Rational design to control multipotent stromal cell migration for applications in bone tissue engineering and injury repair

by

Shan Wu

B.S., Computer Engineering and Electrical Engineering
University of Virginia, 2004

Submitted to the Department of Biological Engineering in
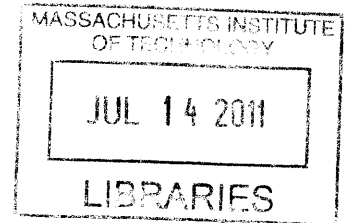Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN BIOLOGICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

© 2011 Massachusetts Institute of Technology. All rights reserved.

Signature of Author: _____
Department of Biological Engineering
May 20, 2011

Certified by: _____
Douglas A. Lauffenburger
Ford Professor of Biological Engineering, Biology, and Chemical Engineering
Thesis Supervisor

Certified by: _____
Forest M. White
Associate Professor of Biological Engineering
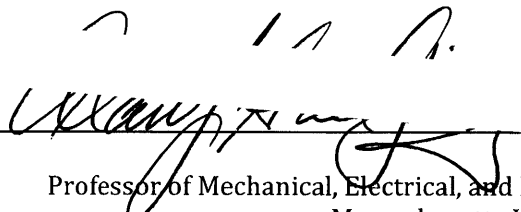Biological Engineering Graduate Program Committee Co-Chairman

Thesis Committee

Douglas A. Lauffenburger
Ford Professor of Biological Engineering, Biology, and Chemical Engineering
Massachusetts Institute of Technology
Thesis Supervisor

Linda G. Griffith
SETI Professor of Biological and Mechanical Engineering
Massachusetts Institute of Technology

Alan Wells
Thomas J. Gill III Professor of Pathology
University of Pittsburgh

Approved by: _____

Alan J. Grodzinsky
Professor of Mechanical, Electrical, and Biological Engineering
Massachusetts Institute of Technology
Thesis Committee Chairman

3

Rational design to control multipotent stromal cell migration for applications in bone tissue engineering and injury repair

by

Shan Wu

Submitted to the Department of Biological Engineering on May 20, 2011 in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy in
Biological Engineering

Abstract

Multipotent stromal cells derived from bone marrow hold great potential for tissue engineering applications because of their ability to home to injury sites and to differentiate along mesodermal lineages to become osteocytes, chondrocytes, and adipocytes to aid in tissue repair and regeneration. One key challenge, however, is the scarcity of MSC numbers isolated from *in vivo*, suggesting a role for biomimetic scaffolds in the cells' *ex vivo* expansion before reintegration into target tissue. Toward this end, immobilized epidermal growth factor (tEGF) has recently been found to promote MSC survival and proliferation and is a prime candidate to be incorporated into scaffolds to control MSC behavior. To rationally and effectively design scaffolds to drive MSC responses of survival, proliferation, migration, and differentiation, we must first understand these responses and the underlying protein signaling pathways that mediate them. While our knowledge of MSC behavior is limited as a field, MSC migration is particularly less studied despite being critical for tissue and scaffold infiltration. In this thesis, we quantitatively investigate the effects of tEGF and extracellular matrix (ECM) on MSC migration response and signaling. We take a systems level computational view to show a combined biomaterials and small molecule approach to control MSC migration.

Cell migration is a delicately integrated biophysical process involving polarization and protrusions at the cell front, adhesion and translocation of the cell body through contractile forces, followed by disassembly of adhesion complexes at the cell rear to allow detachment and productive motility. This process is mediated by a multitude of crosstalking signaling pathways downstream of integrin and growth factor activation. Using a poly(methyl methacrylate)-grafted-poly(ethylene oxide) (PMMA-g-PEO) copolymer base, we modify the PEO sidechains with immobilized epidermal growth factor (tEGF) as a model system for biomimetic scaffolds. We systematically adsorb fibronectin, vitronectin, and collagen ECM proteins to alter surface adhesiveness and measure MSC migration responses of speed and directional persistence alongside intracellular activities of EGFR, ERK, Akt,

and FAK phosphoproteins. While tEGF and ECM proteins differentially affected signaling and migration, univariate correlations between signals and responses were not informative, prompting the need for multivariate modeling to identify key patterns. Using decision tree "signal-response" modeling, we predicted that inhibiting ERK on collagen-adsorbed tEGF polymer surfaces would increase cell mean free path (MFP) by increasing directional persistence.

