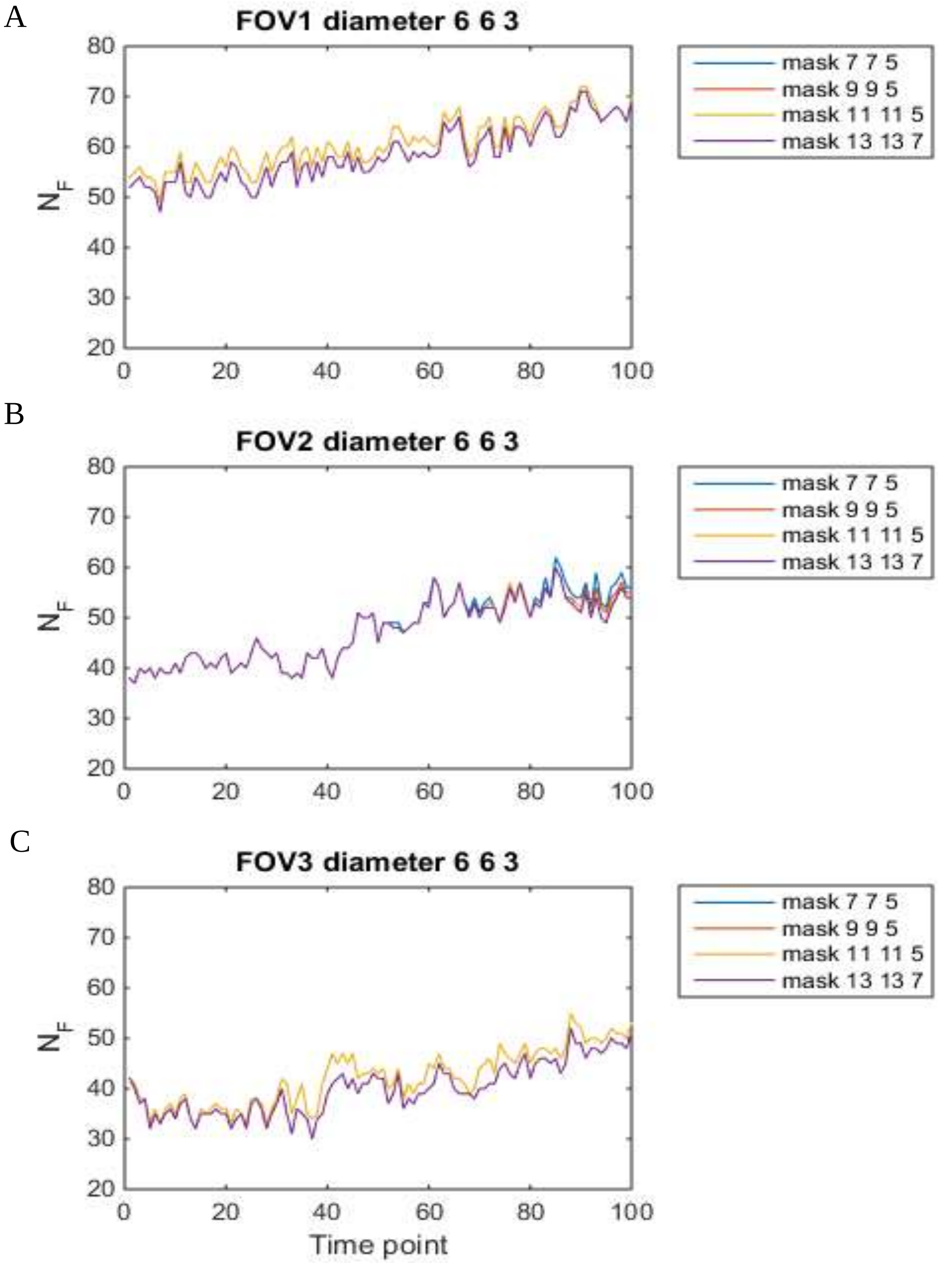


Figure 4-3: Dependence of Number of Features on Mask Size with fixed Diameter

The number of features N_F are plotted for 100 time-points for A. Gel I FOV 1, B. Gel I FOV 2, C. Gel I FOV 3. Superimposed lines are not visible. Diameter is (6,6,3) for all cases. Blue lines represent mask (7 7 5). Red lines represent mask (9 9 5). Yellow lines represent mask (11 11 5). Purple lines represent mask (13 13 7).

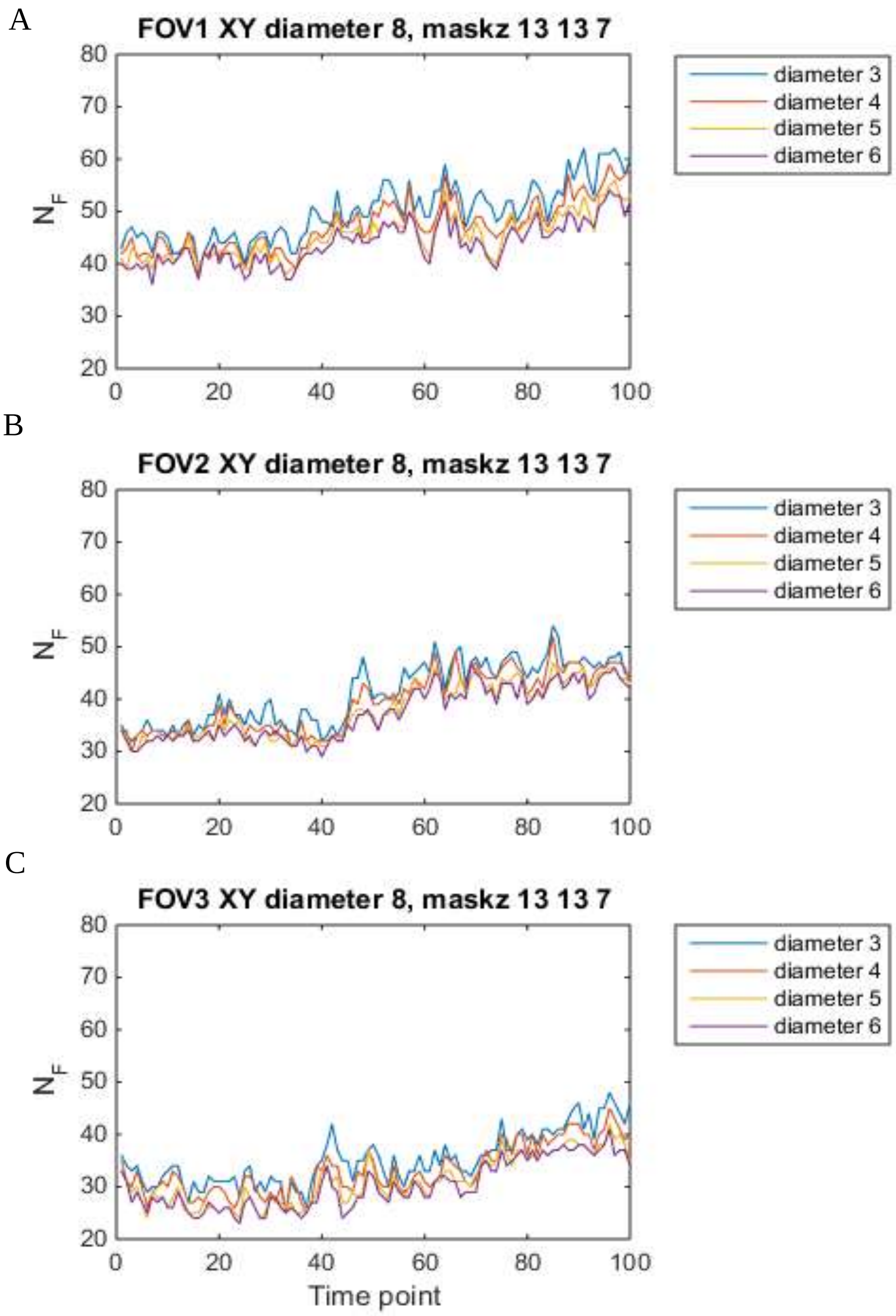
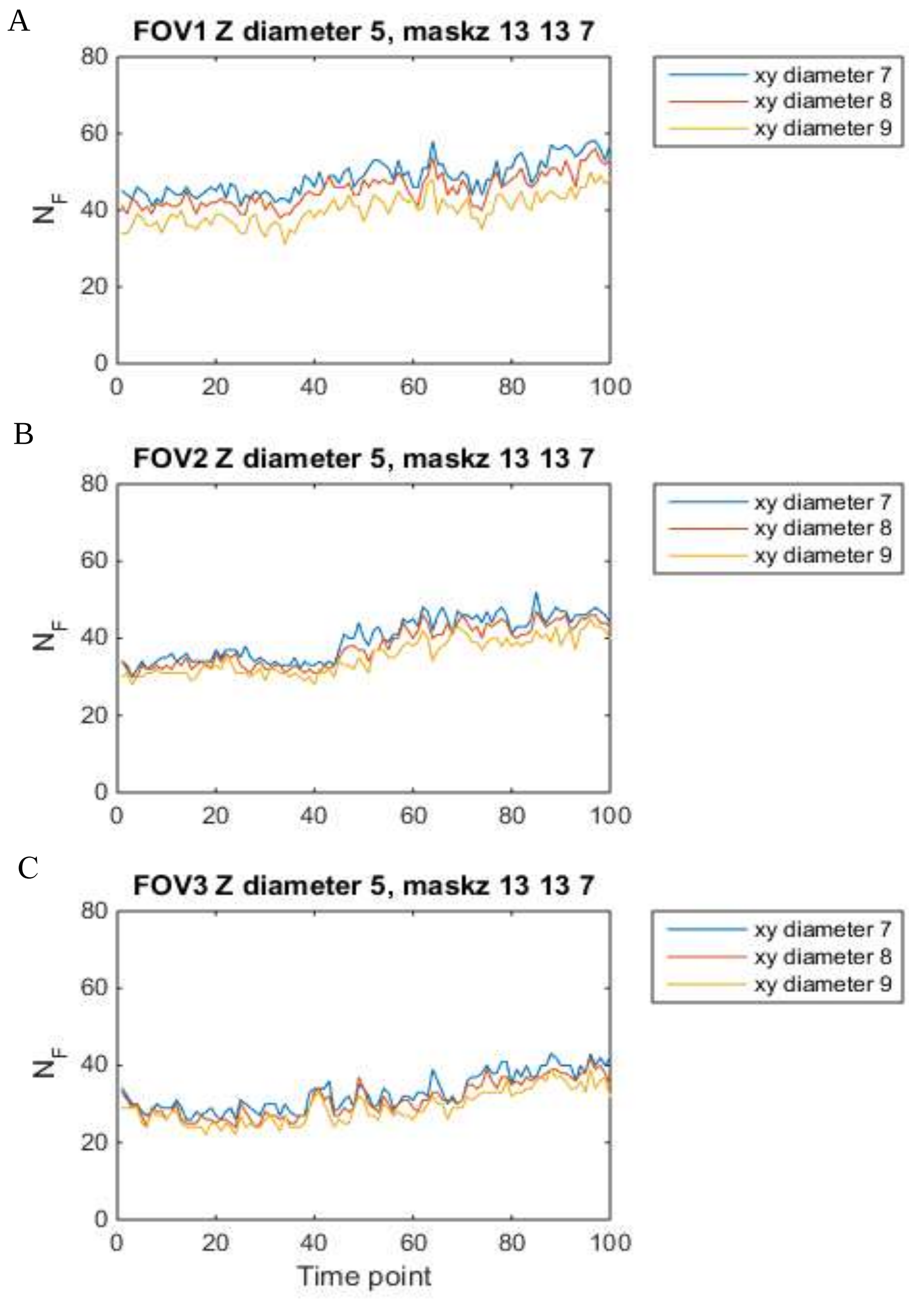


Figure 4-4: Dependence of Number of Features on Z-Diameter Size

The number of features N_F are plotted for 100 time-points for A. FOV1, B. FOV2, and C. FOV3 from Gel I. Only the Z-value of the diameter is varied to 3 (blue), 4, (red), 5 (yellow) and 6 (purple). XY diameter is 8 and mask is (13 13 7)



<u>Figure 4-5: Dependence of Number of Features on XY Diameter Size</u>

The number of features N_F are plotted for 100 time-points for A. FOV1, B. FOV2, and C. FOV3 from Gel I. Only the XY-value of the diameter is varied to 7 (blue), 8 (red) and 9 (yellow). Diameter is 5 for Z and mask is (13 13 7).

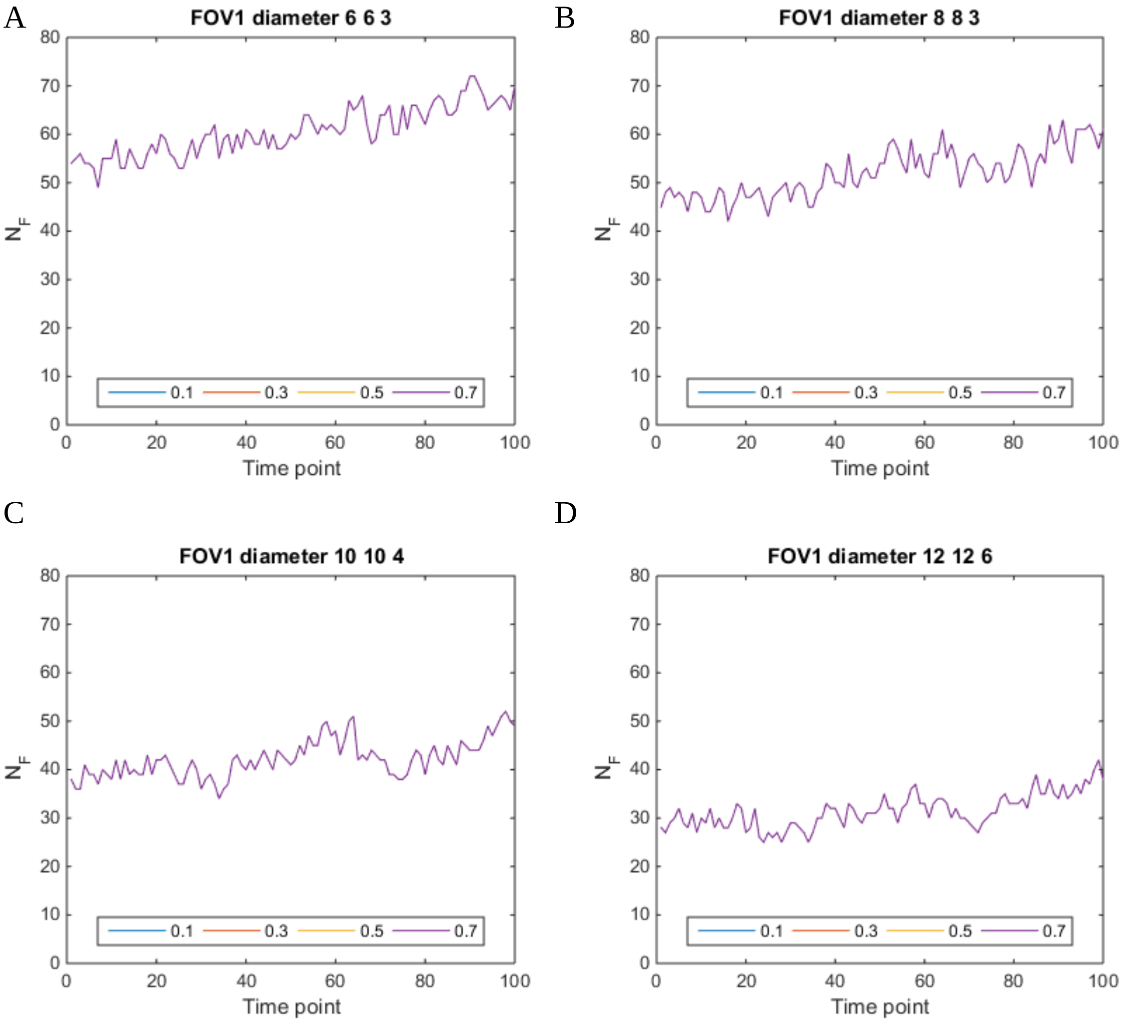


Figure 4-6: Dependence of Number of Features on Threshold

The number of features N_F are plotted for 100 time-points for one fields of view (FOV) from Gel I with A. diameter (6 6 3), B. diameter (8 8 3), C. diameter (10 10 4), and D. diameter (12 12 6). Threshold values are 0.1 (blue), 0.3 (red), 0.5 (yellow), and 0.7 (purple). Superimposed lines are not visible indicating changing threshold does not alter output.

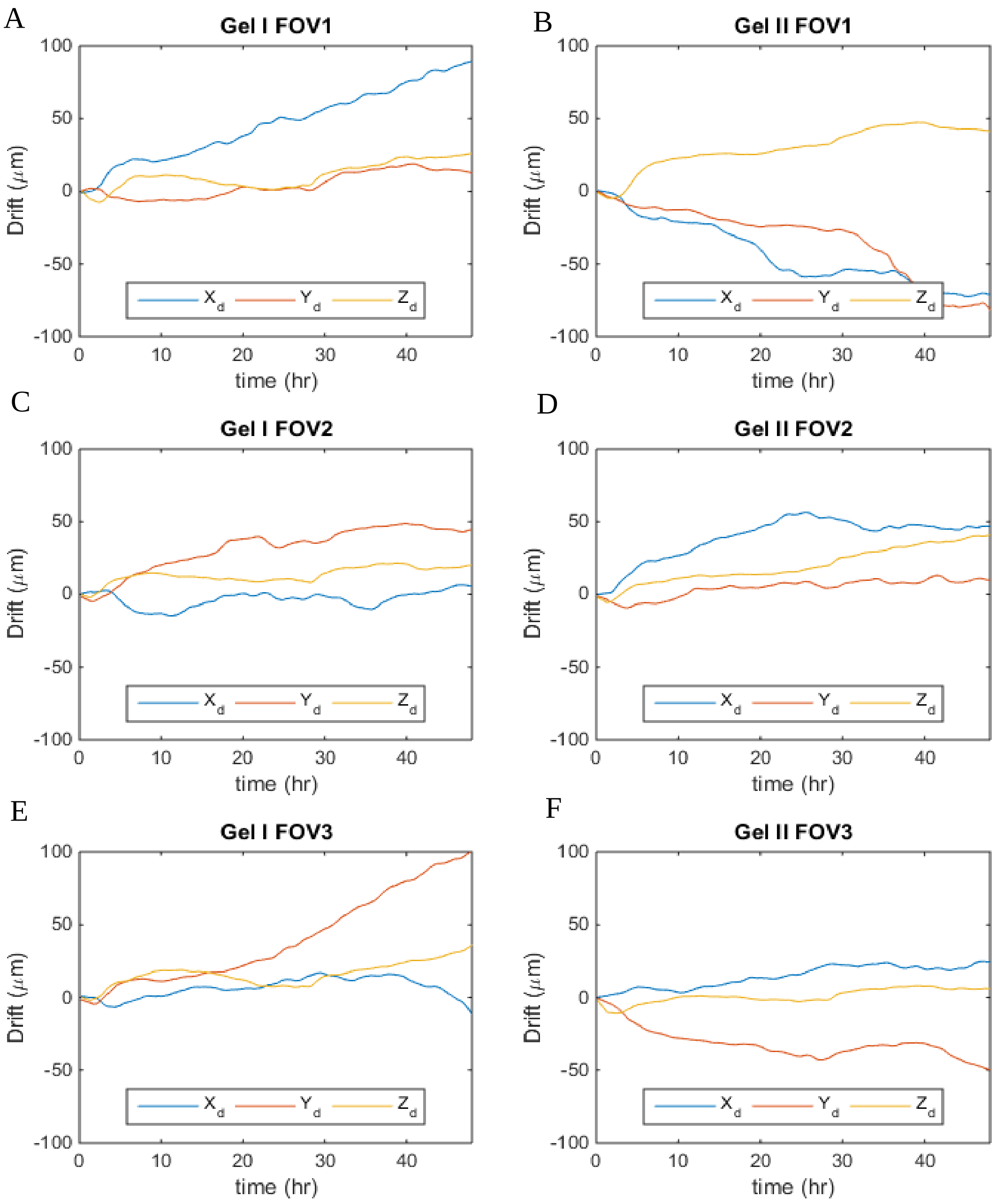


Figure 4-7: Drift eliminated from datasets for 48 hours

The drift observed in the X(blue), Y(red) and Z(yellow) dimensions in 2 gels and 3 fields of view over 48 hours for A. Gel I FOV 1, B. Gel II FOV 1, C. Gel I FOV 2, D. Gel II FOV 2, E. Gel 1 FOV 3, and F. Gel II FOV 3. This drift is eliminated to inform final positions of cellular trajectories.

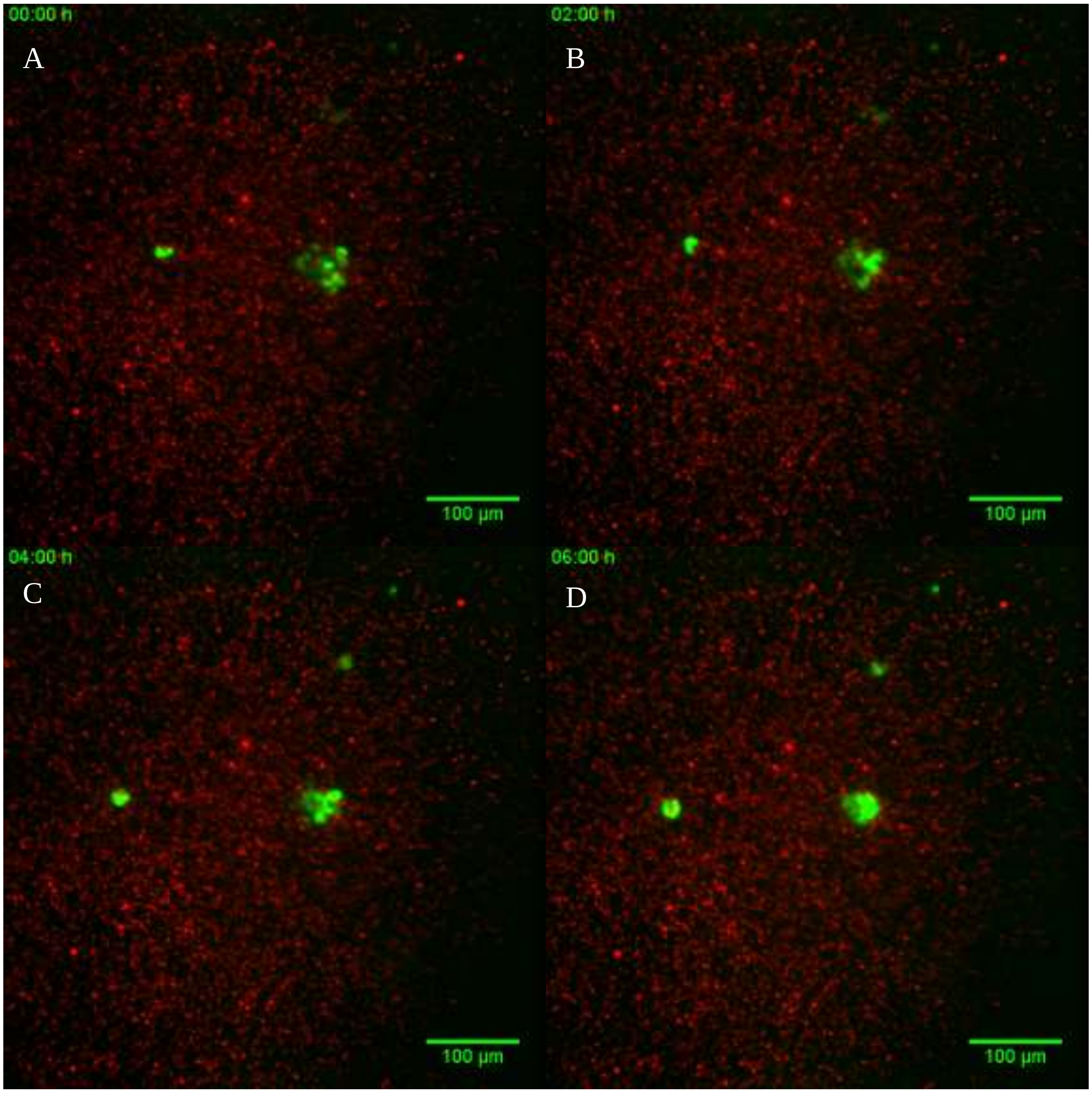


Figure 4-8: Z-projections of cellular clusters with embedded beads

Green colors show nuclei of cohorts, red beads are embedded in a 2 mg/ml collagen matrix. Panels are at A. 0 h, B. 2 h, C. 4 h and D. 6 h into the experiment. The beads are not undergoing local flows in the matrix