We confirmed this experimentally, successfully demonstrating a two-layer approach—"coarse" biomaterials followed by small molecules "fine-tuning"—to precisely and differentially control MSC migration speed and persistence, setting the stage for combination therapies for bone tissue engineering.

Thesis Supervisor: Douglas A. Lauffenburger
Title: Ford Professor of Biological Engineering, Biology, and Chemical Engineering

to mom & dad, with all the love in the world

Acknowledgements

I saved this part to write last because I thought it would be a nice little treat at the end of everything else. Turns out, this part is really tough. So many people have walked with me these last seven years and influenced my development as a student, scholar, scientist, and citizen. No words would ever seem worthy of capturing the magnitude of my gratitude for having had all of them by my side.

First and foremost, this thesis would not have been possible without the undying support and encouragement of my advisor, Doug Lauffenburger. My path was not the typical path, sometimes to the detriment of my own productivity, but Doug always emphasized and even prioritized the importance of forging my own exploration of science and of a career. Doug has been absolutely instrumental in my development as an analytical thinker, always managing to provide that tiny missing piece of connection between ideas that makes all the difference. Thank you Doug, for all of your trust in me.

I am forever indebted to my thesis committee members who all helped so much along the way. I could not have gotten this far without Al Grodzinsky's enthusiastic and continuous support. Linda Griffith was an endless resource for all things biomaterials and tissue engineering related, not to mention a great mentor and role model for science. Alan Wells provided critical insights during times of the most need on the signaling components of this thesis, without which this thesis would have had a hard time progressing.

My various colleagues in the Lauffenburger and Griffith labs have given countless hours to aid me in my research endeavors. Lisa Joslin and HD Kim helped me set up and get started as a new lab member on the migration team. Aran Parillo was always available to fix/replace/ reinstall computer parts and software to ensure my work could go on with minimal delays. Shannon Alford's forever willingness to share protocols, guide me through techniques, and discuss problems makes us all strive to be as good of a mentor to others as she has been to us. Without Linda Stockdale's meticulous and relentless pursuit to standardize and understand all of our polymer variants, this thesis would not have been possible. And most of all, I am so lucky to be able to call all of these wonderful colleagues my friends.

The Lauffenburger lab has had a phenomenal dynasty of lab managers who work behind the scenes to enable all of our research. From my early days working with Christina Lewis to Stacey Pawson's thoroughness to Hsinhwa Lee's warmth and heart, they have made a tremendous impact on us all.

And now comes the fun parts.

My friends make me smile every day. It is those fleeting moments of shared laughter that help us to look back on our meandering paths through grad school with the utmost joy, fondness, and nostalgia. I am thankful to have met all of you. Hong Ma and Kaity Ryan, thank you for being some of the best people I have ever met and for introducing me to the great outdoors of New England. My BE classmates made my first year survivable, especially Kristen Naegle, Megan Palmer, Hector Mobine, and Phil Dextras. Ed Lee kept us all together having fun from the first day of

orientation on that late August morning many years ago. My fellow GSCers - Eric Weese, Obrad Scepanovic, Jehanzeb Noor - thanks for all the humor. Jed Wartman, you have been an immense inspiration to me and to countless other students. Awesom-o, I am so glad I found you guys this past year to keep me sane and hydrated. Danilo and Tina Scepanovic you have both enriched my life so much with your spirit and joy throughout the years. Nan Gu, I will always pause what I am doing and grab a cup of coffee with you. New House 4, thanks for providing so many laughs, late nights, and Thursday night study breaks.

Victor, thank you for being such a guiding force through some of my toughest times in grad school and in life. You inspired me to love science as much as you do and showed me the passion and determination that is possible if we only just put our minds to it. You challenged me to aim for my highest potential, to achieve not for anyone else but for my own goals, aspirations, and abilities. Most of all, you taught me how much integrity, care, and beauty there is in our world and helped me to become a better citizen of humanity.

Last but not least, my family has always played a tremendous part in my lifelong career in learning. From my grandfather who invested so much hope and love in a little girl, to my dad who inspired my fascination with chemistry and with science before I even knew what science was, to my mom who would sacrifice anything to provide me with a better life. Mom and dad, we have come such a long way since our Saturday morning grocery walks to Food Lion, but those will always remain some of my favorite memories. Thanks for teaching me the importance of meeting challenges head on and for providing me with the best life growing up that you knew how. Words simply cannot express my love and my gratitude. This thesis is for you, with all the love in the world.

Table of Contents

1. Chapter 1 – Introduction

1.1. Multipotent stromal cells for bone tissue engineering

Bone-marrow derived multi-potent stromal cells (MSCs), or sometimes referred to as

mesenchymal stem cells, are nonhematopoietic stem cells that can give rise to multiple lineages

(Figure 1.1) [1,2]. These cells were first discovered in 1968 by Friedenstein et al. as fibroblast-like

cells from bone marrow aspirates that can adhere to tissue culture plastic [3]. Friedenstein defined

them as colony-forming unit fibroblasts (CFU-F), and showed that they were able to differentiate

into cells of the mesodermal lineage such as adipocytes, chondrocytes, and osteocytes. Since their

discovery more than 40 years ago however, there are no standard definitive cell surface markers to

isolate these cells from humans [4]. As such, reports of MSCs in literature can be muddled and

conflicting. Depending on the exact source of MSC isolation, ranging from bone marrow [1], to

skeletal muscle [5] to umbilical cord [6] to adipose tissue [7] to the circulatory system [8], the cells

are not necessarily functionally equivalent. In fact, even adherent populations isolated from bone

marrow are actually heterogeneous populations containing not only true self-renewing

15

undifferentiated stem cells, but also downstream progenitors and more lineage-restricted

precursor cells [9]. These complications have prompted the International Society for Cellular

Therapy to define the following set of minimum criteria for MSCs [10]:

1) Be able to adhere to tissue culture plastic under normal culture conditions;

2) Be positive for the expression of CD105, CD73, and CD90, and furthermore absent for
   the expression of hematopoietic surface markers of CD34, CD45, Cd11a, CD19, and HLA-
   DR;

3) Have the ability to differentiate into osteocytes, adipocytes, and chondrocytes *in vitro*
   under specific stimulatory culture conditions.
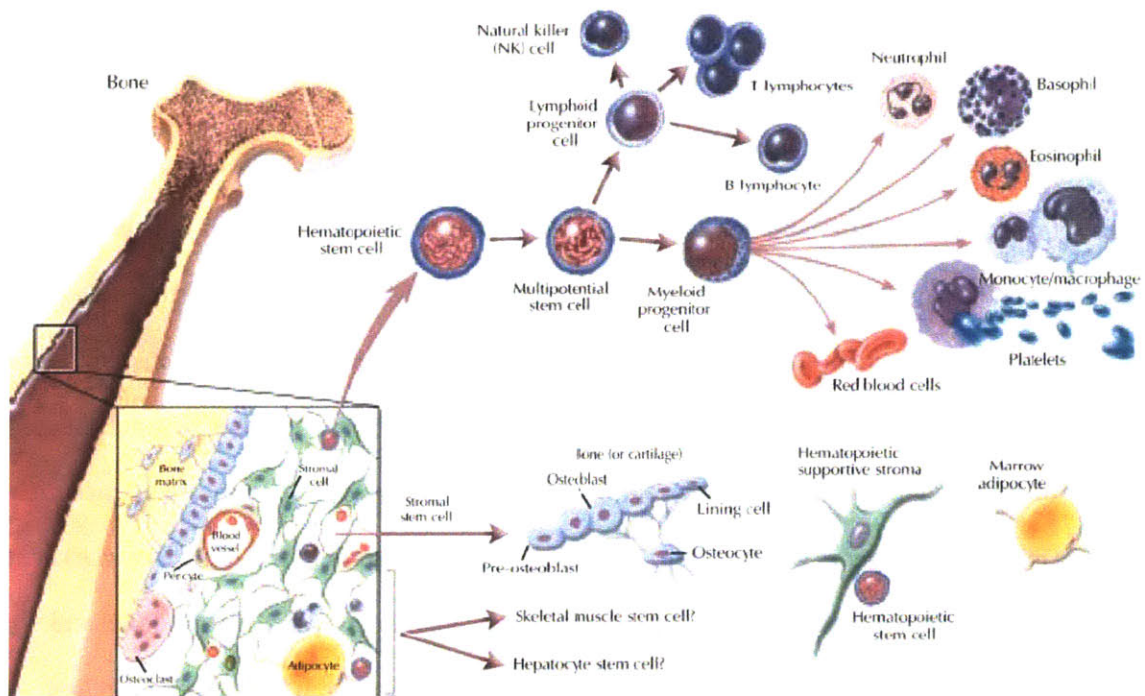
Figure 1.1. Bone marrow derived hematopoietic stem cells and stromal cells have the potential
to differentiate down multiple lineages. *Source: Zech, 2004* [2].