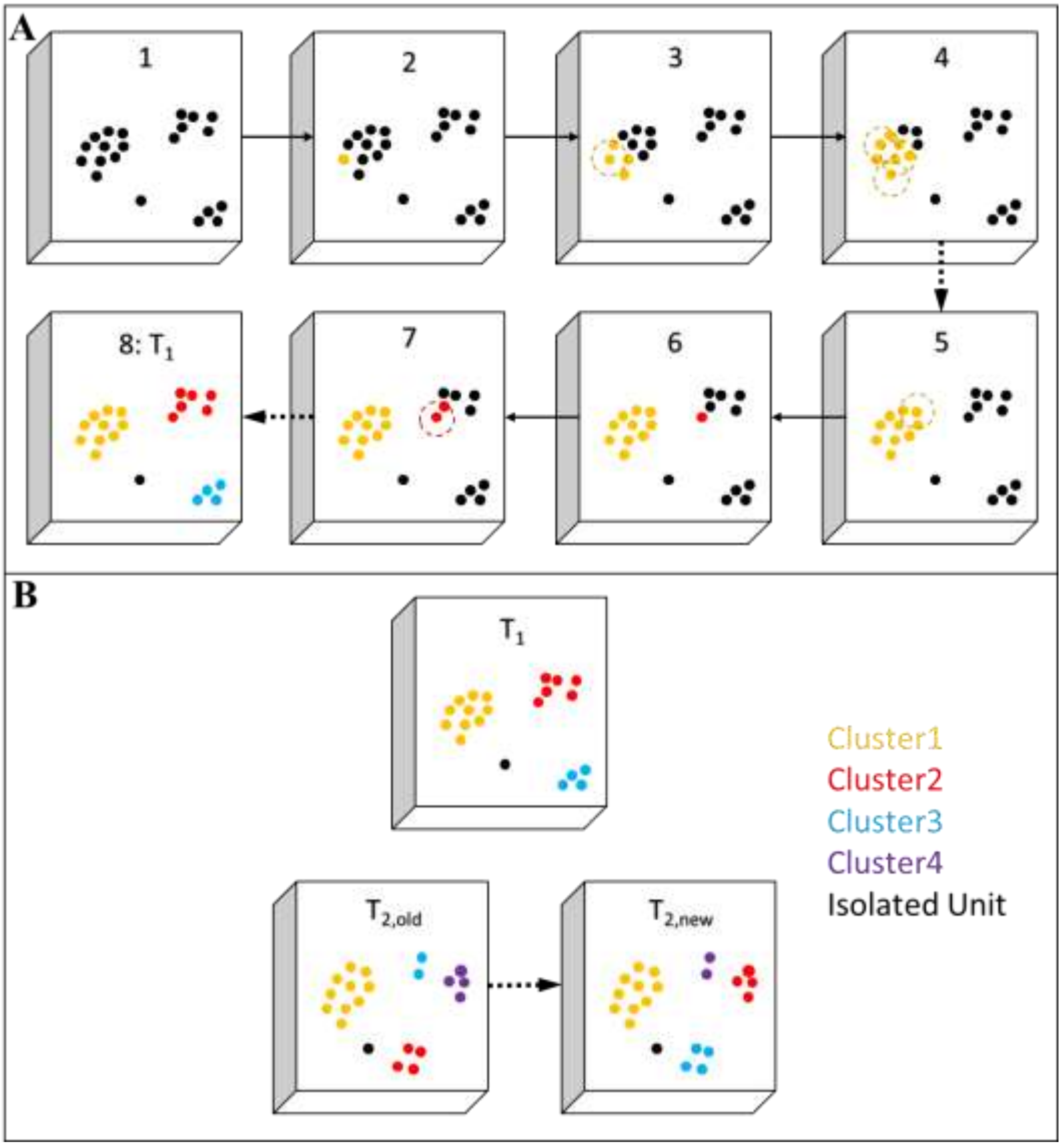


Figure 4-9: Schematic of Clustering Algorithm for Position-Time-Cell ID data

A. Clustering Algorithm for a single time point T1. Box 1- Position data Box 2- Yellow cell is assigned Cluster ID 1 Box 3- All cells within radius ri of this cell are assigned to Cluster 1 Box 4- All cells within radius ri of these cells are assigned to Cluster 1 Box 5- No more cells can be assigned to Cluster 1 Box 6- Red cell is assigned to Cluster 2 Box 7- Cells within radius ri of red cell are assigned to Cluster 2 Box 8- Process repeats until all cells are assigned a Cluster ID and single cells are assigned Cluster ID 0 B. Clusters are correlated across 2 consecutive time-points, T_1 and T_2. For each cluster in T_1 a corresponding cluster in T_2 is identified by finding the cluster in T_2 with maximal overlap of Cell ID's. At the end of the process any unlabeled clusters in T_2 are assigned a new Cluster ID.

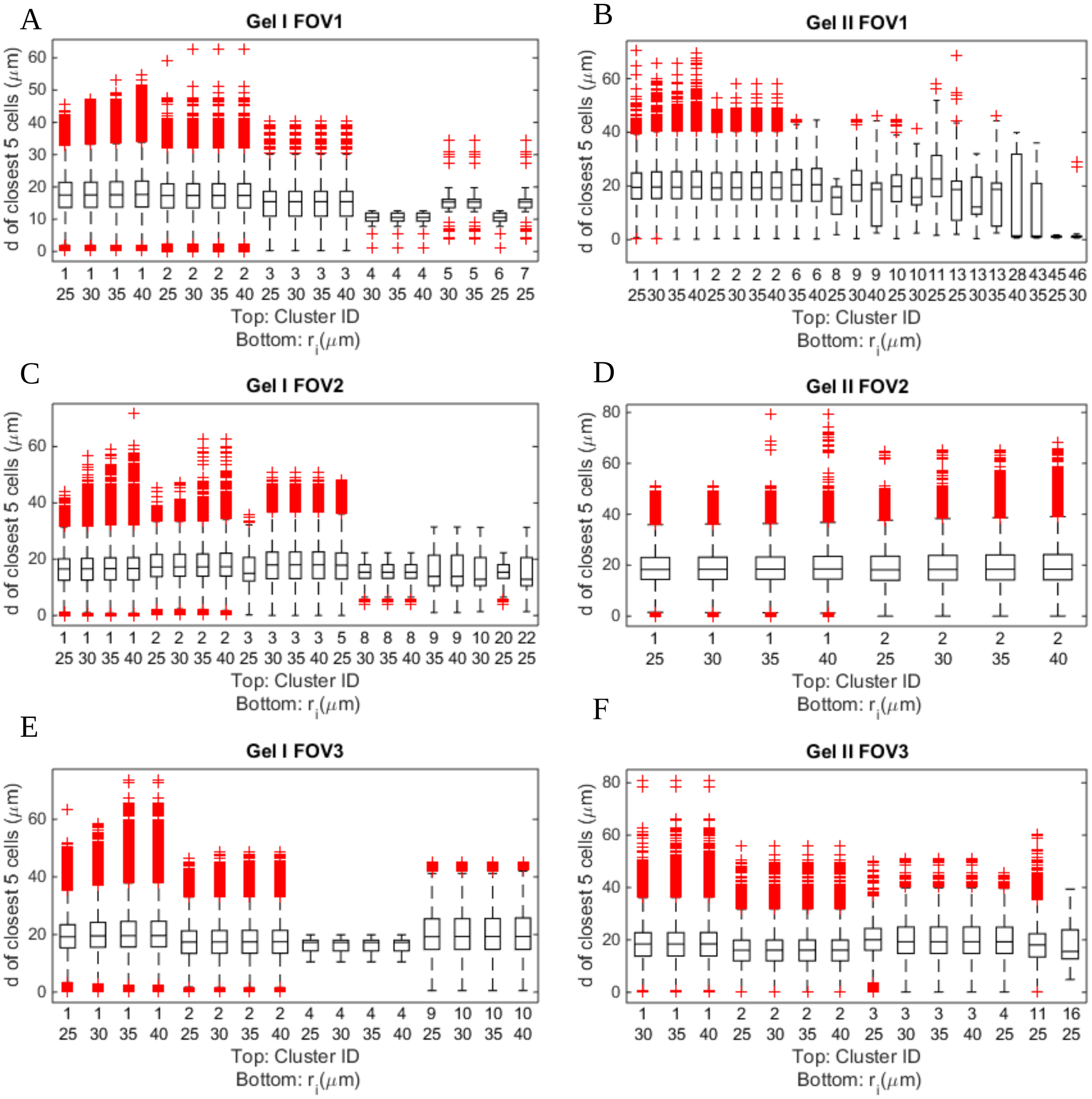


Figure 4-10: Clusters and Neighbors with varying interaction radius

These boxplots how the distribution of distances of five nearest cells for each cluster along all time with various cutoff r_i values for A. Gel I FOV 1, B. Gel II FOV 1, C. Gel I FOV 2, D. Gel II FOV 2, E. Gel 1 FOV 3, and F. Gel II FOV 3. For all cases, median distance of nearest cells is ~20 µm, or 1 3D cell length. Outliers are caused when the nearest neighbors are more than 1 3D cell length away, since a cell may not have five immediate neighbors.

CHAPTER 5. COMPUTATIONAL MODEL FOR COLLECTIVE CELL MOTILITY

5.1. Introduction

Collectively moving entities such as locust swarms, bird flocks, fish schools, bacterial swarms and cells appear to share distinct characteristic traits and motion patterns. Computational modeling techniques employ Self Propelled Particle (SPP) models to recreate and interpret such motion patterns (Vicsek & Zafeiris 2012). These systems self-organize to behave as ‘larger than the sum of their parts’, displaying behavior in which the dominating influence on a unit within a collective is of the properties of the units around it (Meschede & Hallatschek 2013). These collectives are composed of similar individual units that move with nearly the same velocity, are subject to a varying magnitude of noise, and are capable of changing direction and interacting with each other to improve alignment.

This study expands on a previous use of SPP models for cell collectives to propose a 3D SPP model for cell collectives. The most simple (Vicsek et al. 1995) SPP models assign a constant speed to all particles $p \in \{1,N\}$ that are initially placed at random on a 2-dimensional plane; each particle’s direction is then updated according to Equation 5-1, where θ is the direction, v is the velocity, q are all neighbors within a specified radius of attraction, and $\eta\xi$ is a noise term. This allows all particles to evolve over time to have some collective properties and phases that are dependent on noise as well as radius of attraction.

Equation 5-1

$$\theta_p^{t+1} = arg\left[\sum_{p\sim q} \overrightarrow{v_q}\right] + \eta\xi_q^t$$

Simplistic SPP particle models have been expanded further to be more cell like (Grégoire et al. 2003) by scaling a direction dependent term and a force dependent term, but also maintaining constant speed according to Equation 5-2. Here α and β are scaling factors that contribute to the effect of neighbor directionality and forces, respectively; N_p is the number of neighbors that influence a noise term η and a random unit vector u.

Equation 5-2

$$\theta_p^{t+1} = arg\left[\alpha\sum_{p\sim q} \overrightarrow{v_q} + \beta\sum_{p\sim q} \overrightarrow{f_{pq}} + N_p\eta\vec{u}_p^t\right]$$

In this study, first the model presented in Equation 5-2 is adapted for 3D and evaluated to emulate 3D cellular clusters. Further complexity is added to this 3D-SPP model by scaling individual cell speed and accounting for inter-cellular adhesion dynamics.

5.2. SPP Model in 3D

5.2.1. Model Set-up

To initialize the system, 512 cells are seeded randomly onto positions of a 3D cube of dimensions 1280 x 1280 x 1280 with periodic boundaries; each cell is also assigned a starting spherical direction by randomly assigning by two angle vectors (θ,φ) such that θ ∈ [0 2π] and φ ∈ [0 π]. The constant speed value is set at 2 in order to emulate a cellular speed of 2 μm/10 min, such that 1 time step of the computational model is equivalent to 10 minutes. SPP models assume constant speed of all units, this speed is set at 2 for this model.

The 3D direction vector (θ,φ) of any cell at time t+1 updates according to the rules of Equation 5-3. Here θ_p and φ_p are derived by projecting the contributing force C_p onto spherical coordinates– thus θ_p is the argument or arctangent of C_y/C_x and φ_p is the arccosine of C_z/C_r.

Equation 5-3

$$\begin{bmatrix} \theta_p^{t+1} = arg\left(C_{x,p}^{t+1} + iC_{y,p}^{t+1}\right) \\ \varphi_p^{t+1} = arccos\left(\frac{C_{z,p}^{t+1}}{\sqrt[2]{\left(C_{x,p}^{t+1}\right)^2 + \left(C_{y,p}^{t+1}\right)^2 + \left(C_{z,p}^{t+1}\right)^2}}\right) \end{bmatrix}$$

The value of contributing forces C_p is also updated for three dimensions where instead of one degree of freedom in angle there are now two. Two forces affect the change in direction: the first being the direction or velocity $\vec{v}_q$ of each individual neighbor q, which is scaled by α; and the second being an attraction-repulsion force term $\vec{f}_{pq}$ between cells p and neighbors q, which is scaled by β. A noise function is also added to the direction by randomizing two angles: Θ and Φ, and multiplying these with an input noise scaling factor η as well as number of neighbors N_p. The sum total of these forces to the contributing force vector C_p in the x, y and z direction is described in Equation 5-4; η is set to a value of 1 for 3D SPP evaluation. N_p is the number of neighbors within an equilibrium distance $r_e = 20$; this value is $\cong$ 1 3D cell diameter in µm in order to emulate cellular systems. The N_p value is set at a maximum of 12, which would facilitate maximal packing assuming perfect spheres.

Equation 5-4

$$\begin{bmatrix} C_{x,p}^{t+1} = \alpha \sum_{p\sim q} \vec{v}_{q,x} + \beta \sum_{p\sim q} \vec{f}_{pq,x} + N_p \eta \cos\Theta \sin\Phi \\ C_{y,p}^{t+1} = \alpha \sum_{p\sim q} \vec{v}_{q,y} + \beta \sum_{p\sim q} \vec{f}_{pq,y} + N_p \eta \sin\Theta \sin\Phi \\ C_{z,p}^{t+1} = \alpha \sum_{p\sim q} \vec{v}_{q,z} + \beta \sum_{p\sim q} \vec{f}_{pq,z} + N_p \eta \cos\Phi \end{bmatrix}$$

The attraction repulsion force term $\vec{f}_{pq}$ between two cells p and q is described in Equation 5-5 and adapted as is from the parallel 2D model (Grégoire et al. 2003). This is essentially a modified Lennard-Jones potential: here $\vec{e}_{pq}$ is the unit vector along the direction between the two cells; r_{pq} is the distance between two cells; r_c is the core repulsion radius, cell pairs with r_{pq} below this have high repulsion; r_a is the attraction radius, cell pairs with r_{pq} above this sense a constant small attraction; re is an equilibrium radius as defined above, cell pairs with an r_{pq} between r_c and r_a sense an attraction scaled by the value of r_{pq}; and r_0 is the limit of interaction beyond which cells sense no force. To emulate cellular systems, r_c is 8, a typical nucleus diameter, at which there should be high repulsion; r_a is 32, ~ 1.5 cell diameters along which there is attraction, and r_0 the limit of interaction is 36.

Equation 5-5

$$\begin{bmatrix} \vec{f}_{pq} = \vec{e}_{pq} \begin{cases} -\infty & if\ r_{pq} < r_c \\ \frac{1}{4}\frac{r_{pq} - r_e}{r_a - r_e} & if\ r_c < r_{pq} < r_a \\ 1 & if\ r_a < r_{pq} < r_0 \\ 0 & if\ r_{pq} > r_0 \end{cases} \end{bmatrix}$$

Unlike the reference model in 2D (Grégoire et al. 2003) this work does not incorporate Voronoi tessellation to achieve volume exclusion in 3D. This volume exclusion is too computational intensive and not considered fundamentally important, since the repulsion term in Equation 5-5 is sufficient to account for volume exclusion.

5.2.2. Testing SPP parameters

The SPP model is allowed to evolve for 20000 time-steps to ensure equilibrium for values of $\alpha,\beta \in \{0,0.01,0.03,0.05,0.07,0.1,0.3,0.5,0.7,1,3,5,7,10\}$. A schematic of the final time-step for these conditions is in Figure 5-1. It is apparent that low values of α and β do not result in formation of a lot of clusters whereas high values of both result in worm like structures.

For the last 288 time-steps, cluster values are assigned to each output cell using the same clustering algorithm described in Section 4.3.4. Briefly, a hierarchical clustering algorithm assigns clusters at a single time point by finding all cells within an interaction radius r_i of each other and assigning them to the same cluster. The interaction radius is set to 30 to emulate cell clusters. Clusters are then correlated across consecutive time points T_i and T_{i+1} according to the following rules:

1. For all clusters in T_i, the cluster in the T_{i+1} which has maximal overlap of same cells is found and assigned the corresponding identity.
2. If two or more clusters from T_i have maximal overlap with the same cluster in T_{i+1}, the identity of largest of these clusters in T_i is assigned to the merged cluster in T_{i+1}.

3. After all the clusters in T_i have been accounted for or merged, unassigned clusters in T_{i+1} are assigned a new identity.
4. If T_i has no clusters and T_{i+1} has clusters, all clusters in T_{i+1} are assigned new identities.

The average number of clusters over the final 288 time-steps for 512 cells for different α - β scaling parameters is mapped in the top panel of Figure 5-2. The median size of clusters formed over this time-range is depicted in the bottom panel. The number of clusters formed has a biphasic relationship with cluster size– as more clusters form cluster size decreases by increasing either α or β, but upon increasing them both clusters are so large that all 512 cells form 5-10 clusters. Based on number of clusters formed and average size, desired α and β would be larger than 1 but not as large as 10 concurrently. The average distance from the center of the cluster for the clusters form is mapped in Figure 5-3. It is apparent that when both values of α and β are high the clusters are very large (average distance from center > 60), thus target values to emulate cellular cohorts would be combinations of high α (>5) and low β (<1), or high β (>0.03) and low-medium α (<5).

5.2.2.1. Order Parameter

Order parameters provide an effective tool to gauge the cohesiveness of SPP models, and thus a translation dependent order parameter ϕ is introduced as in Equation 5-6. Here p is an individual cells in a system of N total cells, and $\vec{v}_p$ is the velocity at time t.

Equation 5-6

$$\phi(t) = \frac{\left|\sum_{p=1}^{N} \vec{v}_p\,(t)\right|}{\sum_{p=1}^{N}\left|\vec{v}_p(t)\right|}$$

The order parameter for the whole system (N = 512), and the median order parameters of all clusters (a cluster is defined as a group of 2 or more cells) is mapped in Figure 5-4. In order to compare with cellular cohorts, the order parameter ϕ of individual cohorts presented in Figure 4-1 is evaluated to be between 0.6-1 over 1 hour intervals, and the cellular speed between 3-7 μm/h of each individual cell. To further test which values of α and β values could emulate cellular clusters, all datasets are tested to find those which fit three criteria– a median cluster size between 4 and 31 units; a median order parameter ϕ over 6 time-points between 0.6-1; and a cellular distance covered over 6 time points between 3-7. Results of such a criteria analysis are depicted in Table 5-1, suggesting that values of $\alpha \in [0.3\ 0.7]$ and $\beta \in [1\ 10]$ are ranges in which simulated clusters could emulate experimental cohorts

5.2.2.2. Stability of formed clusters

Once clusters form as the model evolves, they may either stay in their established form or have some internal fluctuations. To map out whether these internal fluctuations influence positions of individual cells within clusters, the number of neighbors in the center and extremes of the cluster are plotted in Figure 5-5. The similar trends in the two left and two right plots indicate that once a cluster forms, individual units tend to retain their positions relative to each other in both the center and the extreme edges of the cluster. Thus there is no size dependent or position dependent instability in these clusters that leads to more diffusion within them in certain conditions.

To evaluate the change in shape of clusters over time, an inertia tensor is evaluated for each cluster assuming that the cluster is a collection of spherical rigid bodies. This inertia tensor is described by Equation 5-7 where T is the inertia tensor; p is a unit and the total number of units is N; m is a mass, considered 1 for all points in this system; I is the identity tensor; r_p is the position vector of particle p (x_p,y_p,z_p); and ⊗ is the tensor product. Once the inertia tensor T is known, the change of shape can be evaluated by calculating the angle change of any eigenvector across consecutive time-points, the average angle change of the smallest eigenvector of the inertia tensor in consecutive time-steps is mapped in Figure 5-6. Conditions that form smaller clusters or even no clusters have large angle change between consecutive time-points for the principal eigenvectors of the inertia tensor, whereas larger cluster forming conditions have more fidelity of shape.

Equation 5-7

$$T = \sum_{p=1}^{N} m\left(\left(r_p \cdot r_p\right)I - \ r_p \otimes r_p\right)$$

5.3. 3D-SPP model with Variable Speed

5.3.1. Model Set-up

To reduce computation time, the scale of the world is decreased to 100 units in a 743 × 743 × 743 sized cube with periodic boundaries. In order to better emulate cellular cohorts, further complexity is added to the existing 3D SPP model by allowing individual units to have variable speed. Cell speed is scaled linearly and decreases with the presence of neighbors that form multiple Adherens Junction (AJ) complexes formed by membrane bound E-cadherin molecules; numbers of AJs formed for each cell are evaluated based on neighboring cells (Vargas et al. 2015). The time-scale for each step of the model is 10 minutes, which is significantly larger than single E-cadherin bond lifetime (Evans & Calderwood 2007); since homotypic E-cadherin bonds are persistent and do not vary in strength based on intercellular forces, speed is scaled linearly with the number of bonds to be at 86% of maximal value at an equilibrium AJ concentration (more neighbors) and 100% value at 2 μm/10 min or 2 box-units/simulation-time unit.

5.3.1.1. Verifying Role of E-cadherin

MDCK GFP-Ecad cells (Adams 1998), modified to express fluorescent marker GFP-tagged E-cadherin molecules are cultured according to the protocols described in Section 4.2.1-4.2.4. Upon visualization, the cell lines show E-cadherin molecules at cell surfaces as displayed in Figure 5-7.

5.3.2. Testing SPP parameters

The number of clusters formed and the median cluster size for various α-β conditions is mapped in Figure 5-8. The cluster size has at values of (α,β) of (10,1) and (10,0.1) and is anti-correlated with the number of clusters as expected. The median order parameter ϕ and average AJ concentration for the same data is mapped in Figure 5-9.

The order parameter is not linearly correlated with the number of clusters any more, since larger clusters at (α,β) of (10,1) and (10,0.1) have low order. The mean adhesiveness also does not correlate linearly with the other parameters and peaks at (α,β) values of (1,3).

Even though the datasets at (α , β) of (10,1) and (10,0.1) have low order, their shape remains conserved as mapped in Figure 5-10; in fact, these datasets have the least change in angle compared to others, even though they have high distance from the center of the cluster. On the other hand, clusters spread very far apart in the low (α , β) values of (0.01,1) also have high angle change in consecutive time-points, indicating an unstable system. However, visualizing the neighbors in the center and edge of clusters as well as their conservation does indicate a generally stable cluster system as in Figure 5-11 for these ranges. This means that for (α , β) values of (0.01,1), shape change and instability across consecutive time-points occurs without altering neighbors. Based on the large distance to center in this region, Figure 5-10, it is likely that all clusters fluctuate within consecutive time-points.

It is already apparent that adding the complexity of variable speed does not allow straightforward characterization of the system. Thus in order to further test which values of α and β could emulate cellular clusters, the analysis of section 0 is repeated. All datasets are tested to find those which fit three criteria– a median cluster size between 4 and 31 units; a median order parameter ϕ over 6 time-points between 0.6-1; and a cellular distance covered over 6 time points between 3-7. Results of such a criteria analysis are depicted in Table 5-2. The noise function η is also varied between 0.1 and 10 for a few interesting conditions. Results of criteria analysis of this are in Table 5-3. For a few of the datasets that could be emulating cellular cohorts, the last time-step is visualized in Figure 5-12.

5.4. Discussion

The creation of a 3D SPP model allows multiple phases and systems to form, from worm like creatures to cohorts to single cells and cell pairs. Visualizing maps of various quantities shows clear trends that increasing not just the ratio of α to β influences the output states, but the values themselves. Adding just a little complexity to such a model system in the form of scaling speed based on adhesion with neighboring members creates a dynamic model with unpredictable outcomes. There are also several possible conditions that may be emulating cellular systems and may contain crucial information about intercellular and intracohort heterogeneity. In order to obtain valuable information from this model, analytical tools that are able to extract relevant data from a large input pool are needed. However with the simple checkpoints presented in Table 5-1, Table 5-2, and Table 5-3 there are already are conditions at hand that can serve as a computational model for cellular systems.

5.5. Tables

↓α/β→	0	0.01	0.03	0.05	0.07	0.1	0.3	0.5	0.7	1	3	5	7	10
0	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓	✓	✓
0.01	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓	✓	✓
0.03	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓ ✓	✓ ✓
0.05	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓	✓	✓
0.07	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓	✓ ✓	✓ ✓	✓ ✓
0.1	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓	✓ ✓
0.3	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓ ✓	✓ ✓	✓ ✓	✓ ✓ ✓	✓ ✓
0.5	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
0.7	✓ ✓	✓	✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
1	✓	✓	✓	✓	✓	✓	✓	✓	✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
3	✓	✓	✓	✓	✓	✓	✓	✓ ✓	✓	✓	✓	✓	✓ ✓	✓ ✓
5	✓	✓	✓	✓	✓	✓ ✓	✓ ✓	✓	✓	✓	✓ ✓	✓	✓ ✓	✓
7	✓	✓	✓	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓	✓	✓	✓	✓	✓
10	✓	✓	✓	✓ ✓	✓ ✓	✓	✓	✓	✓ ✓	✓	✓	✓ ✓	✓	✓

Table 5-1: SPP-3D Values Compared with Cellular Cohorts

Black checkmarks indicate that the median cluster size not counting single cells is between 4 to 31 cells. Red checkmarks indicate that the cells in each dataset move 3-7 distance units every 6 time-steps. Blue checkmarks indicate that the order parameter of individual clusters is between 0.6-1. Values of α ∈ [0.1 1] and β ∈ [1 10] are ranges in which simulated clusters could emulate experimental cohorts.

↓α/β→	1	2	3	6	10	20	30	60	100
0	✓ ✓	✓	✓		✓	✓	✓	✓	
0.01	✓ ✓	✓	✓	✓	✓	✓	✓		
0.03	✓ ✓	✓	✓	✓	✓	✓	✓	✓	
0.1	✓ ✓	✓	✓	✓	✓	✓	✓	✓	
0.3	✓ ✓	✓	✓	✓	✓	✓	✓	✓	✓
1	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
3	✓ ✓	✓ ✓	✓ ✓	✓	✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓	✓ ✓ ✓
10	✓	✓	✓	✓ ✓	✓ ✓	✓ ✓	✓	✓ ✓ ✓	✓ ✓

Table 5-2: Variable Speed SPP Values Compared with Cellular Cohorts

Black checkmarks indicate that the median cluster size not counting single cells is between 4 to 31 cells. Red checkmarks indicate that the cells in each dataset move 3-7 distance units every 6 time-steps. Blue checkmarks indicate that the order parameter of individual clusters is between 0.6-1. Values of (α,β) equaling (3,20), (3,30), (3,100), and (10,60) could emulate experimental cohorts.

α	↓β/ η→	0.1	0.2	0.5	1	2	5	10
1	1	✓ ✓	✓ ✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
3	3	✓ ✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
3	20	✓	✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓	✓ ✓	✓
10	60	✓	✓ ✓ ✓	✓ ✓	✓ ✓ ✓		✓ ✓ ✓	✓ ✓
1	3	✓	✓	✓	✓ ✓	✓	✓	✓ ✓
3	1	✓	✓ ✓	✓	✓ ✓	✓	✓ ✓	✓ ✓

Table 5-3: Role of Noise-factor η for Various α-β Combinations

Black checkmarks indicate that the median cluster size not counting single cells is between 4 to 31 cells. Red checkmarks indicate that the cells in each dataset move 3-7 distance units every 6 time-steps. Blue checkmarks indicate that the order parameter of individual clusters is between 0.6-1. Values of (α,β,η) equaling (1,1,0.2), (3,3,0.1), (3,20,0.5), (3,20,1), (10,60,0.2), (10,60,1), and (10,60,5) could emulate experimental cohorts.

5.6. Figures

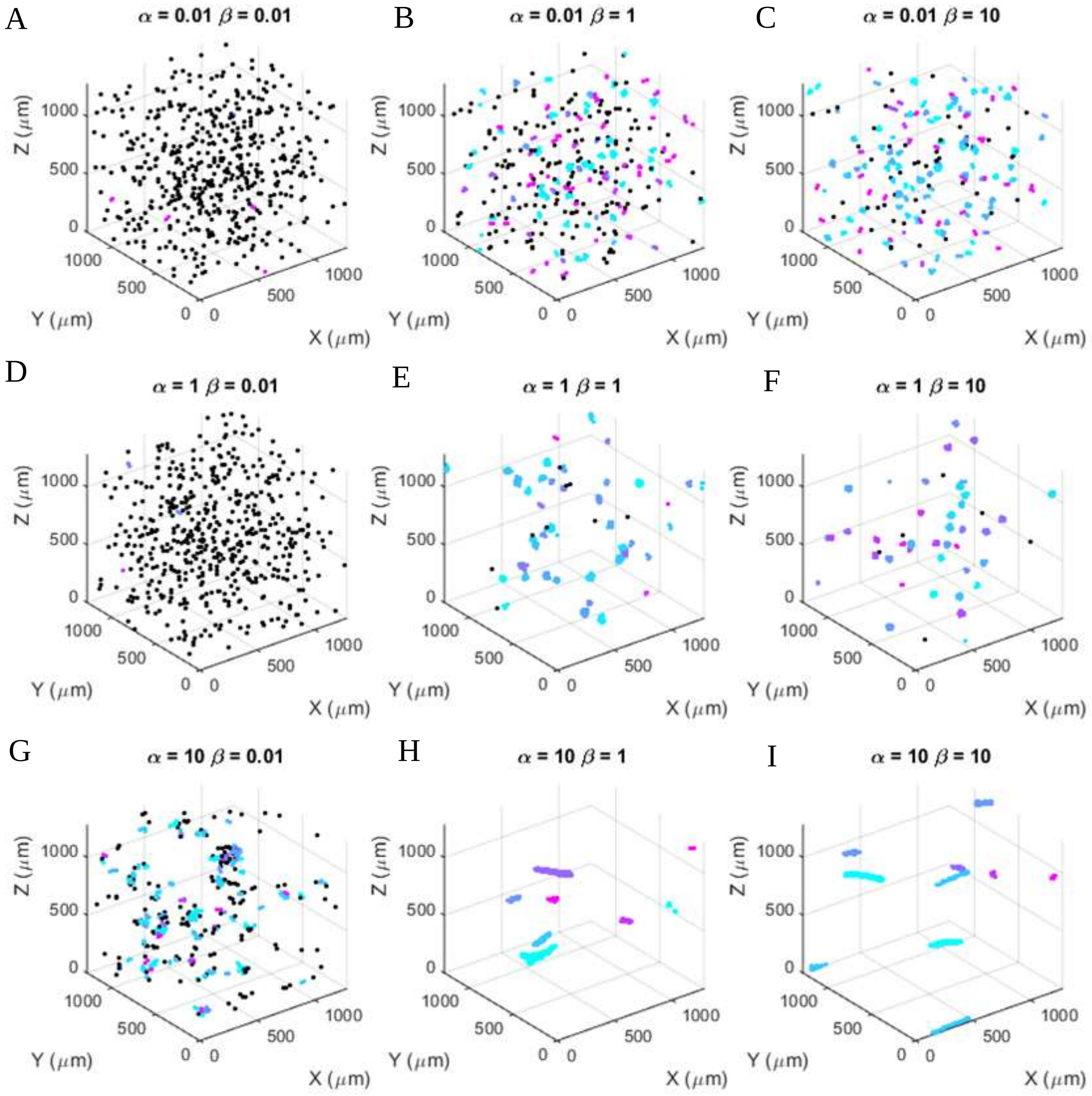


Figure 5-1: Visualization of SPP model for 9 α-β combinations

The circles mark cells or units in an XYZ frame for (α,β) values of A. (0.01, 0.01), B. (0.01, 1), C. (0.01, 10), D. (1, 0.01), E. (1, 1), F. (1, 10), G. (10, 0.01), H. (10, 1), and I. (10, 10). Units belonging to a single cluster are assigned the same color, units without any cluster are black in color. Increasing magnitude of α and β forms clusters in the final time point, and can even form worm like organized clusters in cases of high α.

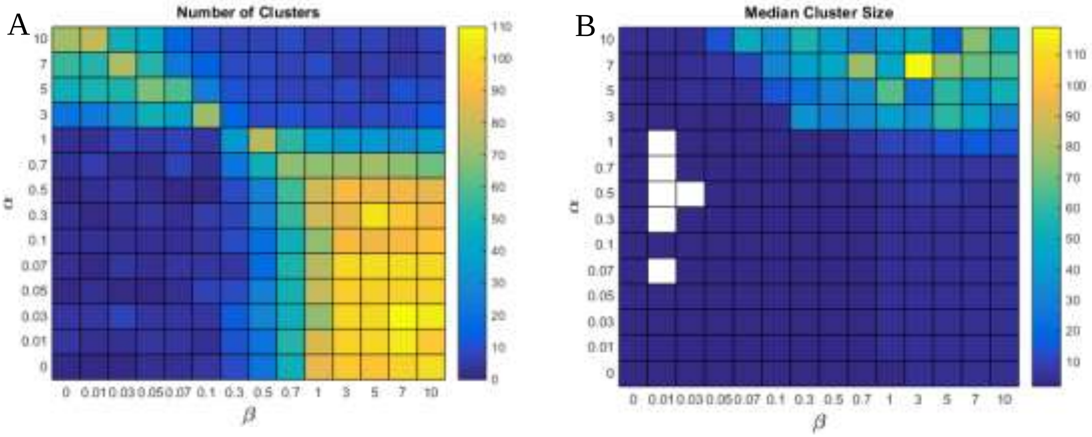


Figure 5-2: Number of Clusters and Median Cluster Size for ranges of α and β

A. The number of clusters formed over different alpha beta values. B. the median cluster size of clusters formed over final 288 time-steps. The number of clusters formed has a biphasic relationship with cluster size–as more clusters form cluster size decreases by increasing either α or β, but upon increasing them both clusters are so large that all 512 cells form 5-10 clusters. Desired range of α and β to emulate cells would be in the light blue region.

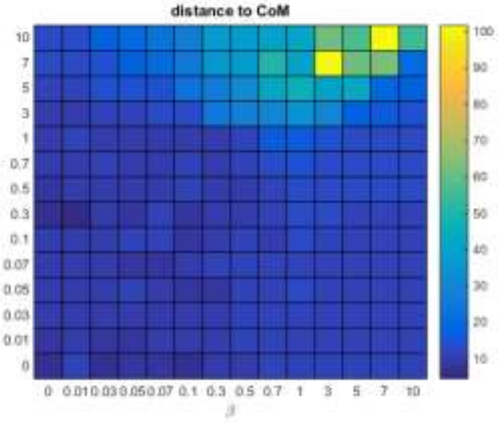


Figure 5-3: Distance to Center of Mass of Clusters for α and β values

The average distance of cells from the Center of Mass or units of the cluster. White spots indicate the clusters were too small for computing distance.

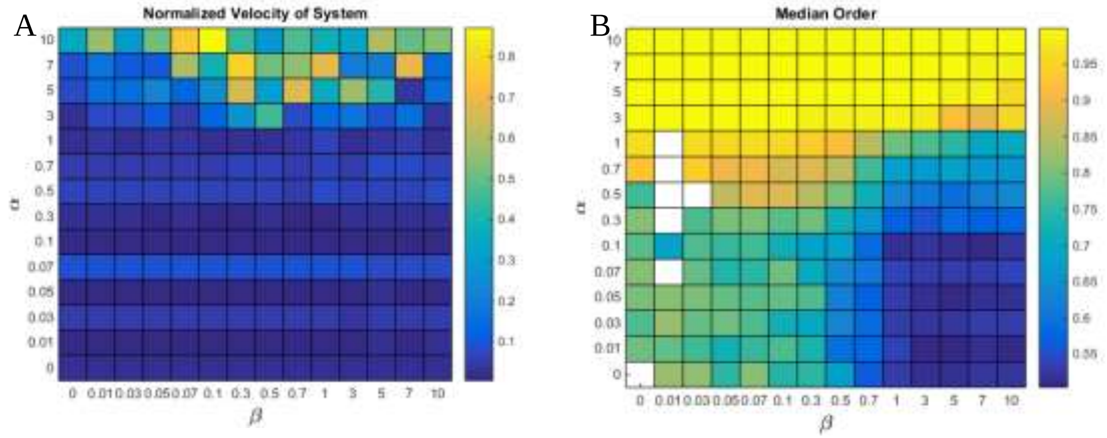


Figure 5-4: Order Parameter for SPP model in 3D

A. Order parameter averaged over all cells. B. Median order parameter of clusters over all time. The white boxes occur when clusters either do not exist or do not exist for consecutive frames, thus resulting in no values for the order, which is velocity dependent. High α values result in highly ordered clusters. High β values result in highly ordered clusters when corresponding with high α.

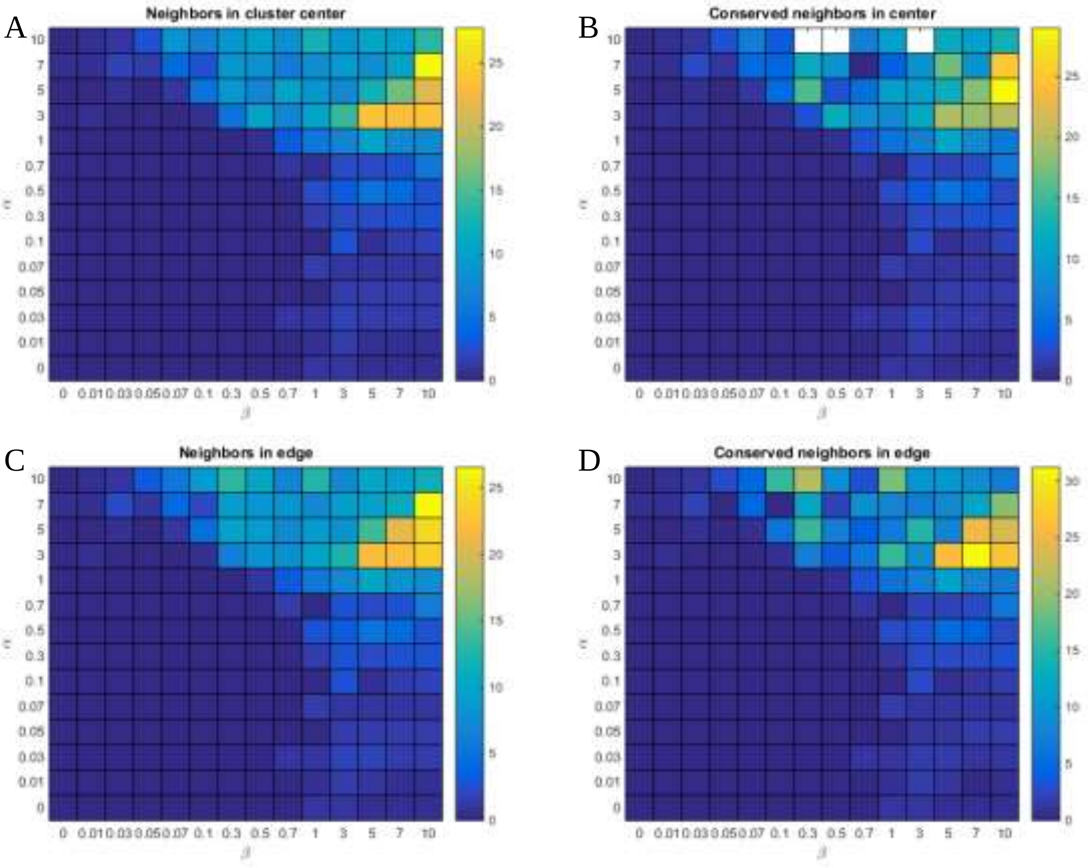


Figure 5-5: Neighbors and Conservation for SPP 3D model

The metric for A. neighbors in the cluster center, B. conserved neighbors in the center, C. neighbors in the edge and D. conserved neighbors in the edge on average for clusters in each conditions allows us to map diffusiveness within individual clusters. The similar trends in the two left and right plots indicate that once a cluster forms, individual units tend to retain their positions relative to each other.

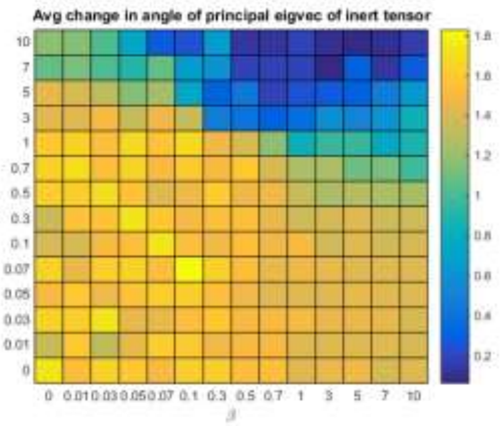


Figure 5-6: Shape Change over α β values

Map of average change in angle of smallest vector of inertia tensor between consecutive timepoints. The change in the angle of orientation of inertia tensor between consecutive time-points indicates the fidelity of shape formed for clusters in each condition. Conditions that form smaller clusters correspond with higher angle change up ~π/2 in magnitude, whereas larger cluster forming conditions correspond with lower shape change.

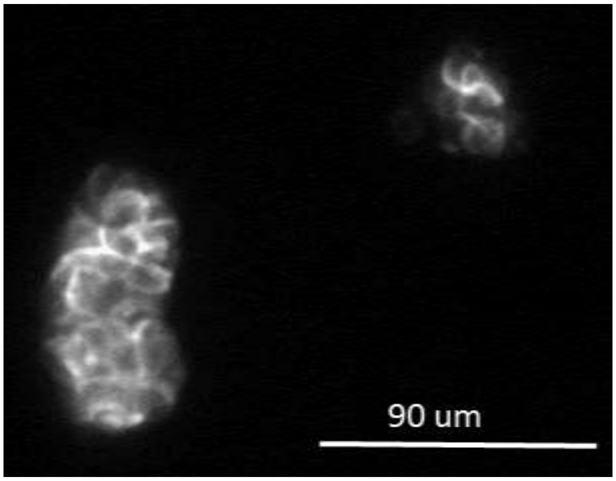


Figure 5-7: Z-projection image of Ecad-GFP MDCK cell cohorts in a collagen matrix

Z-projection sum of all slices imaging two cohorts of GFP-Ecad MDCK cells embedded in 2 mg/ml collagen.

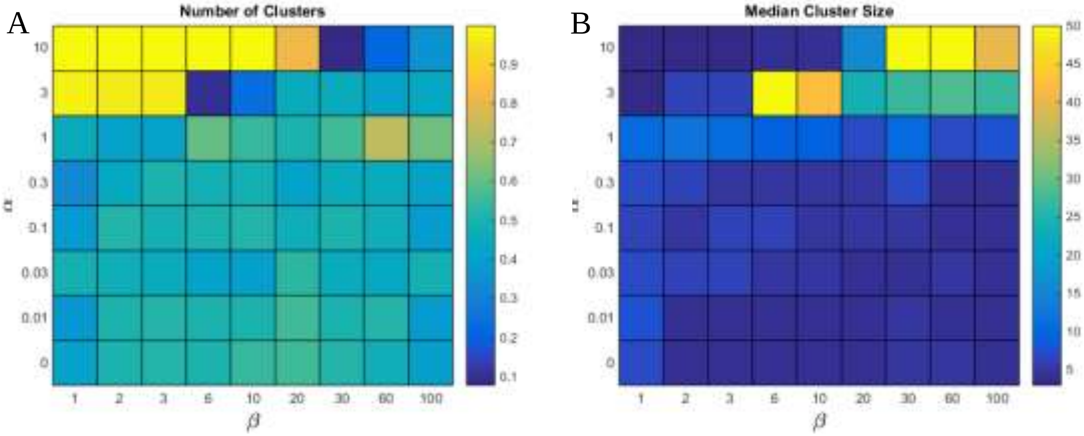


Figure 5-8- Number of Clusters and Median Cluster Size for Variable Speed 3D SPP

A. Number of clusters in Variable Speed 3D-SPP for combinations of α-β values B. Median Cluster Size of existing clusters in the same dataset. While a few data sets form 2-3 large clusters of 35-50 each (total N = 100), a lot of conditions form smaller clusters from 2-20 units big.

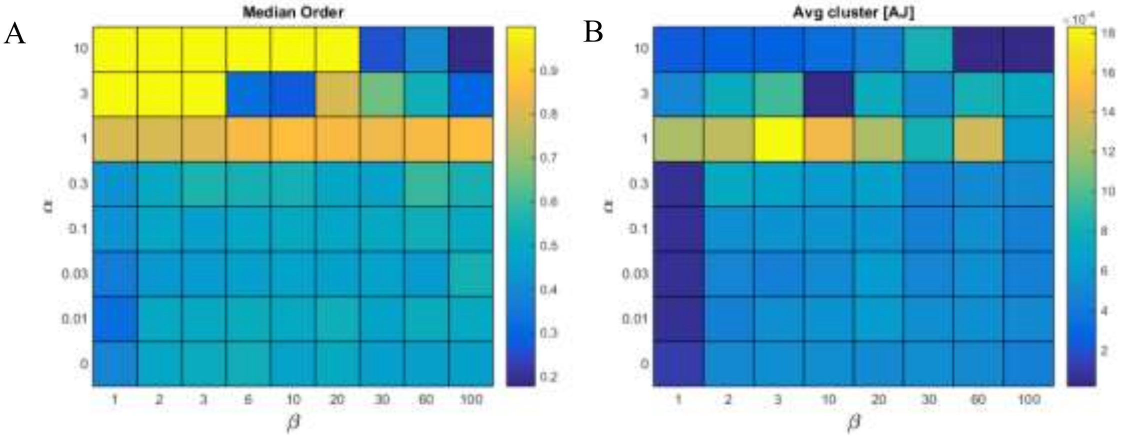


Figure 5-9: Order and Adhesion for Varying Speed 3D-SPP model

A. Median Order in Variable Speed 3D-SPP for combinations of α-β values B. Average AJ concentration of existing clusters in the same dataset. The least ordered clusters also have low AJs, whereas very ordered clusters do not necessarily have large AJs. Clusters with medium order tend to correlate with maximum AJ concentration.

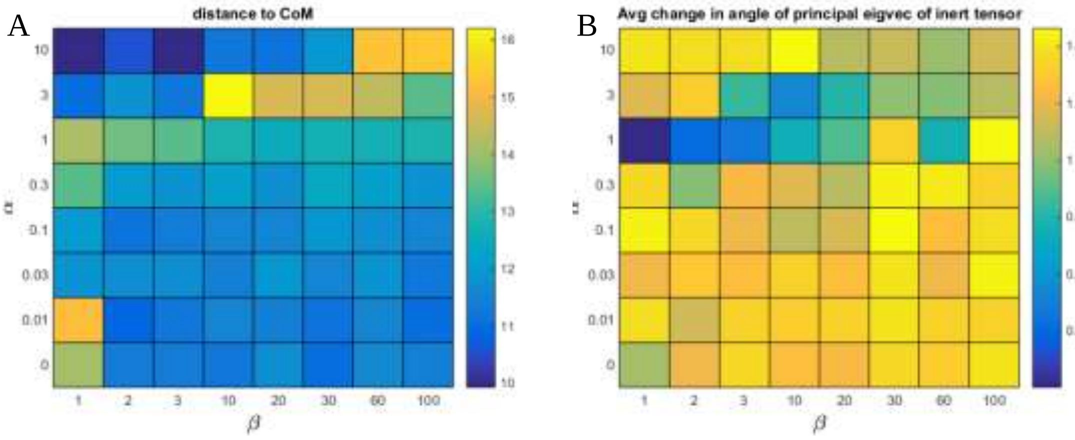


Figure 5-10: Distance to Center and Average Shape Change for Varying Speed 3D-SPP

A. Average distance to center of cluster for α-β values. B. Average angle change of smallest eigenvector of inertia tensor across consecutive time-points. The largest clusters retain fidelity in shape. Some medium sized clusters have high shape change or fluctuation for (α,β) values of (0.01,1), indicating a fluctuating system.

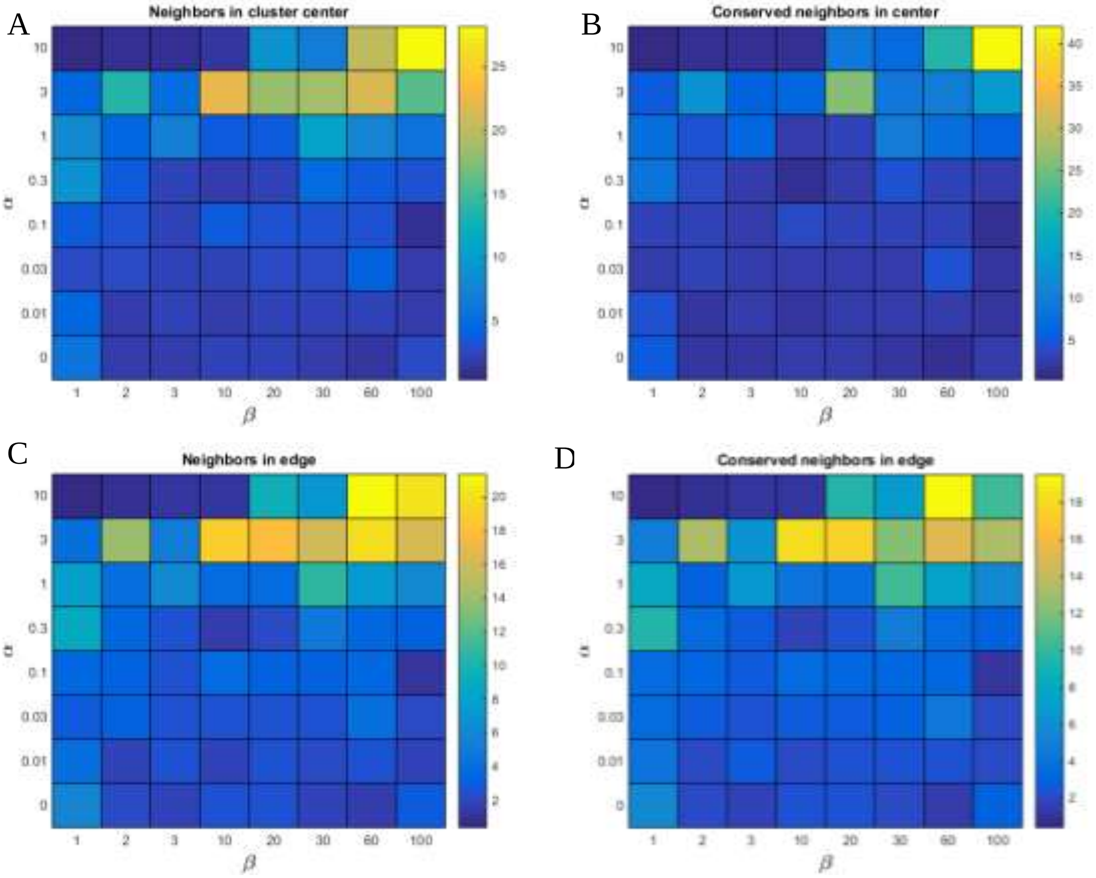


Figure 5-11: Neighbors and conservation for Varying Speed 3D SPP model

The metric for A. neighbors in the cluster center, B. conserved neighbors in the center, C. neighbors in the edge and D. conserved neighbors in the edge on average for clusters in each conditions allows us to map diffusiveness within individual clusters. The similar trends in the two left and right plots indicate that once a cluster forms, individual units tend to retain their positions relative to each other, despite adding variable speed as a parameter to the SPP model.

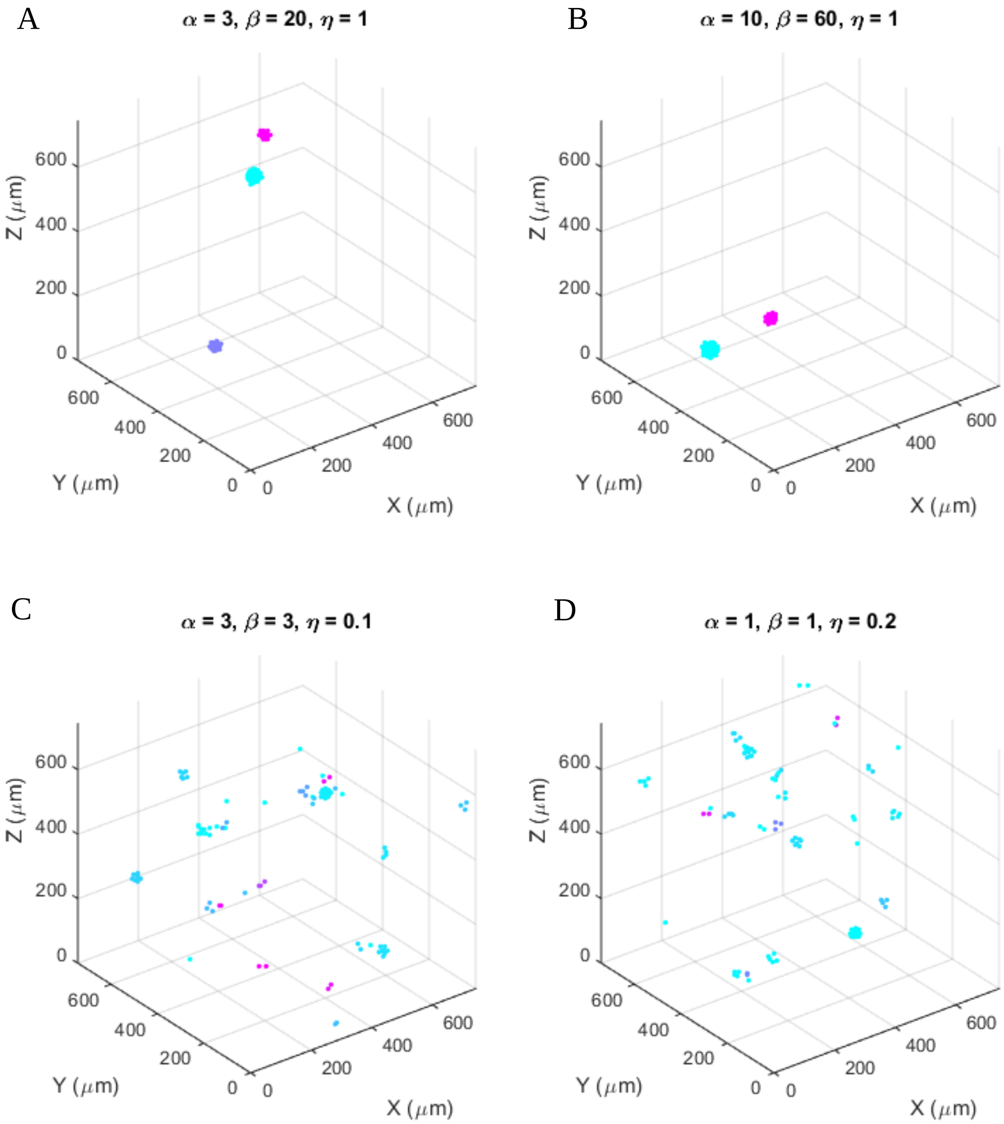


Figure 5-12: Visual Depictions of a few α-β-η combinations

Finale timepoints for (α,β,η) values of A. (3,20,1), B. (10,60,1), C. (3,3,0.1), and D. (1,1,0.2). The circles mark cells or units in an XYZ frame. Units belonging to a single cluster are assigned the same color, units without any cluster are black in color. In all four of these α-β,-η combinations, output data can serve as a computational model system for cellular clusters.

CHAPTER 6. ANALYTICAL METHODS FOR EVALUATING 3D CELL COLLECTIVES

6.1. Introduction

Conventional cell tracking research assumes that the behavior of cells in 2D and 3D is homogenous in time; information from all time points is averaged to deduce the timescales for various cellular behaviors (Zaman et al. 2006). This is typically achieved by fitting cell trajectories to a stochastic random walk model (Dickinson & Tranquillo 1993) according to Equation 6-1 where $<D^2(\tau)>$ is the mean squared displacement (MSD) of aca cells trajectory, a is dimensionality, S is the speed, τ is the time lag, and P is the persistence time, or on average the time for which a particle appears to move in a straight direction before changing direction. This analysis is effective at finding speed and persistence of cells in 2D or 3D, however the basic assumptions of a persistent random walk fail in a systems of cell collectives. A graphic depiction of experimental data from two experimental datasets is presented in Section 4.3, and Figure 6-1 and Figure 6-2. These systems are not correlated within a single field of view at all time, in fact there are intervals in which clusters move away and intervals in which clusters move toward each other. For some smaller clusters, translation is followed by intermittent rotation. Projections of the Z-slices to show the same cluster translating and rotating at different time intervals are in Figure 6-3 and Figure 6-4 respectively.

<u>Equation 6-1</u>

$$\langle D^2(\tau)\rangle = aS^2P\left(\tau - P + Pe^{-\frac{\tau}{P}}\right)$$

Cell collective studies typically calculate the velocity correlation length of monolayers and order parameters. Order parameters are easily adapted for the 3D scenario on a cohort-by-cohort basis, however the correlation length proves elusive. On a two-dimensional monolayer, the radius for which cells are correlated with each other is termed the correlation length r_0 at which a velocity correlation function of the form in Equation 6-2 equals 0. This is effective at distinguishing differences between monolayers as well as characterizing them, yet requires a dense system for a 3D adaptation.

Equation 6-2

$$C(r) =< v(\vec{r}' + \vec{r}, t\,) \times v(\vec{r}', t\,) >$$

This work presents a quantitative approach to 3D cell collectives that have spatial, temporal, and dynamic heterogeneity– it treats each cellular cohort as unique, as well as each moment of time and space as unique and finds motility events within experimental and simulated cell collectives for each individual cohort (Sharma et al. 2015). This study investigates each individual cohort– experimental or simulated, independently, and all consequential analysis methods are followed on a cohort-by-cohort basis.

6.2. Analytical Methods

6.2.1. Order Parameter

Order parameters are useful in characterizing phases and collectivity of a system (Vicsek & Zafeiris 2012). To eliminate noise for calculating order parameter, a smoothing function is run on experimental XYZ position data obtained in Section 4.3 between consecutive time points according to Equation 6-3 where x represents position and t represents time; the interval between consecutive data points ΔT is 10 minutes.

Equation 6-3

$$\vec{x}_t = \frac{\vec{x}_{t-\Delta T} + \vec{x}_t + \vec{x}_{t+\Delta T}}{3}$$

The T_{int} order parameter ϕ is calculated for the cohort between time t+ 0.5*T_{int} and t-(0.5*T_{int} +ΔT) as shown in Equation 6-4 where v is velocity and N is the number of cells in the cohort. T_{int} is set to 1 hour for experimental data

Equation 6-4

$$\phi(t) = \frac{\left|\sum_{i=1}^{N} \vec{v}_{i\,T_{int}}\right|}{\sum_{i=1}^{N} \left|\vec{v}_{i\,T_{int}}\right|}$$

T_{int} is selected by studying time lag dependent Mean Squared Displacement (MSD) vs. time lag τ plots for all cells in a selected cluster. These plots would traditionally be fitted to Equation 5-7 to yield a speed and persistence time, however MSD plots for single clusters, as shown in Figure 6-5: MSD of cells from 1 cluster, suggest that the cells in these high heterogeneity of behavior over intervals as low as 30 minutes. In order to account for tracking, de-drifting and noise induced bias, we doubled this number to set T_{int} = 1 h.

6.2.2. Displacement Quantiles

Displacements of each cell in a cohort between time t+ 0.5*T_{int} and t-(0.5*T_{int} +ΔT) are calculated across the entire timespan of the experiment, where T_{int} = 1 h and ΔT = 10 min (gap between consecutive time points). This results in a distribution with as many values as number of cells in the cohort at each time point. Displacements are squared, and the median, upper and lower quartiles of this distribution are evaluated for all time points of the experiment. The median of these, the Squared Displacement Median (SDM) when plotted against time as in Figure 6-6, has distinct peak like signatures corresponding to intervals of high motion, dubbed motility events.

6.2.3. Motility Event Selection

Since the SDM vs. time data represents a “signal” for each individual cohort, an algorithm can be devised to deduce motility events of individual cohort. This algorithm is created to be dependent on two parameters– T_{int}, as defined above, which also serves as a threshold minimal time for a motility event to occur; and P_{min}, the minimal peak value of SDM for an event to occur.

Initially, Matlab’s built-in peak finding algorithm is used to find all peaks in the SDM vs. time data as shown in Figure 6-6A. The borders of individual peaks are calculated by finding minimum valleys between successive peaks. Then the following order dependent rules are applied to identify motility events:

1. Peaks are merged if the valley between them $> 0.5 \times P_{min}$. This keeps motility events separated by half the minimum peak amount as distinct events, thus providing a threshold for selection. Such functionality is not available in the in-built peak finding algorithm.
2. Peaks are merged if the time gap between them $< 1.5 \times T_{int}$. The results of such merging are shown in Figure 6-6B for one scenario. Again, merging peaks based on a critical width is not a functionality in Matlab’s built in algorithm.

3. Peaks with width < T_{int} are eliminated at this point. Thus peaks with width < T_{int} that do not merge with wider peaks are removed.
4. Peaks with height < P_{min} are eliminated at this point. Similarly peaks with height < P_{min} that do not merge with larger peaks are removed. The results of such elimination for a sample dataset are illustrated in Figure 6-6C.

P_{min}, or the minimum peak height for a motility event, is conservatively set at 60 μm^2, in order to track motion of ~1 3D nucleus diameter and minimize the loss of relevant information. The borders between final peaks are the beginning and end points of individual motility events, as illustrated in Figure 6-6D.

6.2.4. Directional Correlation Function

Once an event is identified, smoothed positions of cells for that event inform correlation functions between all possible cell pairs within a cohort. This correlation function is represented by Equation 6-5 where *i* and *j* are the cell pair, τ is time difference, t is time, and *v* is the velocity (Nagy et al. 2010).

Equation 6-5

$$C_{i,j}(\tau) = \left\langle \left(\frac{\vec{v}_i}{|\vec{v}_i|}\right)_t \cdot \left(\frac{\vec{v}_j}{|\vec{v}_j|}\right)_{t+\tau} \right\rangle$$

This correlation function reaches a maximum peak value at a lag time τ_c; a correlation is considered significant and τ_c is retained for all correlation peaks with a height >0.5. For positive τ_c, cell i lags cell j with duration τ_c; conversely, for negative τ_c, cell j lags cell i with duration τ_c.

6.3. Results

6.3.1. Cellular cohorts have spatial and temporal heterogeneity

Graphs of the SDM plots, corresponding order parameter ϕ, and evaluated motility events are displayed for the 3 clusters each in Gel I FOV 1 and Gel I FOV 2 in Figure 6-7; and for the two clusters each in Gel I FOV 3, Gel II FOV 1 and Gel II FOV 2 in Figure 6-8. Analysis of Gel II Field of View (FOV) 3 is skipped because the cells sink to the bottom and appear to have a 2D morphology after 48 hours. Cohorts do not behave in an identical manner to each other or within themselves across time. Motility events generally concur with high order, although when an event concurs with lowered order rotation is observed corresponding with lesser total displacement. The order, number of cells, and displacement of 61 motility events of experimental clusters from Gel I and II section 4.3 are displayed in Table 6-1 arranged in descending order of total cohort displacement $|d|$. The highest observed translation event has a duration of 6 hours and measures ~90 μm, corresponding to ~8 3D cell diameters. To visualize a few motility events, cell trajectories derived from raw data with events from six different motility events are mapped in Figure 6-9. Panels A and B show rotation and translation events from the same cohort, Panel C is an example of a larger cohort (23 cells) undergoing rotation, Panels D, E and F are translation examples.

6.3.2. Cohorts have somewhat correlated trajectories

The directional correlation function peak values τ_c for all cell pairs from selected motility events are displayed in Figure 6-10 where the y-axis or rows represent cell i and the columns represent cell j according to Equation 6-5. Thus a positive τ_c value indicates that the cell in the row lags behind the cell in the column. A negative τ_c value indicates the cell in the row leads the cell in the column. For the two examples of rotating cohorts, cells are either generally not correlated with each other (Panel B), which is to be expected when observing rotating units moving against each other; or cells that are leading most of the group (Panel C, Cell B), are also following other cells (Panel C, Cell U). For the translating cohorts, correlation lags are spread up to half the observed motility interval apart and no clear cells are leading or lagging behind all others. The percentage of all correlation lags for these intervals are plotted in Figure 6-11. Most cell-pairs have correlated trajectories with no lag, however in the case of translation, trajectory lags can be up to 3 hours. Even within translating cohorts, all cell pairs are not correlated. This may be due to loss and gain of trajectories due to tracking errors or cell proliferation, as well as actual lack of correlation.

6.3.3. Simulated cohorts have motility events

For the simulated conditions that could be replicating cellular systems revealed in the tables in Section 5.5, motility events are extracted using the algorithm presented in section 6.2.3. The same T_{int} and P_{min} values as experimental cohorts are applied, to verify the scaling of the system. A few of these are selected and displayed, along with the order parameter, in Figure 6-12. It is apparent that the event selection algorithm fails on homogenous data that is highly ordered (Panel B), but is able to extract peaks from data with more variation across time. Motility events in simulated cohorts coincide with high order for the most part, with some exceptions (Panel B), which could also be indications of rotation. Additionally for these cases, SPP simulations before applying variable speed have larger fluctuations in order parameter ϕ. Eight events from these datasets are further analyzed– ϕ, |d| and number of cells N_c for these are displayed in Table 6-2. These suggest at least 1 translation event of ~82 μm over ~10 hours, and 1 likely rotation event with low order ϕ and a |d| value of only ~4 μm. Trajectories of simulated cohorts from SPP simulations are displayed in Figure 6-13 and those from variable-speed SPP simulations are in Figure 6-14. Since no tracking induced errors occur, these trajectories are continuous for the entire duration of the motility event. Additionally, these trajectories show a distinct effect of cellular repulsion in high β cases with noise-like traces.

6.3.4. Simulated cohorts are not perfectly correlated

Cell-pairwise correlation maps within the motility events considered are depicted for SPP simulated cohorts in Figure 6-15 and SPP variable-speed simulated cohorts in Figure 6-16. Interestingly, not all cell pairs are correlated with each other in direction. Since these simulated cohorts do not have acceptable correlation peaks between all cell pairs, we can deduce that the pairwise correlations failing within the experimental cohorts are not necessarily due to tracking related errors. The highest incidence of correlations are observed in Figure 6-16 Panel A and D, which are cohorts that translate 82 and 24 μm respectively. Within these correlations there are clear cell columns that are largely blue, indicating these are at the front of the cohort or leading; and those that are largely pink, indicating these are at the back and following. However Figure 6-16 Panel C represents a cohort that translates 34 μm and is almost entirely uncorrelated, likely due to curved trajectories.

The percent incidence of trajectory lags τ_c for events from simulated SPP cohorts are shown Figure 6-17. These cohorts have signatures similar to experimental cohorts, with maximal correlations at τ_c values of 0 min, the difference being that these have a more even spread both directions from 0 min, corresponding more to the rotating cohorts in panels A and C of Figure 6-11. The percent incidence of trajectory lags τ_c for events from simulated SPP cohorts with variable speed are shown in Figure 6-18. These cohorts have maximum correlations at the extremes of the range tested, indicating the effect of high β values– the system itself undergoes fluctuations so motion is correlated over longer intervals.

6.4. Discussion

The analytical tools presented in this section are able to distinguish between motility events and modes within individual cellular cohorts. This kind of motility has been suspected but unobserved in 3D cohorts, and to date not many analytical methods exist. Some cohorts are extremely motile throughout the observed 48 hour interval, for example Gel I FOV 1 Cohort #3 Figure 6-7E; whereas others such as Gel II FOV 2 Cohort #1 have only single motility events as displayed in Figure 6-8E. Therefore the algorithm is able to distinguish between clusters with more and less potential energy to move, simply by identifying the frequency and occurrence of motility events. The events themselves generally correspond with higher order or translation, but may correspond with lower order and translation in which case rotation is likely occurring. Within simulated cohorts, the event finding algorithm is able to distinguish events when a temporal heterogeneity exists.

The larger trajectory lags between cell-pairs that occur within a translating cohort as observed in our system may represent information transfer and polarity along the cohort. Correlation functions serve as a method to understand leader follower relationships within the cohort. It is apparent that in the case of the rotating cohort presented in Figure 6-10C, cells that are serving as leaders are also following other cells. In simulated cohorts, the SPP model appears to recreate trajectory signatures of experimental systems but the SPP model with variable speed has large correlation lags as the dominant percentage of correlations. Thus cells in the variable speed simulated cohorts are held together at longer timescales and dominated by repulsion over shorter timescales.

Emergent motility events arise in the absence of external or forced driving agents, and are stochastic, as in the case of the two clusters that merged (Figure 6-1C and D). There are examples of other similarly sized clusters that do not merge and in fact move in opposite directions (Figure 6-2E and F). The transient nature of these events suggests that this system displays stochasticity and plasticity, both suspected to occur in cancer pathologies (Friedl et al. 2004). The methods presented here can also be extended to any collective system that presents with spatial and temporal heterogeneity, where each cluster can be treated independently.

6.5. Tables

FOV	C	Ti (h)	Tf (h)	|d| (μm)	<ϕ>	Nc
Gel I FOV 2	1	3:20	9:30	89.06	0.92	10
Gel II FOV 1	1	1:30	5:50	53.09	0.84	19
Gel I FOV 1	3	1:50	6:50	49.07	0.93	4
Gel I FOV 3	2	1:10	7:0	34.39	0.87	7
Gel I FOV 1	2	0:30	4:10	33.80	0.85	13
Gel II FOV 1	1	18:40	22:40	31.33	0.87	24
Gel II FOV 1	1	23:40	25:40	30.28	0.93	5
Gel I FOV 2	2	0:40	4:0	28.30	0.91	6
Gel I FOV 2	3	0:30	4:20	26.33	0.81	14
Gel II FOV 1	2	2:00	5:30	25.08	0.77	14
Gel I FOV 3	2	40:20	43:40	23.88	0.67	22
Gel I FOV 3	2	29:20	32:10	21.11	0.87	18
Gel I FOV 2	3	21:30	24:30	20.62	0.77	17
Gel II FOV 2	1	1:30	3:50	19.64	0.82	10
Gel I FOV 3	2	26:30	29:20	19.29	0.75	15
Gel I FOV 2	1	16:00	18:20	19.22	0.80	9
Gel I FOV 2	1	22:30	24:40	19.20	0.81	9
Gel I FOV 1	1	1:20	4:20	19.20	0.79	19
Gel II FOV 2	2	1:00	3:0	18.82	0.78	11
Gel I FOV 2	2	15:40	18:40	17.97	0.84	8
Gel I FOV 1	1	40:50	43:20	17.29	0.84	31
Gel I FOV 3	2	35:30	38:0	16.21	0.85	21
Gel I FOV 1	3	37:30	40:10	15.35	0.69	6
Gel I FOV 2	2	36:30	39:0	14.60	0.50	26
Gel I FOV 1	1	26:50	29:40	14.41	0.86	24
Gel I FOV 1	1	37:50	40:10	14.27	0.87	29
Gel II FOV 2	2	3:00	5:40	13.54	0.84	8
Gel I FOV 3	1	0:30	3:50	12.61	0.72	13
Gel I FOV 1	1	33:10	35:10	12.53	0.86	26
Gel I FOV 1	3	23:00	24:40	12.36	0.78	6
Gel I FOV 2	1	0:50	2:10	12.19	0.82	14
Gel II FOV 2	2	44:30	46:40	12.14	0.76	15
Gel II FOV 2	2	25:10	27:20	11.43	0.84	15
Gel I FOV 1	3	9:50	11:20	11.14	0.83	4
Gel II FOV 2	2	41:20	43:0	11.12	0.77	19

Gel I FOV 2	1	30:00	31:20	11.09	0.80	14
Gel I FOV 1	3	34:50	36:40	11.08	0.77	8
Gel II FOV 1	1	27:10	28:10	11.03	0.95	3
Gel I FOV 1	3	13:00	14:50	10.82	0.84	5
Gel I FOV 1	1	20:20	22:10	10.81	0.79	24
Gel I FOV 3	2	44:40	46:20	10.64	0.74	22
Gel I FOV 1	3	28:50	30:30	10.43	0.87	8
Gel I FOV 1	2	28:30	29:40	10.31	0.85	17
Gel I FOV 3	1	26:00	27:30	9.76	0.60	18
Gel I FOV 1	2	17:40	19:30	9.69	0.90	14
Gel I FOV 2	2	23:30	25:50	9.58	0.75	10
Gel I FOV 1	2	36:40	37:50	9.10	0.85	22
Gel I FOV 2	1	9:50	11:10	8.92	0.79	13
Gel I FOV 1	3	11:40	13:0	7.27	0.76	5
Gel II FOV 1	2	28:10	29:40	6.69	0.80	15
Gel I FOV 1	2	5:00	6:30	6.57	0.75	14
Gel I FOV 2	3	32:30	33:50	6.33	0.72	23
Gel I FOV 2	1	27:50	29:40	5.96	0.86	13
Gel I FOV 2	1	34:40	35:40	5.88	0.82	13
Gel I FOV 1	3	7:10	9:0	5.46	0.73	5
Gel I FOV 1	3	0:50	1:50	5.39	0.87	3
Gel I FOV 1	3	27:00	28:10	4.71	0.69	8
Gel I FOV 1	2	21:10	22:10	4.61	0.66	15
Gel I FOV 1	3	42:30	44:40	3.84	0.78	8
Gel I FOV 1	3	44:40	45:40	3.84	0.56	8
Gel I FOV 1	3	17:10	18:50	3.35	0.39	7

Table 6-1: Experimental Motility Events

FOV is the field of view for a given Gel, C is the cohort ID as identified by a clustering algorithm, T_i is the starting time of the event, T_f is the final time of the event, $|d|$ is the displacement of the center of the cohort during the event, $\langle\phi\rangle$ is the average order parameter during the event, and N_c is the average number of cells for the cohort during the event. Events are organized from highest to lowest $|d|$.

Simulation	**C**	**T_i (h)**	**T_f (h)**	**$\|d\|$ (μm)**	**$\langle\phi\rangle$**	**Nc**
SPP α 1 β 10	17	20:20	22:10	9.89	0.79	12
SPP α 1 β 10	17	39:30	42:40	19.86	0.73	12
SPP α 0.7 β 1	22	20:30	22:10	9.30	0.69	8
SPP α 0.7 β 1	22	33:20	35:20	3.95	0.50	8
Var $\|v\|$ SPP α 10 β 60 η 1	3	10:00	19:50	82.86	0.83	27
Var $\|v\|$ SPP α 10 β 60 η 5	3	33:50	36:40	12.69	0.84	24
Var $\|v\|$ SPP α 3 β 20 η 1	2	18:20	25:30	34.00	0.81	24
Var $\|v\|$ SPP α 3 β 20 η 1	2	43:50	47:20	24.00	0.87	24

Table 6-2: Motility Events for a Few Simulated Cohorts

For the presented simulated conditions, C is the cohort ID as identified by a clustering algorithm, T_i is the starting time of the event, T_f is the final time of the event, $|d|$ is the displacement of the center of the cohort during the event, $\langle\phi\rangle$ is the average order parameter during the event, and N_c is the average number of cells for the cohort during the event. Events are organized from highest to lowest $|d|$.

6.6. Figures

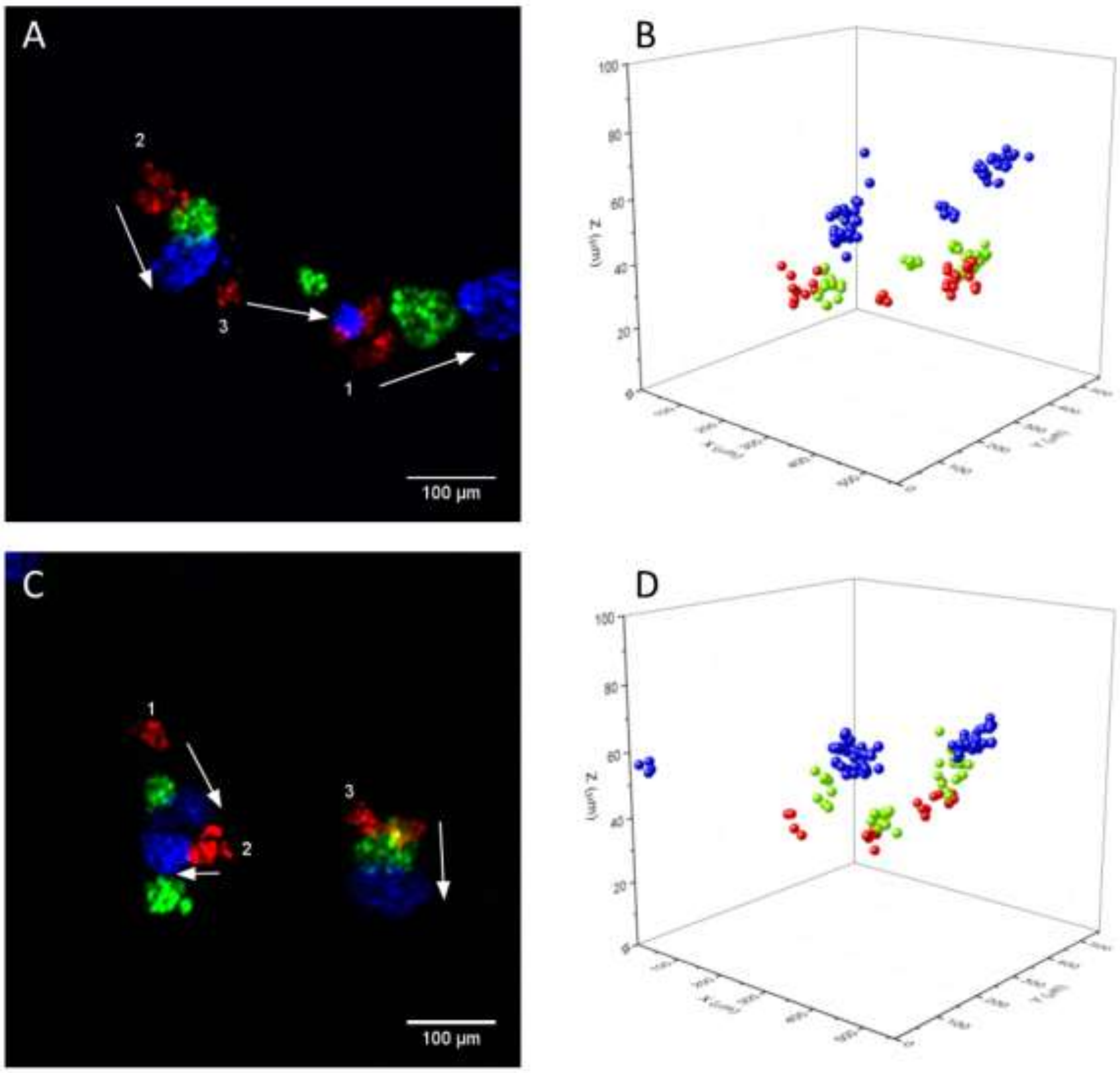


Figure 6-1: Z-Sum projection and visualization of Gel I FOV 1, Gel I FOV2

A. Z-projection of Gel I FOV 1, B. 3D rendering of nuclear tracks of Gel I FOV 1, C. Z-projection of Gel I FOV 2, D. 3D rendering of nuclear tracks of Gel I FOV 1. MDCK GFP-NLS cell cohorts at 0 h (red), 24 h (green) and 48 h (blue). The numbers in white indicate cohort number as determined by a clustering algorithm. Panels B and D are 3D renderings of nuclear tracking corresponding to Panel A and C respectively.

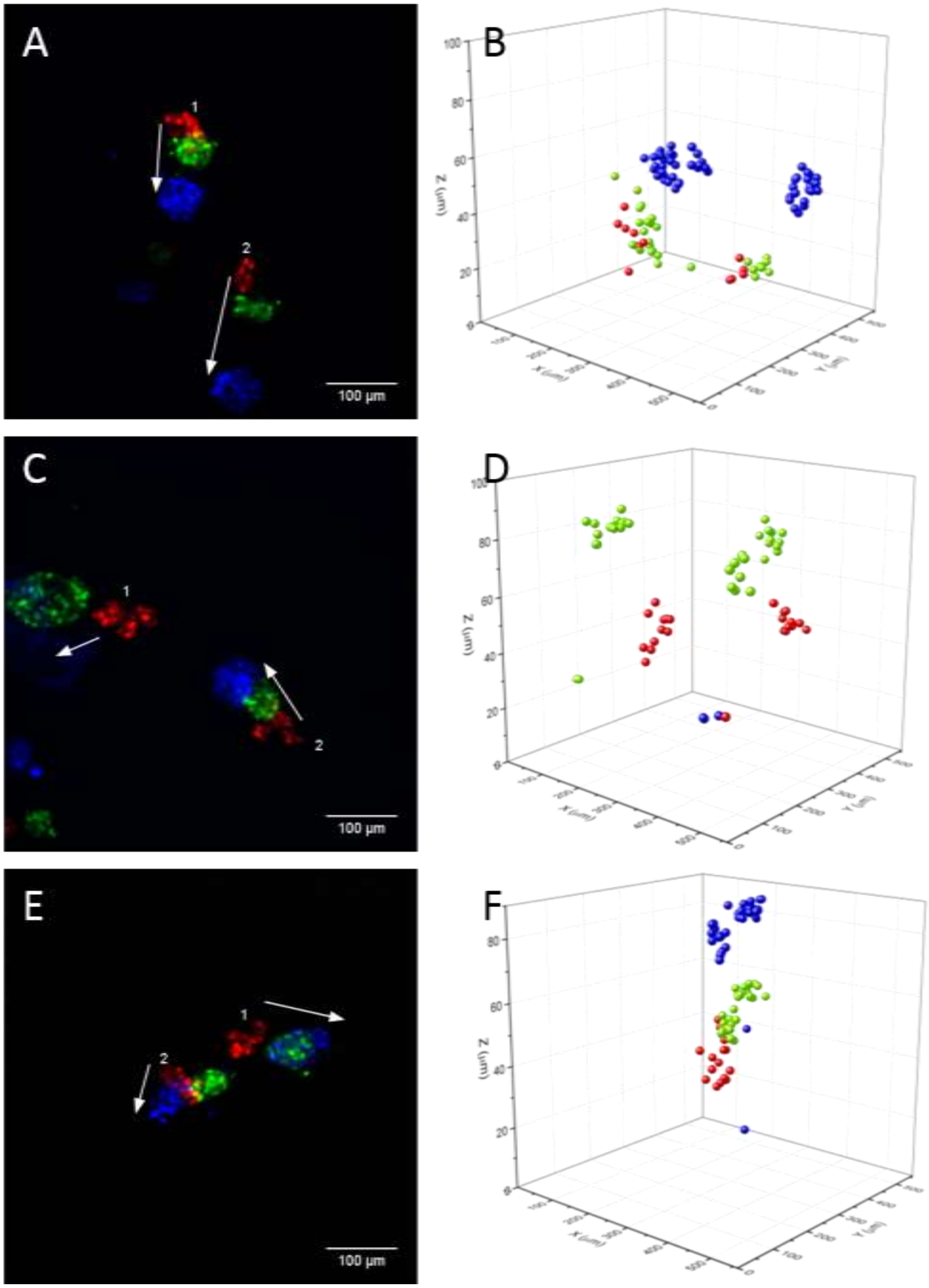


Figure 6-2: Z-Sum projection and visualization of Gel I FOV 3, Gel II FOV1, and Gel II FOV 2

A. Z-projection of Gel I FOV 3, B. 3D rendering of nuclear tracks of Gel I FOV 3, C. Z-projection of Gel II FOV 1, D. 3D rendering of nuclear tracks of Gel II FOV 2, E. Z-projection of Gel I FOV 3, F. 3D rendering of nuclear tracks of Gel II FOV 2. MDCK

GFP-NLS cell cohorts at 0 h (red), 24 h (green) and 48 h (blue). The numbers in white indicate cohort number as determined by a clustering algorithm. Panels B and D are 3D renderings of nuclear tracking corresponding to Panel A and C respectively.

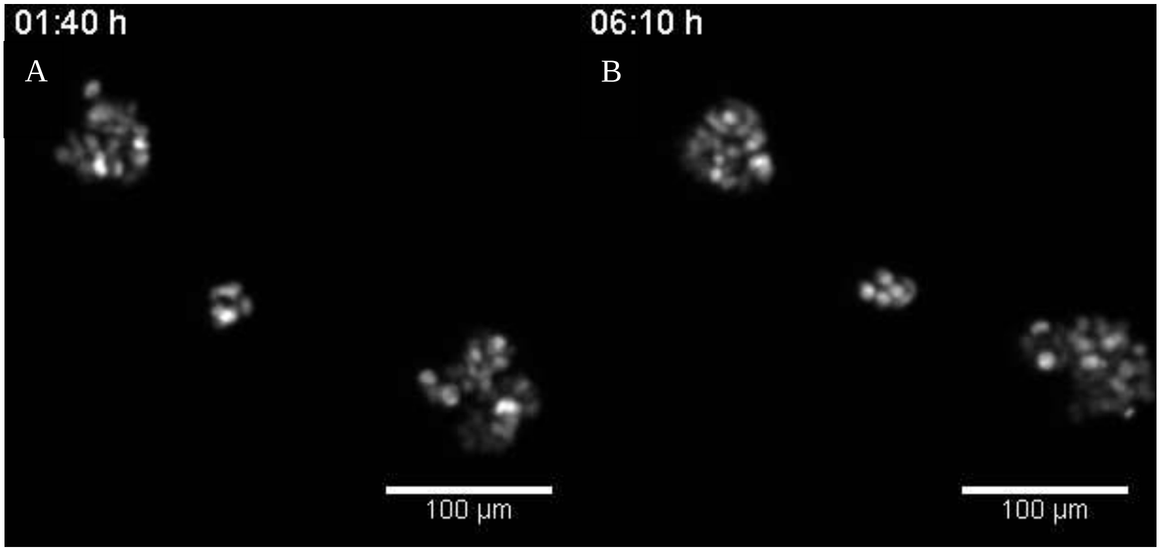


Figure 6-3: Gel I FOV 1 C3 translates

The small cohort (# 3) in Gel I FOV 1 translates toward the larger cluster (Cluster 1) over a few hours. A. Z-projection of a small section 1h 40 minutes after the start of the experiment. B. Z-projection of the same section 6 h 10 min after the start of the experiment.

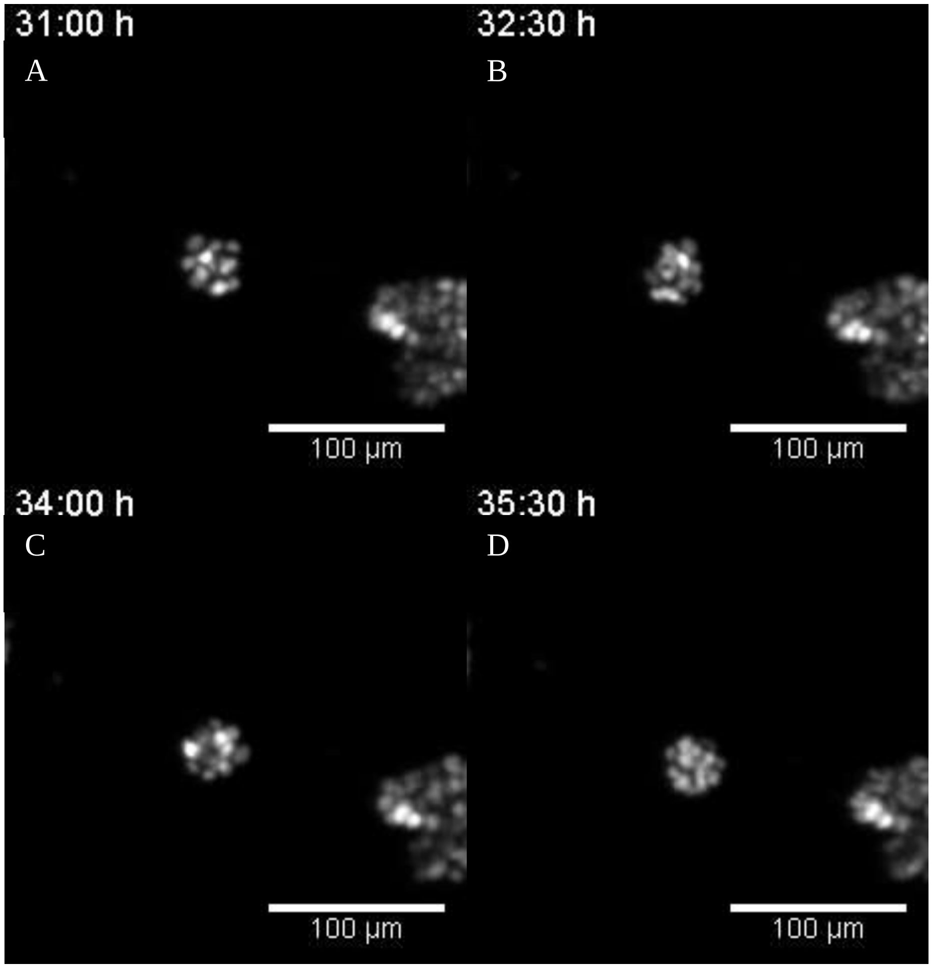


Figure 6-4: Gel 1 FOV 1 C3 rotates

A section of Gel I FOV 1 at A. 31 h, B. 32.5 h, C. 34 h and D. 35.5 h after the start of the experiment. Cohort 3 (smallest) rotates in this period, the entire cluster geometry undergoes a rotation around all dimensions, so that the brightest cell nuclei goes from top left (A) to top right(B) to left again(C) to underneath(D) the cluster, indicating a rotation in all three dimensions available.

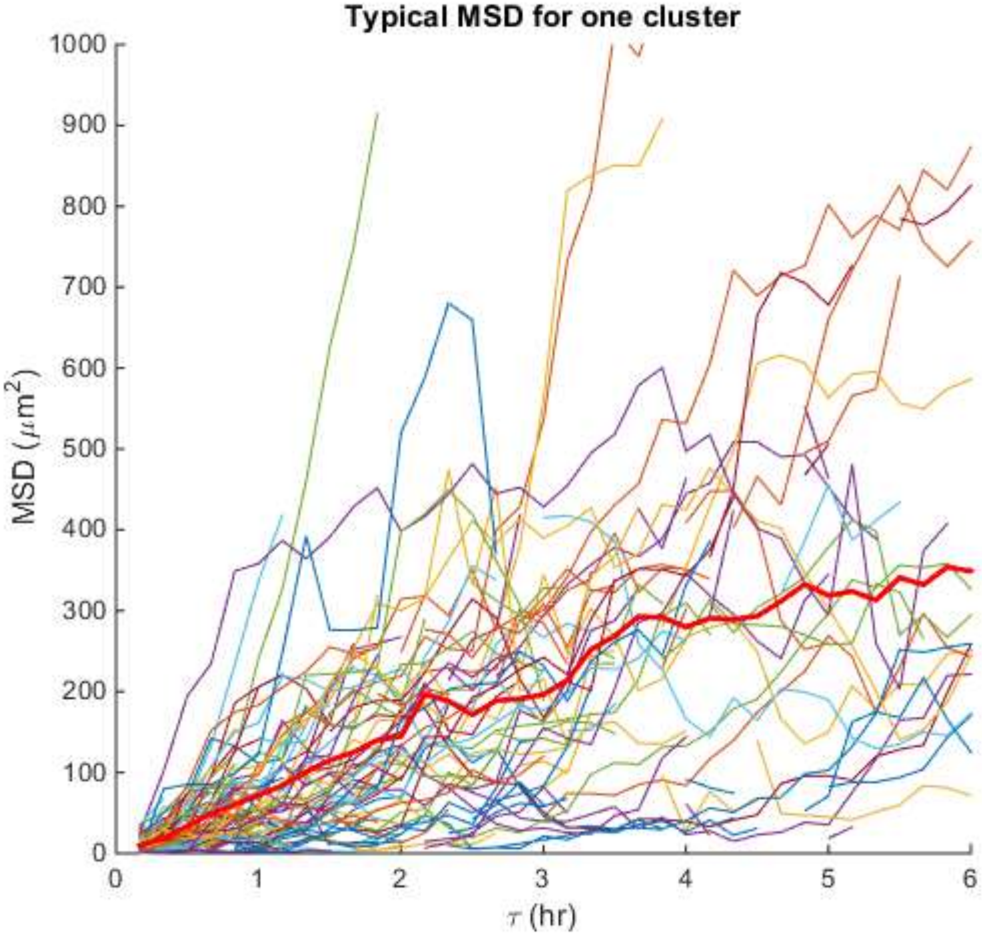


Figure 6-5: MSD of cells from 1 cluster

The Mean Squared Displacement (MSD) of cells in a Gel I FOV 1Cohort #3 from the 48 hour interval plotted up to a time lag τ of 6 hours. Red bold line depicts the average of all cells plotted. Cells within a single cluster have variation between MSD plots over as little as 30 min.

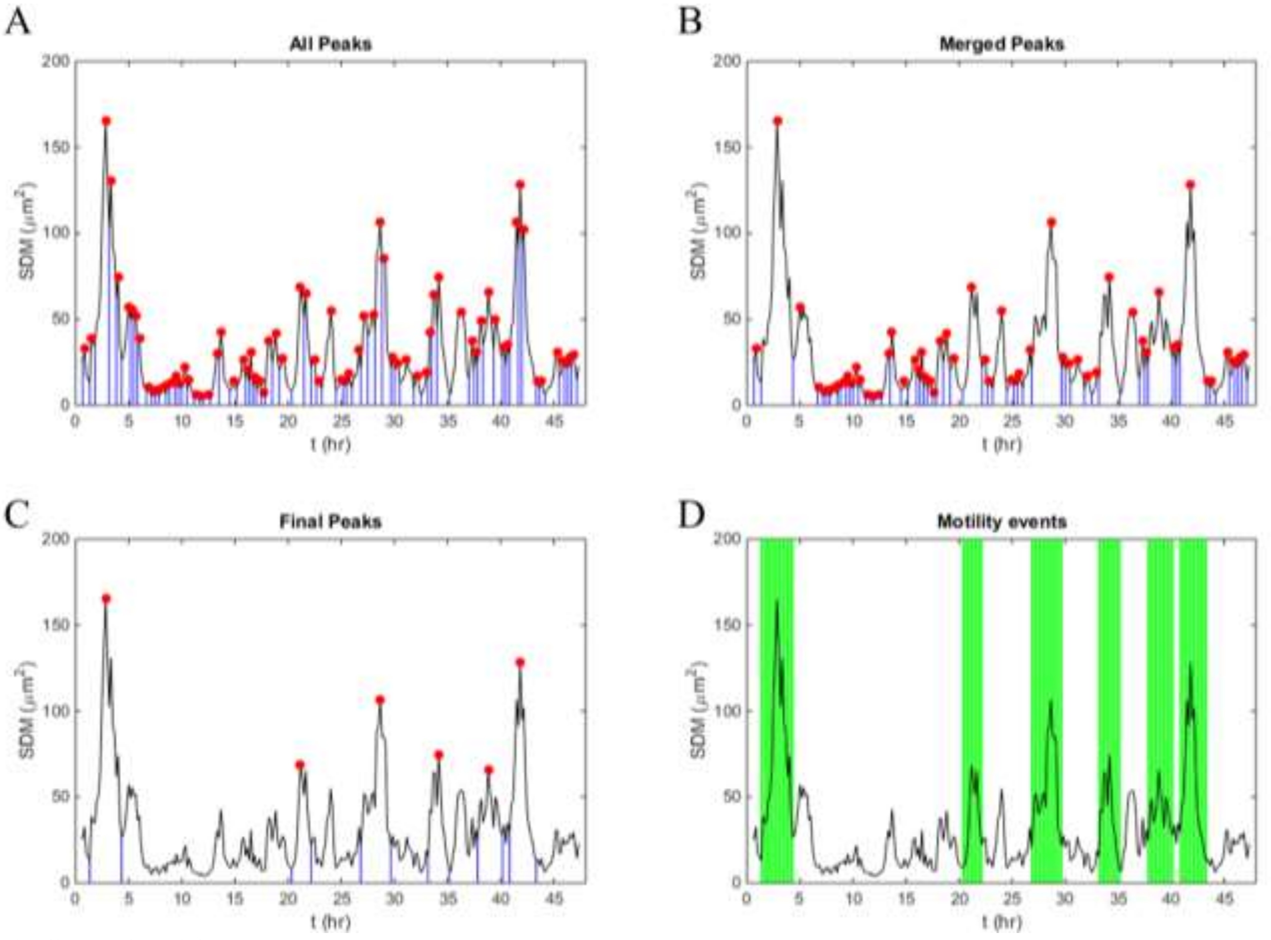


Figure 6-6: Event-Finding Algorithm

SDM data vs. time for Gel I FOV 1 Cohort 1 against time is depicted in all four panels (black). A. All peaks (red dots) and their borders (blue) are evaluated using Matlab's peak finding algorithm. B. Peaks remaining after all peaks with valleys $>P_{min}$ and time gaps $< 1.5\ T_{int}$ have been merged. C. Peaks remaining after all peaks $<P_{min}$ in magnitude and $< T_{int}$ in interval have been eliminated. D. Motility events are shaded in green.

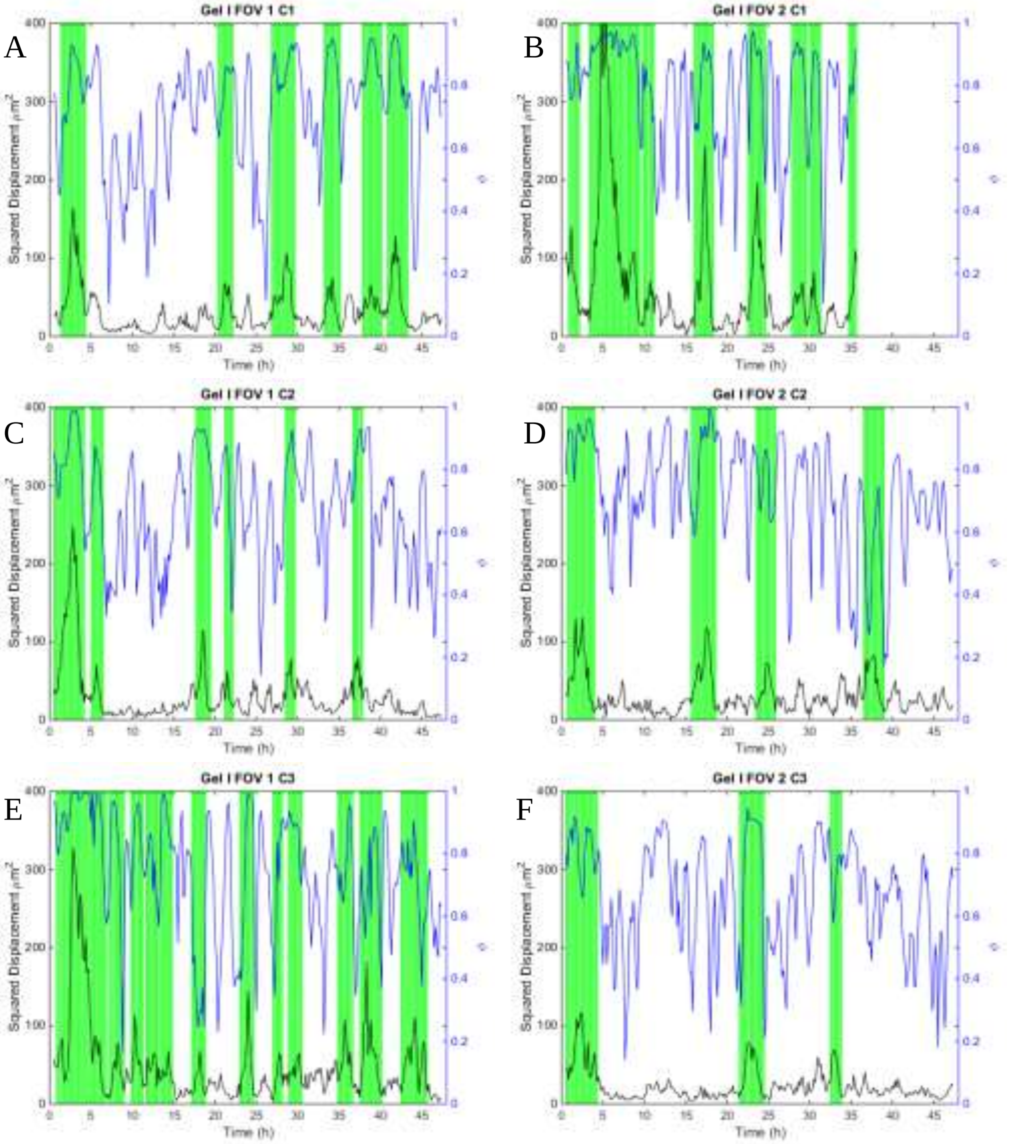


Figure 6-7: Motility Events for individual cohorts in Gel I FOV1 and Gel I FOV 2

Black lines along left axis plot the Squared Displacement Median SDM. Green shaded regions mark identified motility events. Blue lines along right axis mark the order parameter ϕ. Each figure is for an individual cellular cohort A. Gel I FOV 1 Cohort 1, B. Gel I FOV 2 Cohort 1, C. Gel I FOV 1 Cohort 2, D. Gel I FOV 2 Cohort 2, E. Gel I FOV 1 Cohort 3, and F. Gel I FOV 2 Cohort 3.

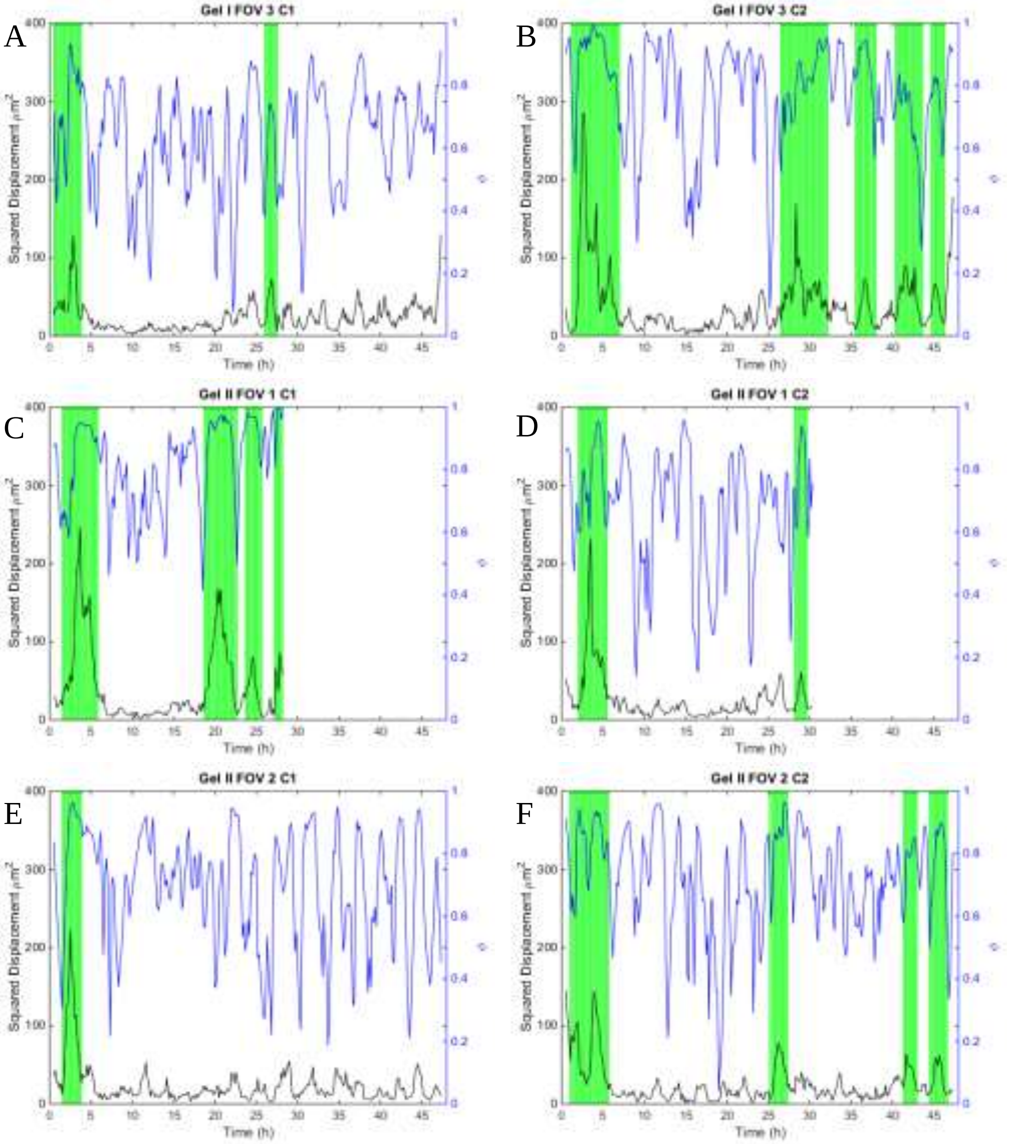


Figure 6-8: Motility Events for individual cohorts in Gel I FOV3, Gel II FOV 1, and Gel II FOV 2

Black lines along left axis plot the Squared Displacement Median SDM. Green shaded regions mark identified motility events. Blue lines along right axis mark the order parameter ϕ. Each figure is for an individual cellular cohort A. Gel I FOV 3 Cohort 1, B. Gel I FOV 3 Cohort 2, C. Gel II FOV 1 Cohort 2, D. Gel II FOV 1 Cohort 2, E. Gel II FOV 2 Cohort 1, and F. Gel II FOV 2 Cohort 2.

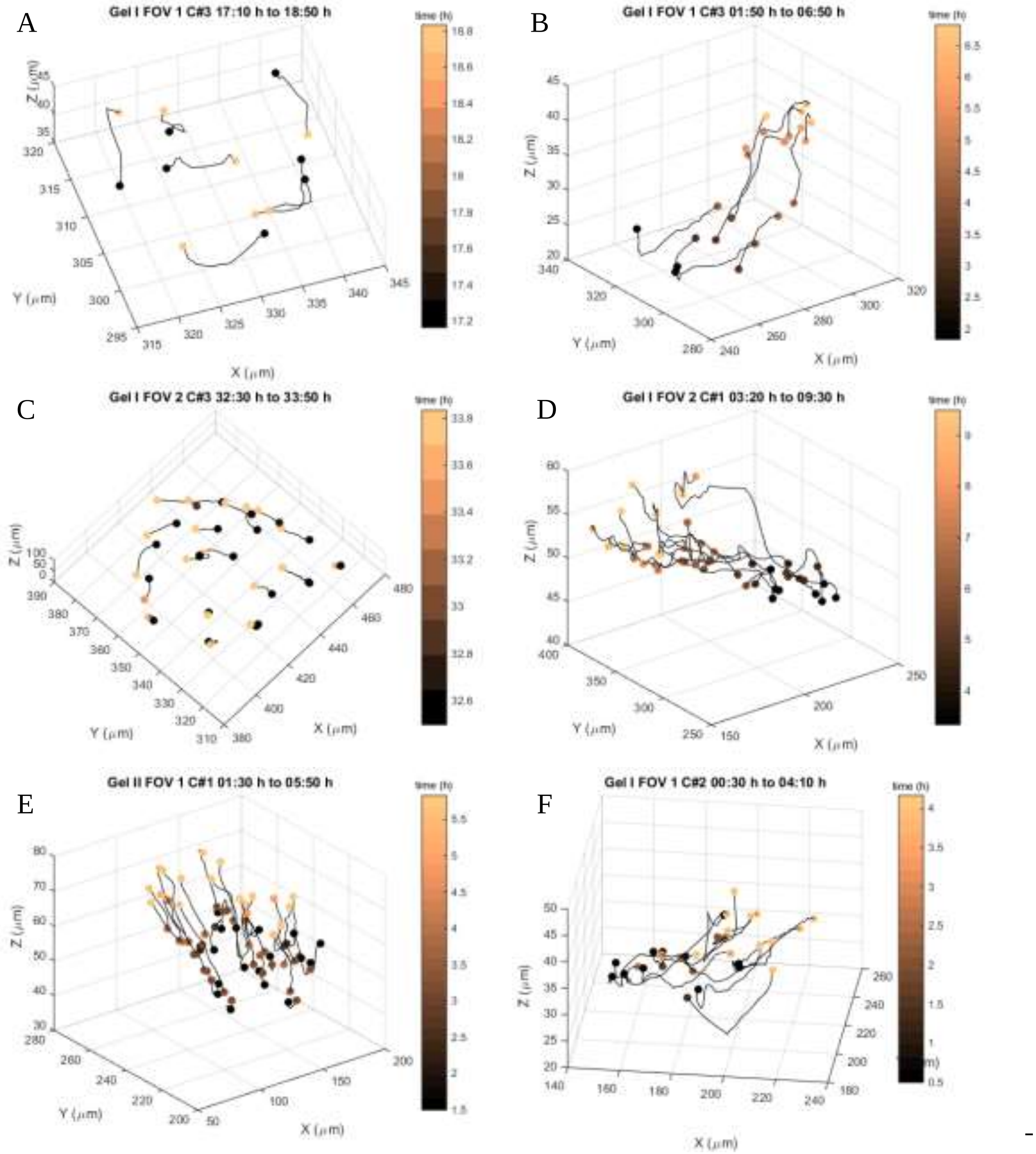


Figure 6-9: Cell Trajectories of Motility Events of a few cohorts in experments

Spots mark nuclei connected by black lines denoting trajectories. Colors of circles represent time of beginning and end of trajectory. A. Gel I FOV 1 Cohort # 3 from 17:10 h to 18:50 h, B. Gel I FOV 2 Cohort # 3 from 1:50h to 6:50 h, C. Gel I FOV 3 Cohort # 3 from 32:30 h to 33:50 h, D. Gel 1 FOV 2 Cohort #1 from 3:20 h to 9:30 h, E. Gel II FOV 1 Cohort #1from 1:30 h to 5:50 h and F. Gel I FOV 1 Cohort #2 from 0:30 h to 4:10 h.

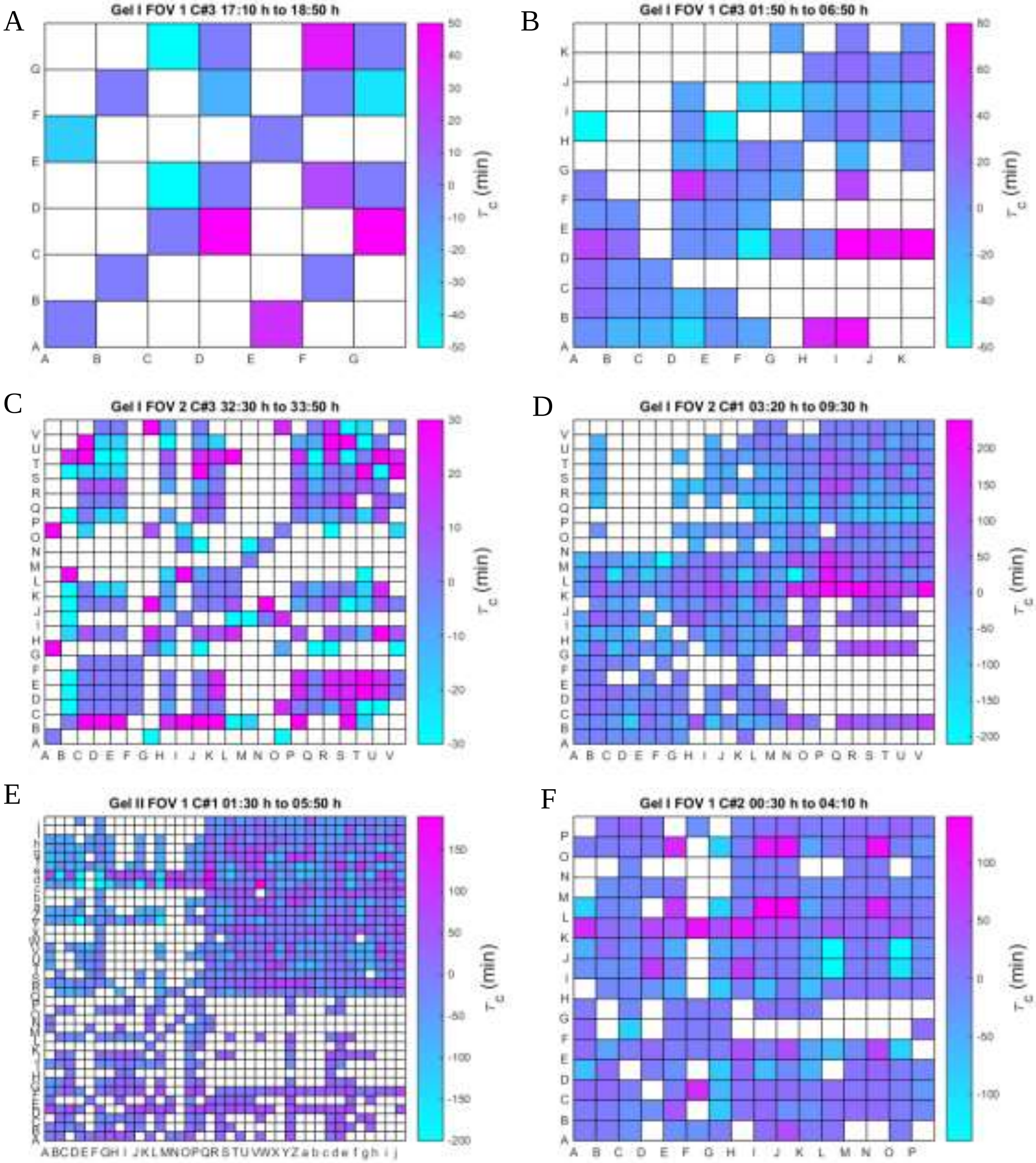
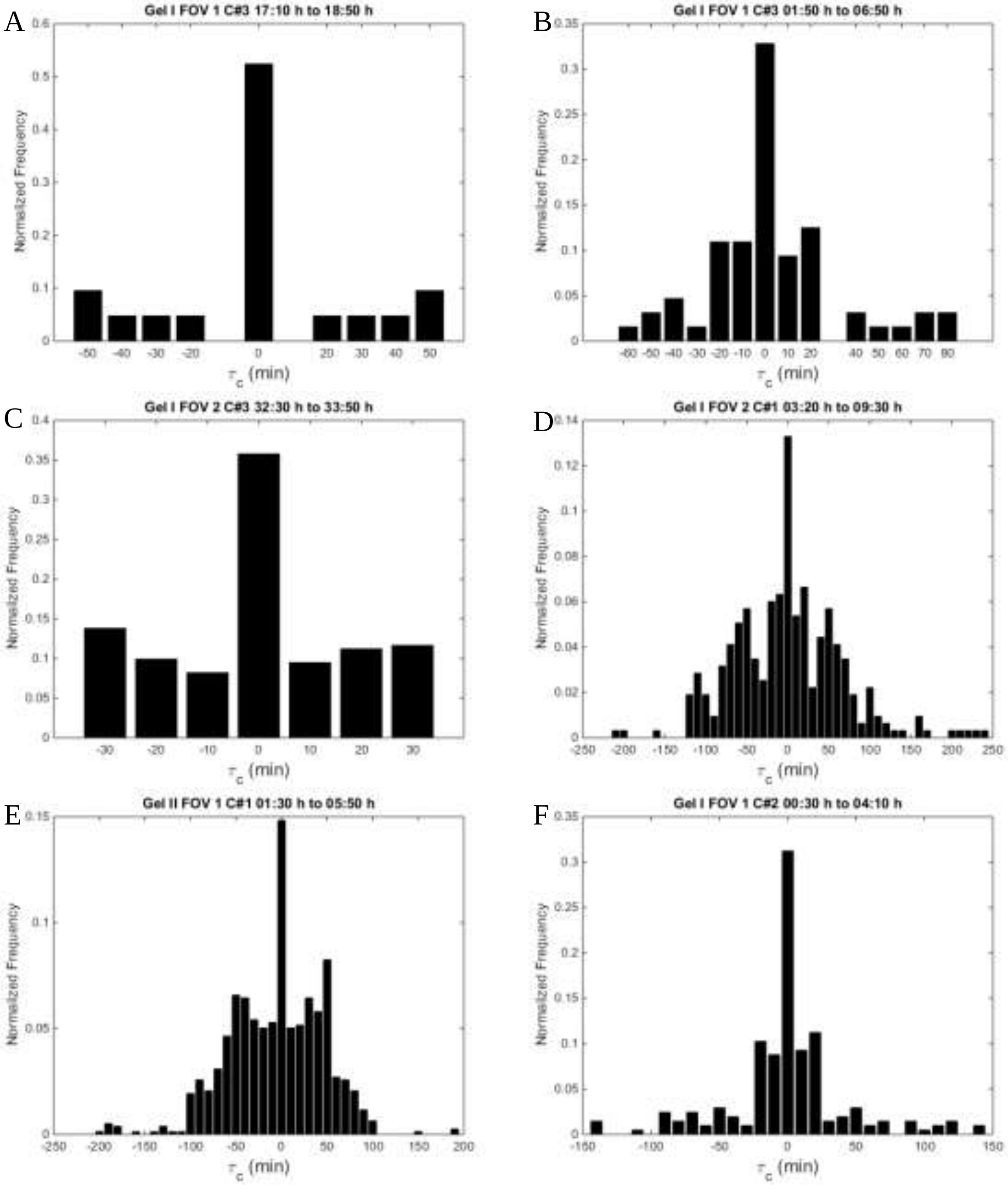


Figure 6-10: Correlation Values τ_c of Experimental Clusters within motility events

Each map shows the correlation time lag τ_c of cell pairs during a given motility event. Cells are named A through Z, followed by lowercase a. Cells are always correlated with each other with a τ_c of 0 minutes. A. Gel I FOV 1 Cohort # 3 from 17:10 h to 18:50 h, B. Gel I FOV 2 Cohort # 3 from 1:50h to 6:50 h, C. Gel I FOV 3 Cohort # 3 from 32:30 h to 33:50 h, D. Gel 1 FOV 2 Cohort #1 from 3:20 h to 9:30 h, E. Gel II FOV 1 Cohort #1from 1:30 h to 5:50 h and F. Gel I FOV 1 Cohort #2 from 0:30 h to 4:10 h.



<u>Figure 6-11: Percentage of τ_c values for experimental motility events</u>

Each graph shows the percentage of the correlation time lag τc of cell pairs during a given motility event. A. Gel I FOV 1 Cohort # 3 from 17:10 h to 18:50 h, B. Gel I FOV 2 Cohort # 3 from 1:50h to 6:50 h, C. Gel I FOV 3 Cohort # 3 from 32:30 h to 33:50 h, D. Gel 1 FOV 2 Cohort #1 from 3:20 h to 9:30 h, E. Gel II FOV 1 Cohort #1from 1:30 h to 5:50 h and F. Gel I FOV 1 Cohort #2 from 0:30 h to 4:10 h.

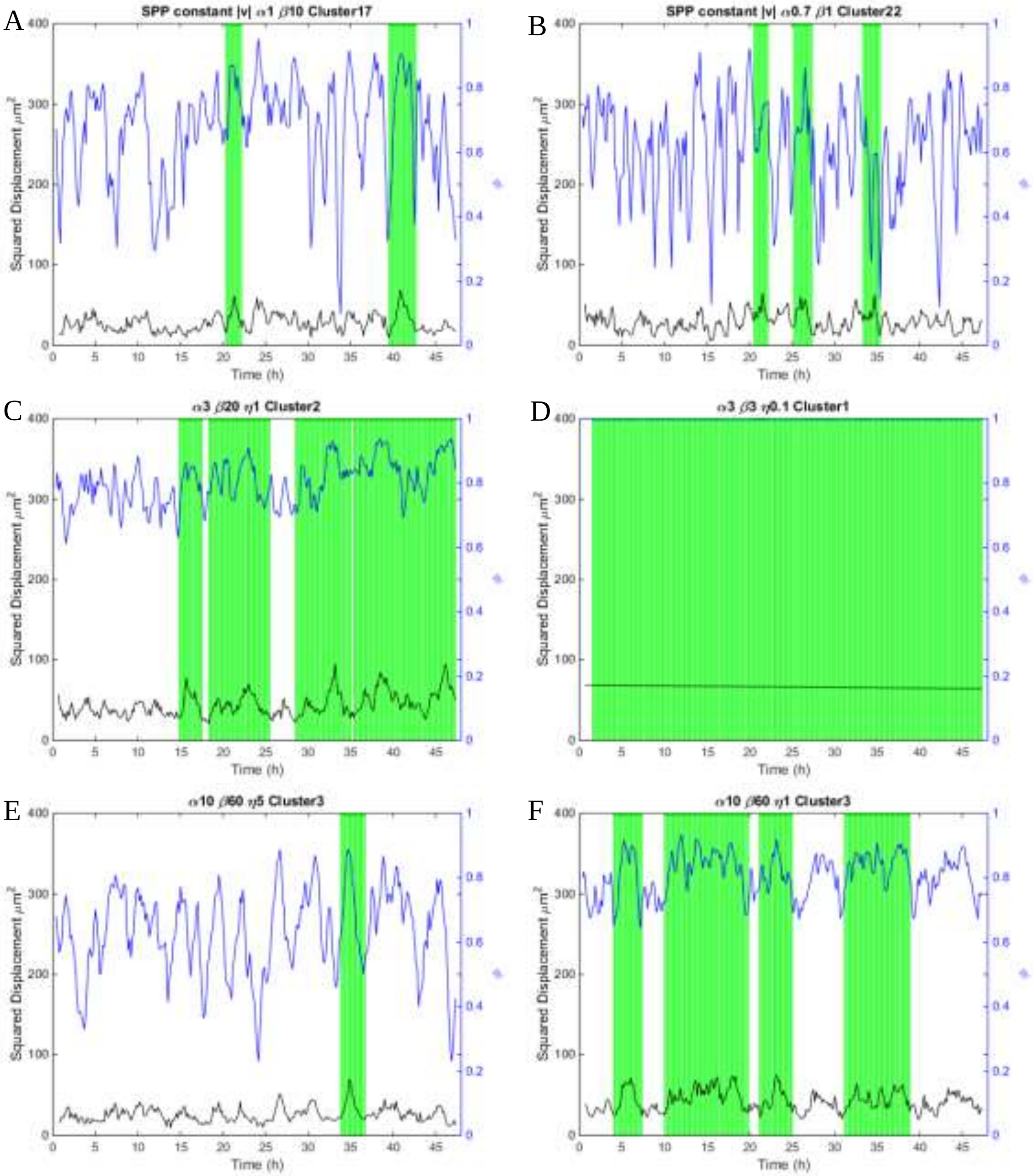


Figure 6-12: Motility events for simulated cohorts

Black lines along left axis plot the Squared Displacement Median SDM. Green shaded regions mark identified motility events. Blue lines along right axis mark the order parameter ϕ. Each figure is for an individual cellular cohort. A. SPP model (α,β) values (1,10) Cohort 17, B. SPP model (α,β) values (0.7,1) Cohort 22, C. Variable Speed SPP (α,β,η) values (3,20,1) Cohort 2, D. Variable Speed SPP (α,β,η) values (3,3,0.1) Cohort 1, E. Variable Speed SPP (α,β,η) values (10,60,1) Cohort 3, and F. Variable Speed SPP (α,β,η) values (10,60,1) Cohort 3.

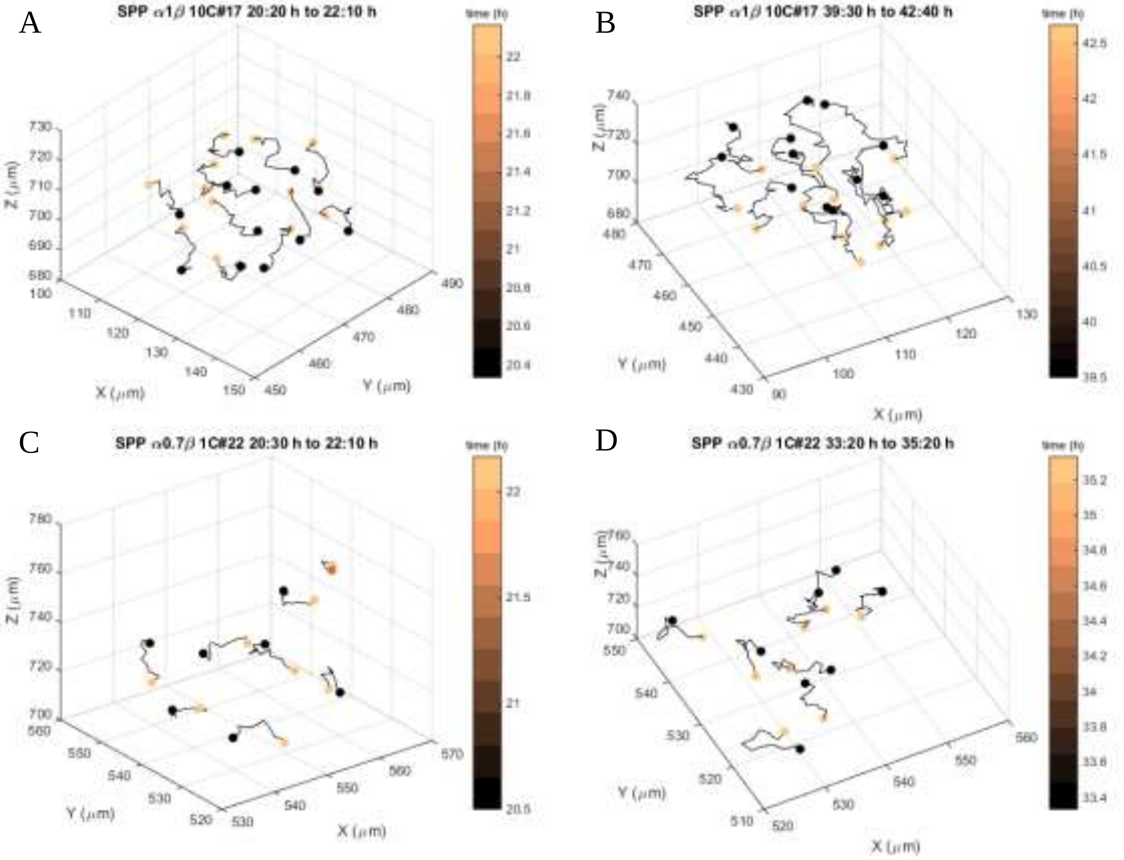


Figure 6-13: Cell Trajectories of Motility Events of SPP simulated cohorts

Spots mark nuclei connected by black lines denoting trajectories. Colors of circles represent time of beginning and end of trajectory. A. (α,β) values of (1,10) Cohort # 17 from 20:20 h to 22:10 h, B. (α,β) values of (1,10) Cohort # 17 from 39:30 h to 42:40 h C. (α,β) values of (0.7,1) Cohort # 22 from 20:30 h to 22:10 h and D. (α,β) values of (0.7,1) Cohort # 22 from 33:20 h to 35:20 h

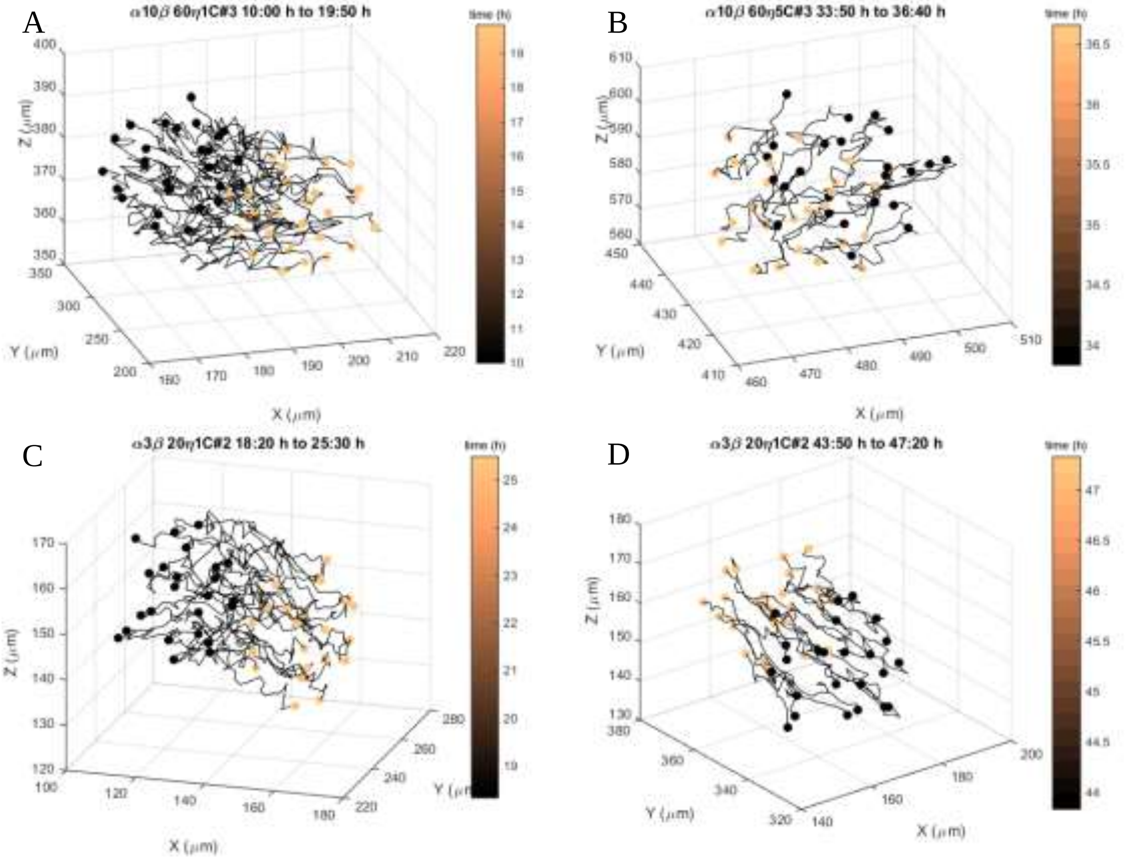
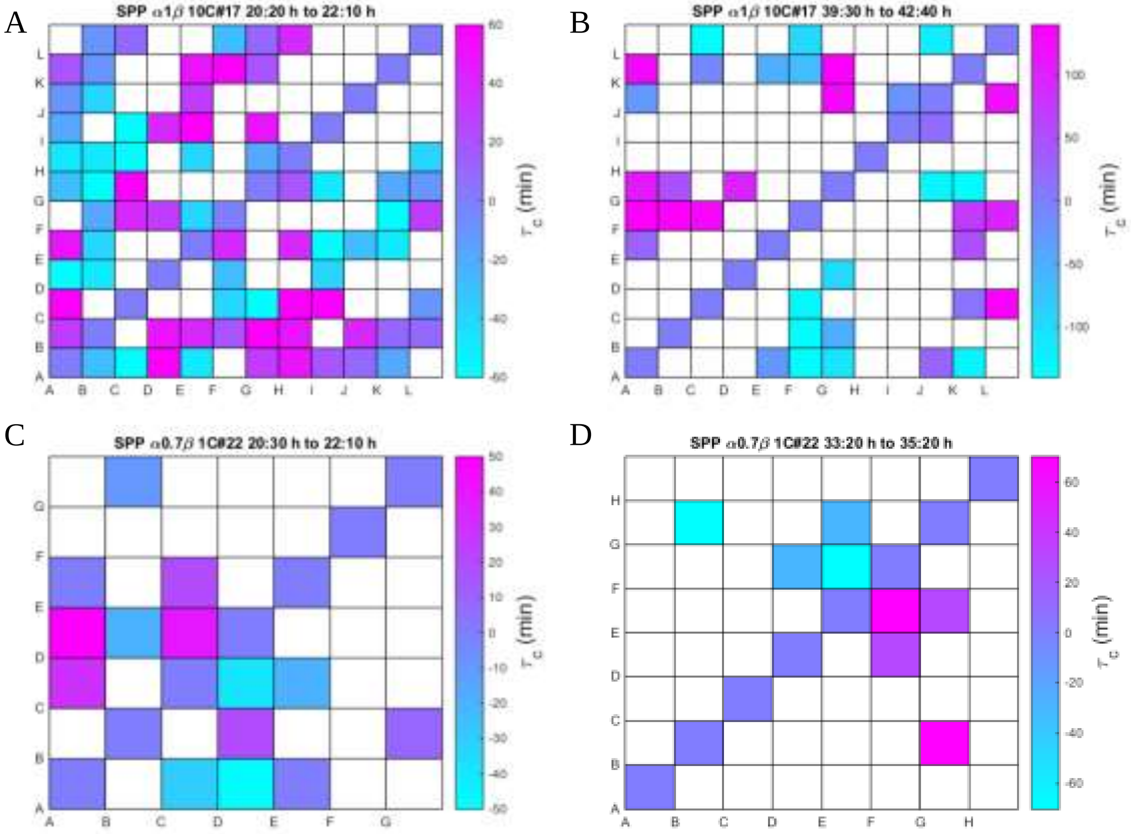


Figure 6-14: Cell Trajectories of Motility Events of variable speed SPP simulated cohorts

Spots mark nuclei connected by black lines denoting trajectories. Colors of circles represent time of beginning and end of trajectory. A. (α,β,η) values of (10,60,1) Cohort # 3 from 10:00 h to 19:50 h, B. (α,β,η) values of (10,60,1) Cohort # 3 from 33:50 h to 36:40 h C. (α,β,η) values of (3,20,1) Cohort # 2 from 18:20 h to 25:30 h and D. (α,β,η) values of (3,20,1) Cohort # 2 from 43:50 h to 47:20



<u>Figure 6-15: Correlation Values τ_c of simulated SPP cohorts within motility events</u>

Each map shows the correlation time lag τ_c of cell pairs during a given motility event. Cells are named A through Z, followed by lowercase a. Cells are always correlated with each other with a τ_c of 0 minutes. A. (α,β) values of (1,10) Cohort # 17 from 20:20 h to 22:10 h, B. (α,β) values of (1,10) Cohort # 17 from 39:30 h to 42:40 h C. (α,β) values of (0.7,1) Cohort # 22 from 20:30 h to 22:10 h and D. (α,β) values of (0.7,1) Cohort # 22 from 33:20 h to 35:20 h

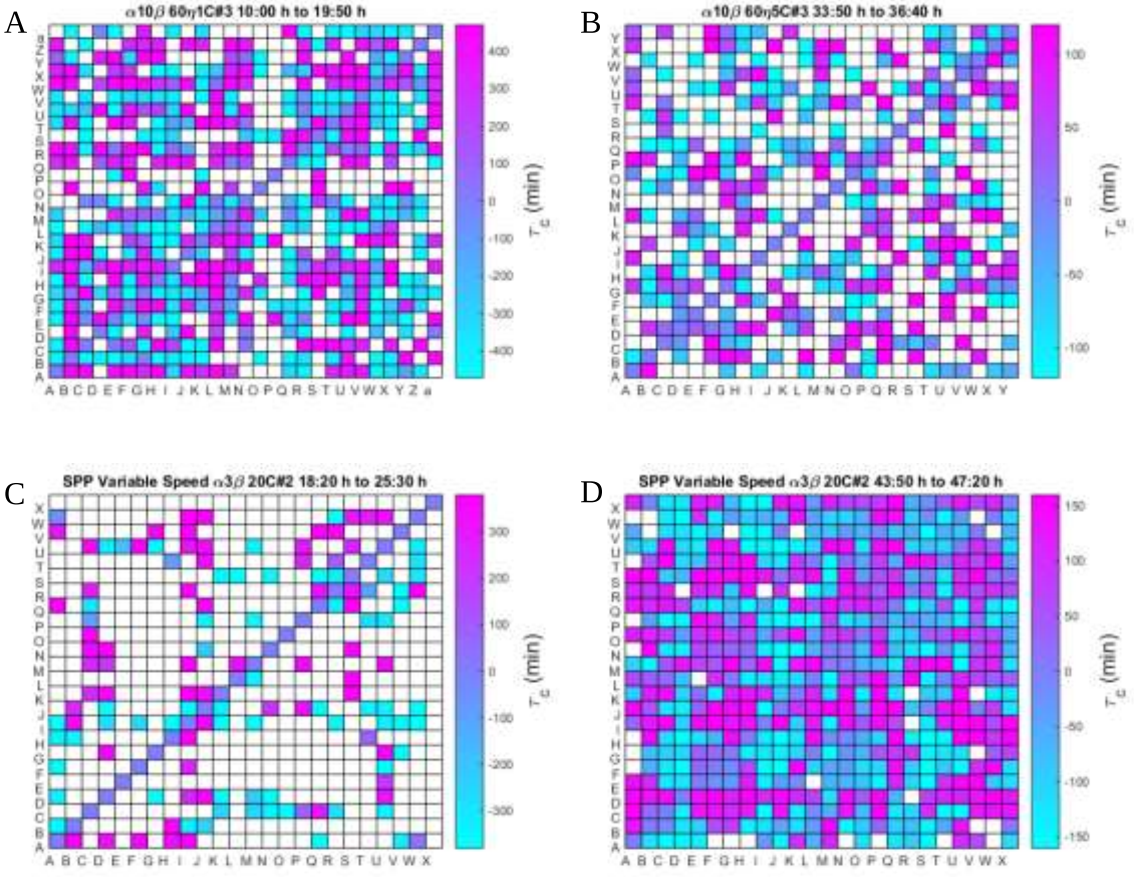
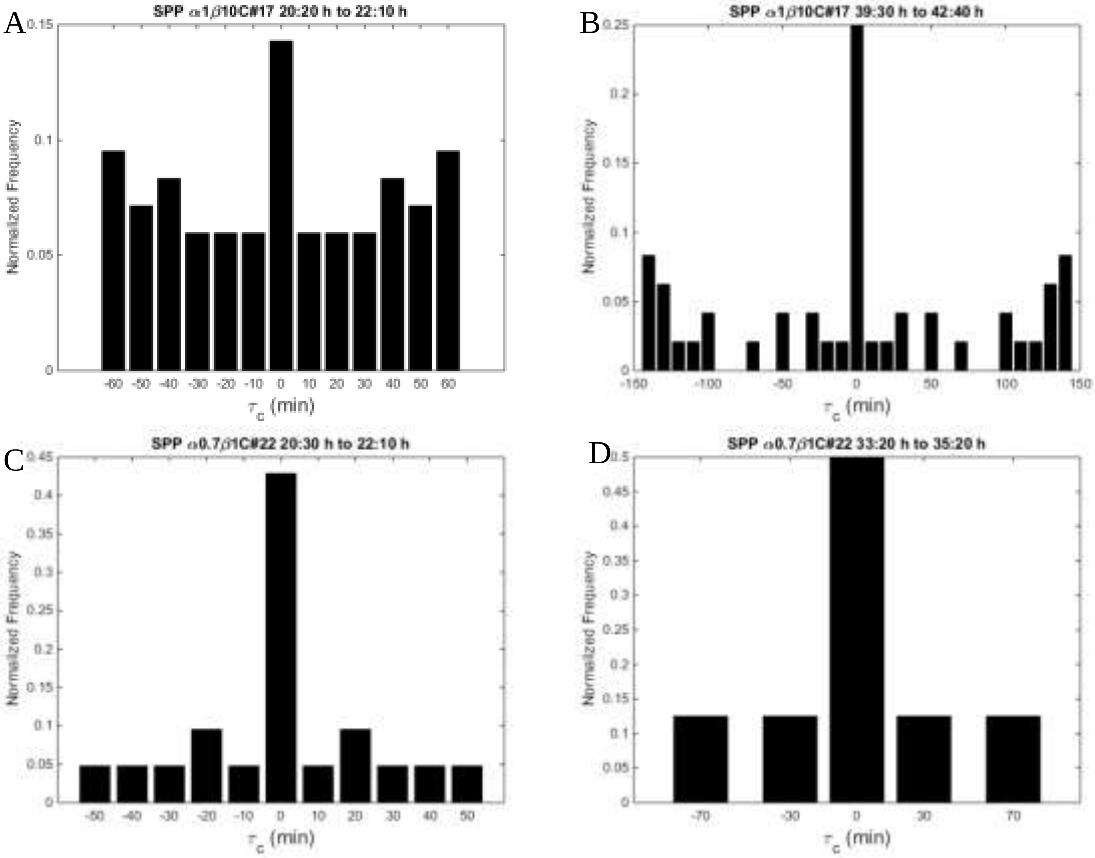


Figure 6-16: Correlation Values τ_c of simulated variable speed SPP cohorts within motility events

Each map shows the correlation time lag τ_c of cell pairs during a given motility event. Cells are named A through Z, followed by lowercase a. Cells are always correlated with each other with a τ_c of 0 minutes. A. (α,β,η) values of (10,60,1) Cohort # 3 from 10:00 h to 19:50 h, B. (α,β,η) values of (10,60,1) Cohort # 3 from 33:50 h to 36:40 h C. (α,β,η) values of (3,20,1) Cohort # 2 from 18:20 h to 25:30 h and D. (α,β,η) values of (3,20,1) Cohort # 2 from 43:50 h to 47:20



<u>Figure 6-17: Percentage of τ_c values for SPP motility events</u>

Each graph shows the percentage of the correlation time lag τ_c of cell pairs during a given motility event. A. (α,β) values of (1,10) Cohort # 17 from 20:20 h to 22:10 h, B. (α,β) values of (1,10) Cohort # 17 from 39:30 h to 42:40 h C. (α,β) values of (0.7,1) Cohort # 22 from 20:30 h to 22:10 h and D. (α,β) values of (0.7,1) Cohort # 22 from 33:20 h to 35:20

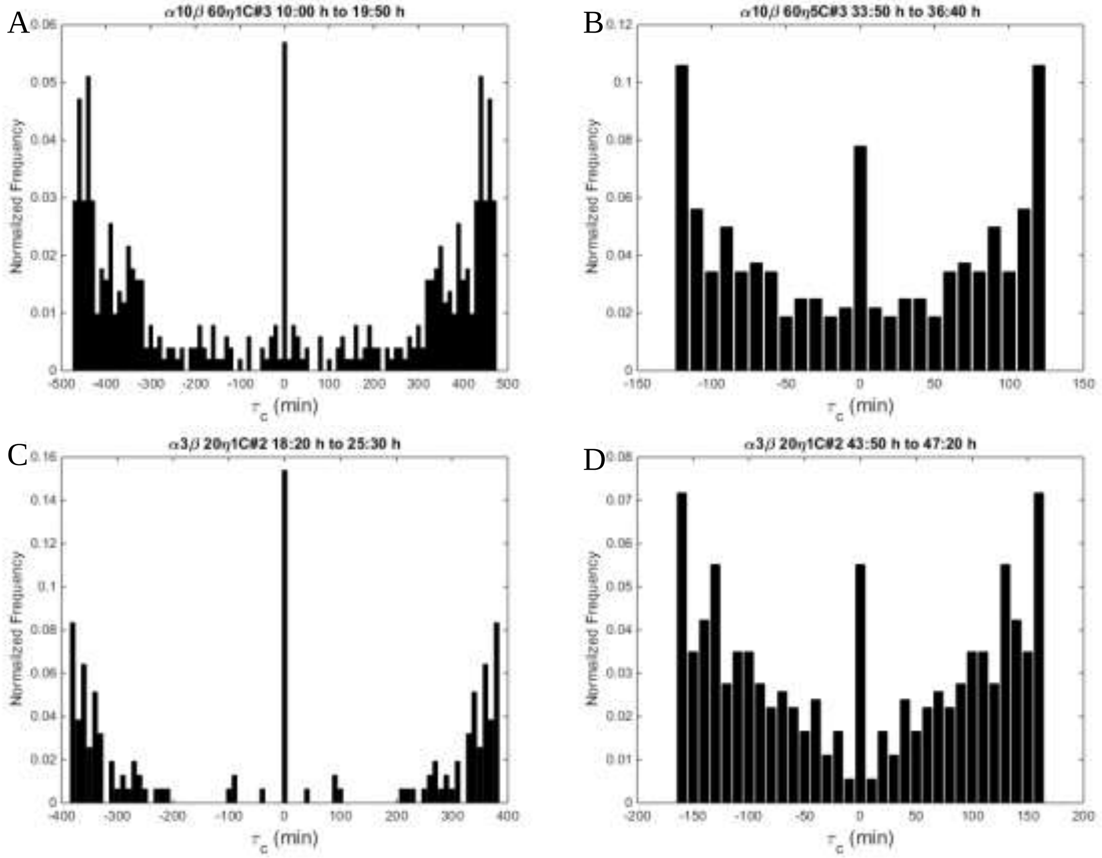


Figure 6-18: Percentage of τ_c values for variable speed SPP motility events

Each graph shows the percentage of the correlation time lag τ_c of cell pairs during a given motility event. A. (α,β,η) values of (10,60,1) Cohort # 3 from 10:00 h to 19:50 h, B. (α,β,η) values of (10,60,1) Cohort # 3 from 33:50 h to 36:40 h C. (α,β,η) values of (3,20,1) Cohort # 2 from 18:20 h to 25:30 h and D. (α,β,η) values of (3,20,1) Cohort # 2 from 43:50 h to 47:20

CHAPTER 7. CONCLUSIONS AND FUTURE DIRECTIONS

7.1. Summary of Specific Aims

The overarching goal of this work has been to observe, analyze and predict the motility of adhesive cellular systems in 3D, as well as to establish standard techniques for their analyses. In Aim 1 we establish an experimental model that is able to identify distinct cellular cohorts and cells after embedding in a 3D collagen gel (Chapter 4). In addition to tracking individual cells within a cohort, we develop an algorithm to assign cellular data to clusters, and are able to track clusters despite change in cell number due to proliferation, cohort merging or dissociation of individuals. This clustering algorithem is dependent on a single parameter, an interaction distance between two cells that defines whether they are in the same cluster. The experimental results provide the first observations of spontaneous motion of cell cohorts comprising more than 4 cells in 3D. In the absence of external driving agents, interactions between cohorts and a collagen matrix are sufficient to drive collective cell motility. Considering the balance of adherence and motility required for collective motion, translation over a few cell lengths in the absence of an external agent provides a promising model for the study of emergent phenomena and collective dynamics.

In Aim 2, we establish computational tools to recreate experimental data (Chapter 5). Self-Propelled Particle models provide a simple and effective tool to not only recreate experimental data, but provide insight into their dynamic properties. A balance of neighbor influence and intercellular forces between neighbors leads to clusters of similar sizes to experiments. Increasing the influence of neighboring cells leads to the creation of large unrealistic worms, and increasing intercellular forces leads to clusters with high fluctuations. We multi-scale the model by accounting scaling Adherens Junctions (AJ) with cellular speed. Cells with higher AJs have lower speed than individual cells, thus we create a variable speed SPP model. We use three criteria of cellular motion every hour, average cluster size, and order parameters to identify simulated conditions that potentially have similar dynamics to experimental data.

The nature of these heterogeneous dynamics led to devising a standard method for studying collectives on a cohort-by-cohort basis, as well as identifying motility events for them (Chapter 6). Our analytical algorithm successfully treates cell collectives as both spatially and temporally heterogeneous, and identifies motility events within individual cellular cohorts. The analytical techniques developed show, for the first time, that cells spontaneously rotate in short bursts and translate for several hours. The short bursts of rotation and comparatively larger spans of translation suggest that an internal stimulus arises within the dynamic cell-matrix system that attempts to drive collective translation. The data suggests that collective cell motility *in vivo* may arise under a combination of ideal conditions, and these analyses lay the foundation for identifying the biochemical mechanisms and conditions that could lead to collective motility. Motility events from simulated cohorts have similar signatures to experimental data in terms of correlation lags between rotating and translating cells, thus validating that the correlations between experimental cell pairs are not significantly affected by tracking and visualization signatures. Thus we are able to create a simulated cohort that has temporal and spatial heterogeneity.

The experimental, analytical, and computational techniques presented here are built to co-evolve. The experimental set-up can be modified to include specific matrix factors, and the result of such an experiment can be incorporated in the computational model. The predictions of the computational model can be tested and verified as well. Even the analytical tools can be modified for any cell type or matrix type, and can further evolve to include even more dynamic information while processing.

7.2. Future Directions

Our approaches can be expanded to study a range of phenomena in 3D, including collective cancer cell migration, density dependent phase transitions, cell jamming, and emergent systems. For instance, 3D Traction Force Microscopy can be conducted on experiments with the incorporation of fluorescent beads within the matrix as in Figure 4-8. This will yield the stresses exerted by cohorts on the ECM as they undergo translation and rotation. Testing experimental cohorts within degradable and non-degradable synthetic matrices, or in matrices with a gradient of degradability, will yield information on motile potential of cohorts and the dependence of motility on the ECM. Calcium chelators can reduce cell adhesivity and test the prediction of the computation model that reduced adhesivity will result in higher frequency of motility events. The computational model can also be multi-scaled further to incorporate differential equation based cellular networks of regulatory molecules in each cell. This holds tremendous potential to inspire questions of the biochemical factors that drive collective cell motility.

7.3. Conclusions

The results establish that cell collectives have diverse responses to similar environments, and that collective cell motility is a complex phenomenon influenced by many factors. However these results also establish that observations of motion alone can yield significant insight into motile potential of each cellular cohort, and classify a range of physical phenomenon. Thus this work provides a visualization of the plasticity of motion of cellular cohorts, and classification of their patterns. The motility classification methods provide anyone in the field of collective motility a simple analytical tool to classify it, and hto any cell type or matrix. In fact these tools can be expanded to anyone in the field of collective motility of any species, for so far most analytical techniques are dependent on uniform spatial density of units. This work also establishes that simple computational models are able to recreate such behavior and hold predictive potential of the biochemical factors driving it. Thus we provide reductionist approaches to understanding a complicated system, without necessarily reducing all of its complexity.

REFERENCES

Adams, C.L., 1998. Mechanisms of Epithelial Cell-Cell Adhesion and Cell Compaction Revealed by High-resolution Tracking of E-Cadherin-Green Fluorescent Protein. *The Journal of Cell Biology*, 142(4), pp.1105–1119.

Baker, B.M. & Chen, C.S., 2012. Deconstructing the third dimension: how 3D culture microenvironments alter cellular cues. *Journal of cell science*, 125(Pt 13), pp.3015–3024.

Baker, E.L. & Zaman, M.H., 2010. The biomechanical integrin. *Journal of biomechanics*, 43(1), pp.38–44.

Bazellières, E. et al., 2015. Control of cell-cell forces and collective cell dynamics by the intercellular adhesome. *Nature cell biology*, 17(4), pp.409–420.

Berrier, A.L. & Yamada, K.M., 2007. Cell-matrix adhesion. *Journal of cellular physiology*, 213(3), pp.565–573.

Bianco, A. et al., 2007. Two distinct modes of guidance signalling during collective migration of border cells. *Nature*, 448(7151), pp.362–365.

Bloom, R.J. et al., 2008. Mapping local matrix remodeling induced by a migrating tumor cell using three-dimensional multiple-particle tracking. *Biophysical journal*, 95(8), pp.4077–4088.

Brown, A., 1982. Neutrophil granulocytes: adhesion and locomotion on collagen substrata and in collagen matrices. *Journal of cell science*, 58, pp.455–467.

Brown, R.A. et al., 2002. Enhanced fibroblast contraction of 3D collagen lattices and integrin expression by TGF-beta1 and -beta3: mechanoregulatory growth factors? *Experimental cell research*, 274(2), pp.310–322.

Brugués, A. et al., 2014. Forces driving epithelial wound healing. *Nature Physics*, 10(9), pp.683–690.

Buhl, J. et al., 2006. From disorder to order in marching locusts. *Science*, 312(5778), pp.1402–1406.

Cai, D. et al., 2014. Mechanical feedback through E-cadherin promotes direction sensing during collective cell migration. *Cell*, 157(5), pp.1146–1159.

Carey, S.P. et al., 2013. Leading malignant cells initiate collective epithelial cell invasion in a three-dimensional heterotypic tumor spheroid model. *Clinical & experimental metastasis*, 30(5), pp.615–630.

Cetera, M. & Horne-Badovinac, S., 2015. Round and round gets you somewhere: collective cell migration and planar polarity in elongating Drosophila egg chambers. *Current opinion in genetics & development*, 32C, pp.10–15.

Charras, G. & Paluch, E., 2008. Blebs lead the way: how to migrate without lamellipodia. *Nature reviews Molecular cell biology*, 9(9), pp.730–736.

Chhabra, E.S. & Higgs, H.N., 2007. The many faces of actin: matching assembly factors with cellular structures. *Nature cell biology*, 9(10), pp.1110–1121.

Clow, P. & McNally, J.G., 1999. In vivo observations of myosin II dynamics support a role in rear retraction. *Molecular biology of the cell*, 10(5), pp.1309–1323.

Czirók, A. et al., 2013. Collective cell streams in epithelial monolayers depend on cell adhesion. *New journal of physics*, 15(7), p.075006.

Czirók, A. et al., 1996. Formation of complex bacterial colonies via self-generated vortices. *Physical Review E*, 54(2), pp.1791–1801.

Dalle Nogare, D. et al., 2014. Leading and trailing cells cooperate in collective migration of the zebrafish posterior lateral line primordium. *Development*, 141(16), pp.3188–3196.

Darnton, N.C. et al., 2010. Dynamics of bacterial swarming. *Biophysical journal*, 98(10), pp.2082–2090.

Deisboeck, T.S. & Couzin, I.D., 2009. Collective behavior in cancer cell populations. *BioEssays*, 31(2), pp.190–197.

Dickinson, R.B., 2000. A generalized transport model for biased cell migration in an anisotropic environment. *Journal of mathematical biology*, 40(2), pp.97–135.

Dickinson, R.B. & Tranquillo, R.T., 1993. Optimal estimation of cell movement indices from the statistical analysis of cell tracking data. *American Institute of Chemical Engineers journal*, 39(12), pp.1995–2010.

Domercq, M. et al., 2007. System xc- and glutamate transporter inhibition mediates microglial toxicity to oligodendrocytes. *Journal of immunology*, 178(10), pp.6549–6556.

Doxzen, K. et al., 2013. Guidance of collective cell migration by substrate geometry. *Integrative biology*, 5(8), pp.1026–1035.

Edelstein, A.D. et al., 2014. Advanced methods of microscope control using μManager software. *Journal of Biological Methods*, 1(2), p.10.

Evans, E. a & Calderwood, D. a, 2007. Forces and bond dynamics in cell adhesion. *Science*, 316(5828), pp.1148–1153.

Fallica, B. et al., 2012. Alteration of cellular behavior and response to PI3K pathway inhibition by culture in 3D collagen gels. *Public Library of Science one*, 7(10), p.e48024.

Fan, Q.-W. et al., 2006. A dual PI3 kinase/mTOR inhibitor reveals emergent efficacy in glioma. *Cancer cell*, 9(5), pp.341–349.

Feroze, R. et al., 2015. Mechanics of blastopore closure during amphibian gastrulation. *Developmental biology*, 398(1), pp.57–67.

Ferrara, N., Gerber, H.-P. & LeCouter, J., 2003. The biology of VEGF and its receptors. *Nature medicine*, 9(6), pp.669–676.

Fournier, M.F. et al., 2010. Force transmission in migrating cells. *The Journal of cell biology*, 188(2), pp.287–297.

Franck, C. et al., 2011. Three-dimensional traction force microscopy: a new tool for quantifying cell-matrix interactions. *Public Library of Science one*, 6(3), p.e17833.

Frascoli, F. et al., 2013. A computational model for collective cellular motion in three dimensions: general framework and case study for cell pair dynamics. *Public Library of Science one*, 8(3), p.e59249.

Frieboes, H.B. et al., 2006. An integrated computational/experimental model of tumor invasion. *Cancer research*, 66(3), pp.1597–1604.

Friedl, P. & Alexander, S., 2011. Cancer invasion and the microenvironment: Plasticity and reciprocity. *Cell*, 147(5), pp.992–1009.

Friedl, P. & Bröcker, E.B., 2000. The biology of cell locomotion within three-dimensional extracellular matrix. *Cellular and molecular life sciences : CMLS*, 57(1), pp.41–64.

Friedl, P. & Gilmour, D., 2009. Collective cell migration in morphogenesis, regeneration and cancer. *Nature reviews. Molecular cell biology*, 10(7), pp.445–457.

Friedl, P., Hegerfeldt, Y. & Tusch, M., 2004. Collective cell migration in morphogenesis and cancer. *International journal of developmental biology*, 48(5-6), pp.441–449.

Friedl, P. & Wolf, K., 2008. Tube travel: the role of proteases in individual and collective cancer cell invasion. *Cancer research*, 68(18), pp.7247–7249.

Fujikawa, H. & Matsushita, M., 1989. Fractal Growth of Bacillus subtilis on Agar Plates.

Journal of the Physics Society Japan, 58(11), pp.3875–3878.

Gaggioli, C., 2008. Collective invasion of carcinoma cells: when the fibroblasts take the lead. *Cell adhesion & migration*, 2(1), pp.45–47.

Gaggioli, C. et al., 2007. Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nature cell biology*, 9(12), pp.1392–1400.

Gao, Y. & Kilfoil, M.L., 2009. Accurate detection and complete tracking of large populations of features in three dimensions. *Optics express*, 17(6), pp.4685–4704.

Gloushankova, N., 2008. Changes in regulation of cell—cell adhesion during tumor transformation. *Biochemistry*, 73(7), pp.742–750.

Gov, N.S., 2007. Collective cell migration patterns: follow the leader. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41), pp.15970–15971.

Grégoire, G., Chaté, H. & Tu, Y., 2003. Moving and staying together without a leader. *Physica D: Nonlinear Phenomena*, 181(3-4), pp.157–170.

Griffith, L.G. & Swartz, M.A., 2006. Capturing complex 3D tissue physiology in vitro. *Nature reviews. Molecular cell biology*, 7(3), pp.211–224.

Guarnieri, D. et al., 2010. Covalently immobilized RGD gradient on PEG hydrogel scaffold influences cell migration parameters. *Acta biomaterialia*, 6(7), pp.2532–2539.

Guck, J. et al., 2010. Critical review: cellular mechanobiology and amoeboid migration. *Integrative biology*, 2(11-12), pp.575–583.

Gunzer, M. et al., 2000. Migration of dendritic cells within 3-D collagen lattices is dependent on tissue origin, state of maturation, and matrix structure and is maintained by proinflammatory cytokines. *Journal of leukocyte biology*, 67(5), pp.622–629.

Guo, M. et al., 2014. Probing the stochastic, motor-driven properties of the cytoplasm using force spectrum microscopy. *Cell*, 158(4), pp.822–832.

Haeger, A. et al., 2014. Cell jamming: collective invasion of mesenchymal tumor cells imposed by tissue confinement. *Biochimica et biophysica acta*, 1840(8), pp.2386–2395.

Haeger, A. et al., 2015. Collective cell migration: guidance principles and hierarchies. *Trends in cell biology*, 25(9), pp.556–566.

Haga, H. et al., 2005. Collective movement of epithelial cells on a collagen gel substrate. *Biophysical journal*, 88(3), pp.2250–2256.

Haigo, S.L. & Bilder, D., 2011. Global tissue revolutions in a morphogenetic movement controlling elongation. *Science*, 331(6020), pp.1071–1074.

Harjanto, D., Maffei, J.S. & Zaman, M.H., 2011. Quantitative analysis of the effect of cancer invasiveness and collagen concentration on 3D matrix remodeling. *Public Library of Science one*, 6(9), p.e24891.

Harjanto, D. & Zaman, M.H., 2010. Computational study of proteolysis-driven single cell migration in a three-dimensional matrix. *Annals of biomedical engineering*, 38(5), pp.1815–1825.

Hayakawa, Y. & Furuhashi, S., 2012. Group-size distribution of skeins of wild geese. *Physical Review E*, 86(3), p.031924.

Hegerfeldt, Y. et al., 2002. Collective Cell Movement in Primary Melanoma Explants : Plasticity of Cell-Cell. *Cancer research*, 62, pp.2125–2130.

Hooper, S., Marshall, J.F. & Sahai, E., 2006. Tumor cell migration in three dimensions. *Methods in enzymology*, 406(1999), pp.625–643.

Humphries, J.D., Byron, A. & Humphries, M.J., 2006. Integrin ligands at a glance. *Journal of cell science*, 119(Pt 19), pp.3901–3903.

Hynes, R.O., 2009. The extracellular matrix: not just pretty fibrils. *Science*, 326(5957), pp.1216–1219.

Ishizaki, T. et al., 2000. Pharmacological properties of Y-27632, a specific inhibitor of rho-associated kinases. *Molecular pharmacology*, 57(5), pp.976–983.

Kabla, A.J., 2012. Collective cell migration: leadership, invasion and segregation. *Journal of the Royal Society, Interface / the Royal Society*, 9(77), pp.3268–3278.

Kenawy, E.R. et al., 2003. Electrospinning of poly(ethylene-co-vinyl alcohol) fibers. *Biomaterials*, 24(6), pp.907–913.

Khalil, A. & Friedl, P., 2010. Determinants of leader cells in collective cell migration. *Integrative biology*, 2(11-12), pp.568–574.

Kim, A., Lakshman, N. & Petroll, W.M., 2006. Quantitative assessment of local collagen matrix remodeling in 3-D culture: the role of Rho kinase. *Experimental cell research*, 312(18), pp.3683–3692.

Kim, T., Hwang, W. & Kamm, R.D., 2007. Computational Analysis of a Cross-linked

Actin-like Network. *Experimental Mechanics*, 49(1), pp.91–104.

King, S.J. & Parsons, M., 2011. Imaging cells within 3D cell-derived matrix. *Methods in molecular biology*, 769, pp.53–64.

Kirfel, G. et al., 2004. Cell migration: mechanisms of rear detachment and the formation of migration tracks. *European journal of cell biology*, 83(11-12), pp.717–724.

Kraning-Rush, C.M. et al., 2011. The role of the cytoskeleton in cellular force generation in 2D and 3D environments. *Physical biology*, 8(1), p.015009.

Kriebel, P.W. et al., 2008. Collective cell migration requires vesicular trafficking for chemoattractant delivery at the trailing edge. *The Journal of cell biology*, 183(5), pp.949–961.

Lämmermann, T. et al., 2008. Rapid leukocyte migration by integrin-independent flowing and squeezing. *Nature*, 453(7191), pp.51–55.

Lee, P. & Wolgemuth, C.W., 2011. Crawling cells can close wounds without purse strings or signaling. *Public Library of Science computational biology*, 7(3), p.e1002007.

Leong, W.S. et al., 2010. Thickness sensing of hMSCs on collagen gel directs stem cell fate. *Biochemical and biophysical research communications*, 401(2), pp.287–292.

Lundquist, E., 2009. The finer points of filopodia. *Public Library of Science biology*, 7(6), p.e1000142.

Luo, Y. & Shoichet, M.S., 2004. A photolabile hydrogel for guided three-dimensional cell growth and migration. *Nature materials*, 3(4), pp.249–253.

Macpherson, I.R. et al., 2007. p120-catenin is required for the collective invasion of squamous cell carcinoma cells via a phosphorylation-independent mechanism. *Oncogene*, 26(36), pp.5214–5228.

Mak, M., Kamm, R.D. & Zaman, M.H., 2014. Impact of dimensionality and network disruption on microrheology of cancer cells in 3D environments. *Public Library of Science computational biology*, 10(11), p.e1003959.

Mason, T.G. et al., 1997. Particle Tracking Microrheology of Complex Fluids. *Physical Review Letters*, 79(17), pp.3282–3285.

Méhes, E. & Vicsek, T., 2014. Collective motion of cells: from experiments to models. *Integrative biology*, 6(9), pp.831–854.

Meschede, M. & Hallatschek, O., 2013. Local disalignment can promote coherent collective motion. *New Journal of Physics*, 15(4), p.045027.

Mierke, C.T., 2013. Physical break-down of the classical view on cancer cell invasion and metastasis. *European journal of cell biology*, 92(3), pp.89–104.

Mogilner, A. & Oster, G., 1996. Cell motility driven by actin polymerization. *Biophysical journal*, 71(6), pp.3030–3045.

Nagy, M. et al., 2010. Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290), pp.890–893.

Narumiya, S. & Watanabe, N., 2009. Migration without a clutch. *Nature cell biology*, 11(12), pp.1394–1396.

Ott, H.C. et al., 2008. Perfusion-decellularized matrix: using nature's platform to engineer a bioartificial heart. *Nature medicine*, 14(2), pp.213–221.

Palecek, S.P. et al., 1996. Integrin dynamics on the tail region of migrating fibroblasts. *Journal of cell science*, 109 (Pt 5, pp.941–952.

Palecek, S.P. et al., 1998. Physical and biochemical regulation of integrin release during rear detachment of migrating cells. *Journal of cell science*, 111 (Pt 7, pp.929–940.

Panorchan, P. et al., 2006. Microrheology and ROCK signaling of human endothelial cells embedded in a 3D matrix. *Biophysical journal*, 91(9), pp.3499–3507.

Park, J.-A. et al., 2015. Unjamming and cell shape in the asthmatic airway epithelium. *Nature materials*, 14(10), pp.1040–1048.

Parkhurst, M.R. & Saltzman, W.M., 1992. Quantification of human neutrophil motility in three-dimensional collagen gels. Effect of collagen concentration. *Biophysical journal*, 61(2), pp.306–315.

Pedersen, J.A. & Swartz, M.A., 2005. Mechanobiology in the third dimension. *Annals of biomedical engineering*, 33(11), pp.1469–1490.

Pelletier, V. et al., 2009. Microrheology of Microtubule Solutions and Actin-Microtubule Composite Networks. *Physical Review Letters*, 102(18), p.188303.

Pietras, K. & Ostman, A., 2010. Hallmarks of cancer: interactions with the tumor stroma. *Experimental cell research*, 316(8), pp.1324–1331.

Poujade, M. et al., 2007. Collective migration of an epithelial monolayer in response to a model wound. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41), pp.15988–15993.

Provenzano, P.P. et al., 2008. Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization. *Biophysical journal*, 95(11), pp.5374–5384.

Provenzano, P.P. et al., 2009. Matrix density-induced mechanoregulation of breast cell phenotype, signaling and gene expression through a FAK-ERK linkage. *Oncogene*, 28(49), pp.4326–4343.

Raeber, G.P., Lutolf, M.P. & Hubbell, J.A., 2007. Mechanisms of 3-D migration and matrix remodeling of fibroblasts within artificial ECMs. *Acta biomaterialia*, 3(5), pp.615–629.

Raeber, G.P., Lutolf, M.P. & Hubbell, J.A., 2005. Molecularly engineered PEG hydrogels: a novel model system for proteolytically mediated cell migration. *Biophysical journal*, 89(2), pp.1374–1388.

Rapanan, J.L. et al., 2014. Collective cell migration of primary zebrafish keratocytes. *Experimental cell research*, 326(1), pp.155–165.

Renkawitz, J. et al., 2009. Adaptive force transmission in amoeboid cell migration. *Nature cell biology*, 11(12), pp.1438–1443.

Rørth, P., 2009a. Collective cell migration. *Annual review of cell and developmental biology*, 25, pp.407–29.

Rørth, P., 2009b. Collective cell migration. *Annual review of cell and developmental biology*, 25, pp.407–429.

du Roure, O. et al., 2005. Force mapping in epithelial cell migration. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), pp.2390–2395.

Roy, P. et al., 1997. An in vitro force measurement assay to study the early mechanical interaction between corneal fibroblasts and collagen matrix. *Experimental cell research*, 232(1), pp.106–117.

Rubenstein, B.M. & Kaufman, L.J., 2008. The role of extracellular matrix in glioma invasion: a cellular Potts model approach. *Biophysical journal*, 95(12), pp.5661–5680.

Sabeh, F., Shimizu-Hirota, R. & Weiss, S.J., 2009. Protease-dependent versus -independent cancer cell invasion programs: three-dimensional amoeboid movement revisited. *The Journal of cell biology*, 185(1), pp.11–19.

Schmeichel, K.L. & Bissell, M.J., 2003. Modeling tissue-specific signaling and organ function in three dimensions. *Journal of cell science*, 116(Pt 12), pp.2377–2388.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), pp.671–675.

Schwartz, M.P. et al., 2013. A quantitative comparison of human HT-1080 fibrosarcoma cells and primary human dermal fibroblasts identifies a 3D migration mechanism with properties unique to the transformed phenotype. *Public Library of Science one*, 8(12), p.e81689.

Schwartz, M.P. et al., 2010. A synthetic strategy for mimicking the extracellular matrix provides new insight about tumor cell migration. *Integrative biology : quantitative biosciences from nano to macro*, 2(1), pp.32–40.

Sepúlveda, N. et al., 2013. Collective Cell Motion in an Epithelial Sheet Can Be Quantitatively Described by a Stochastic Interacting Particle Model M. P. Brenner, ed. *PLoS Computational Biology*, 9(3), p.e1002944.

Shao, D., Levine, H. & Rappel, W.-J., 2012. Coupling actin flow, adhesion, and morphology in a computational cell motility model. *Proceedings of the National Academy of Sciences of the United States of America*, 109(18), pp.6851–6856.

Sharma, Y. et al., 2015. Collective motion of mammalian cell cohorts in 3D. *Integrative Biology*, 7(12), pp.1526–1533.

Sharma, Y. & Zaman, M.H., 2014. Cell Motility in 3D Matrices. In R. Kaunas & A. Zemel, eds. *Cell and Matrix Mechanics*. Boca Raton, FL: CRC Press, pp. 197–218.

Shaw, T.J. & Martin, P., 2009. Wound repair at a glance. *Journal of cell science*, 122(Pt 18), pp.3209–3213.

Slaughter, B. V et al., 2009. Hydrogels in regenerative medicine. *Advanced materials*, 21(32-33), pp.3307–3329.

Sontheimer, H., 2003. Malignant gliomas: perverting glutamate and ion homeostasis for selective advantage. *Trends in Neurosciences*, 26(10), pp.543–549.

Starke, J. et al., 2013. Mechanotransduction of mesenchymal melanoma cell invasion into 3D collagen lattices: filopod-mediated extension-relaxation cycles and force anisotropy. *Experimental cell research*, 319(16), pp.2424–2433.

Szabó, A. et al., 2010. Collective cell motion in endothelial monolayers. *Physical biology*, 7(4), p.046007.

Szabó, A. & Merks, R.M., 2013. Cellular potts modeling of tumor growth, tumor invasion, and tumor evolution. *Frontiers in oncology*, 3, p.87.

Szabó, B. et al., 2006. Phase transition in the collective migration of tissue cells: Experiment and model. *Physical Review E*, 74(6), pp.1–5.

Szabó, P., Nagy, M. & Vicsek, T., 2009. Transitions in a self-propelled-particles model with coupling of accelerations. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 79(2 Pt 1), p.021908.

Takano, T. et al., 2001. Glutamate release promotes growth of malignant gliomas. *Nature medicine*, 7(9), pp.1010–1015.

Tambe, D.T. et al., 2011. Collective cell guidance by cooperative intercellular forces. *Nature materials*, 10(6), pp.469–475.

Tanner, K. et al., 2012. Coherent angular motion in the establishment of multicellular architecture of glandular tissues. *Proceedings of the National Academy of Sciences of the United States of America*, 109(6), pp.1973–1978.

Theveneau, E. & Mayor, R., 2012. Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Developmental biology*, 366(1), pp.34–54.

Thiery, J.P., 2002. Epithelial-mesenchymal transitions in tumour progression. *Nature reviews. Cancer*, 2(6), pp.442–454.

Thiery, J.P., 2009. Metastasis: Alone or Together? *Current Biology*, 19(24), pp.R1121–R1123.

Thiery, J.P. & Lim, C.T., 2013. Tumor Dissemination: An EMT Affair. *Cancer Cell*, 23(3), pp.272–273.

Tkačik, G. et al., 2014. Searching for collective behavior in a large network of sensory neurons. *Public Library of Science computational biology*, 10(1), p.e1003408.

Tunstrøm, K. et al., 2013. Collective states, multistability and transitional behavior in schooling fish. *Public Library of Science computational biology*, 9(2), p.e1002915.

Vargas, D.A. et al., 2011. Computational model for migration of a cell cluster in three-dimensional matrices. *Annals of biomedical engineering*, 39(7), pp.2068–2079.

Vargas, D.A. et al., 2015. Incorporation of N-glycoslylation pathway in numerical model of Wnt/β-catenin reveals its importance in E-cadherin mediated intercellular adhesion. *In Preparation*.

Vicsek, T. et al., 1995. Novel Type of Phase Transition in a System of Self-Driven Particles. *Physical Review Letters*, 75(6), pp.1226–1229.

Vicsek, T. & Zafeiris, A., 2012. Collective motion. *Physics Reports*, 517(3-4), pp.71–140.

Vignjevic, D. et al., 2007. Fascin, a novel target of beta-catenin-TCF signaling, is expressed at the invasive front of human colon cancer. *Cancer research*, 67(14), pp.6844–6853.

Vignjevic, D. et al., 2006. Role of fascin in filopodial protrusion. *The Journal of cell biology*, 174(6), pp.863–875.

Vogler, G. & Bodmer, R., 2015. Cellular Mechanisms of Drosophila Heart Morphogenesis. *Journal of Cardiovascular Development and Disease*, 2(1), pp.2–16.

Weijer, C.J., 2009. Collective cell migration in development. *Journal of cell science*, 122(Pt 18), pp.3215–3223.

Wicki, A. et al., 2006. Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer cell*, 9(4), pp.261–272.

Wirtz, D., 2009. Particle-tracking microrheology of living cells: principles and applications. *Annual review of biophysics*, 38, pp.301–326.

Wolf, K. et al., 2003. Compensation mechanism in tumor cell migration: mesenchymal-amoeboid transition after blocking of pericellular proteolysis. *The Journal of cell biology*, 160(2), pp.267–277.

Wolf, K. & Friedl, P., 2011. Extracellular matrix determinants of proteolytic and non-proteolytic cell migration. *Trends in cell biology*, 21(12), pp.736–744.

Wozniak, M.A. & Keely, P.J., 2005. Use of three-dimensional collagen gels to study mechanotransduction in T47D breast epithelial cells. *Biological procedures online*, 7(1), pp.144–161.

Wu, Y. et al., 2009. Periodic reversal of direction allows Myxobacteria to swarm. *Proceedings of the National Academy of Sciences of the United States of America*, 106(4), pp.1222–1227.

Yamaguchi, N. et al., 2015. Leader cells regulate collective cell migration via Rac activation in the downstream signaling of integrin β1 and PI3K. *Scientific reports*, 5, p.7656.

Zaman, M.H., 2007. A multiscale probabilisitic framework to model early steps in tumor metastasis. *Molecular & cellular biomechanics*, 4(3), pp.133–141.

Zaman, M.H. et al., 2005. Computational model for cell migration in three-dimensional matrices. *Biophysical journal*, 89(2), pp.1389–1397.

Zaman, M.H. et al., 2006. Migration of tumor cells in 3D matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 103(29), pp.10889–10894.

Zhang, H.P. et al., 2010. Collective motion and density fluctuations in bacterial colonies. *Proceedings of the National Academy of Sciences of the United States of America*, 107(31), pp.13626–13630.

VITA

[redacted]

[redacted]

[redacted]