Despite the heterogeneities associated with isolated MSC populations, these cells nevertheless are enthusiastically studied because of their therapeutic potential in a variety of clinical applications such as tissue repair and regeneration, immunomodulation, and gene modification for designing disease- and patient-specific cell-based therapies [11-14]. Bone-marrow derived MSCs naturally motivate applications in bone and cartilage tissue repair and regeneration and have been shown to enhance and lead to bone and cartilage formation [15-18]. The conventional hypothesis for MSCs' role in bone tissue repair has been that MSCs proliferation and differentiate into target cell types to replace injured cells, but recently there is evidence that at least in cases of acute injury, MSCs actually indirectly promote tissue repair by secreting large quantities of bioactive molecules such as cytokines and antioxidants that in turn mediate the stress response of the surrounding tissue [12,19]. Furthermore, the secreted molecules may recruit actual immune and reparative cells to the injury site [20,21].

Regardless of the particular paradigm believed to be MSCs' role in bone tissue repair, common challenges exist that must be addressed to ensure MSC therapeutic success:

1) Cell-based therapies require large number of cells, and MSCs are scarce in number.

2) Design of therapeutics require detailed and comprehensive understanding of cell behavior, and our knowledge of MSC behaviors of survival, proliferation, migration, and differentiation are largely lacking, particularly in the underlying signaling pathways that govern these behaviors.

1.1.1. Expansion of MSCs for bone tissue repair via *ex vivo* scaffolds

MSCs only comprise between 0.001% to 0.01% of the total nucleated cells from isolated bone marrow [1]. Thus, *ex vivo* expansion of isolated MSCs is a critical area of study so that adequate cell numbers can be achieved. One particularly promising approach is the engineering of biocompatible synthetic scaffolds that aid and promote MSC survival, proliferation, and even

17

differentiation [22,23]. Culturing MSCs on a 3D woven scaffold before implantation led to a higher osteogenic potential in rats versus implanting the scaffold with injected MSCs at the time of surgery, presumably because culturing *ex vivo* expanded cell numbers as well as ensured MSC survival under a more controlled environment [24]. Furthermore, there is evidence that culturing MSCs in protein scaffolds such as fibrin and collagen networks help new bone formation around hip implants, leading to better integration of prosthetic devices [25]. These show the therapeutic potential not only of MSCs, but also of scaffolds in the application of bone tissue engineering.

The design and engineering of such scaffolds is a non-trivial process as the landscape of considerations for their properties is vast and varied (Figure 1.2) [26]. The scaffold material must be biocompatible and biodegradeable in addition to being able to promote MSC survival followed by infiltration and population of the scaffold before implantation back into the injured tissue. Numerous designs of scaffolds have been investigated toward this goal, varying engineering parameters to provide a multitude of biophysical as well as biochemical cues appropriate to drive MSC behavior. Materials such as ceramics and bioglass have been shown to be biocompatible as well as to provide effective structural integrity for the expansion of MSCs [27,28], while other materials such as polymer and hydrogel scaffolds have been studied for driving MSC differentiation down osteogenic or chondogenic lineages [29,30], while still other advancements in materials design have been developed for their unique and beneficial geometries and mechanical properties [31].

More recently, scaffolds functionalized with bioactive molecules such as ECM proteins and growth factors have further advanced this field, with the hypothesis that the incorporation of these molecules from MSCs' native *in vivo* environment would better mimic the mechanical, structural, and stimulatory properties of the target tissue. This in turn would provide more realistic *ex vivo* environments during MSC expansion and aid in successful integration back into the *in vivo* sites of injury [32,33].

Most of these approaches however, take a rather "trial-and-error" approach to scaffold design, investigating the effects of one or two parameters. As biomaterials research advances, these studies emerge to determine the effects of new materials and properties on MSCs for possible applications in bone tissue engineering. However, a more systematic and rational approach, on the other hand, is to fundamentally understand MSC behavior and its underlying governing signal pathways so as to design scaffolds with the minimally effective set of materials and functional properties to drive a specific MSC response.
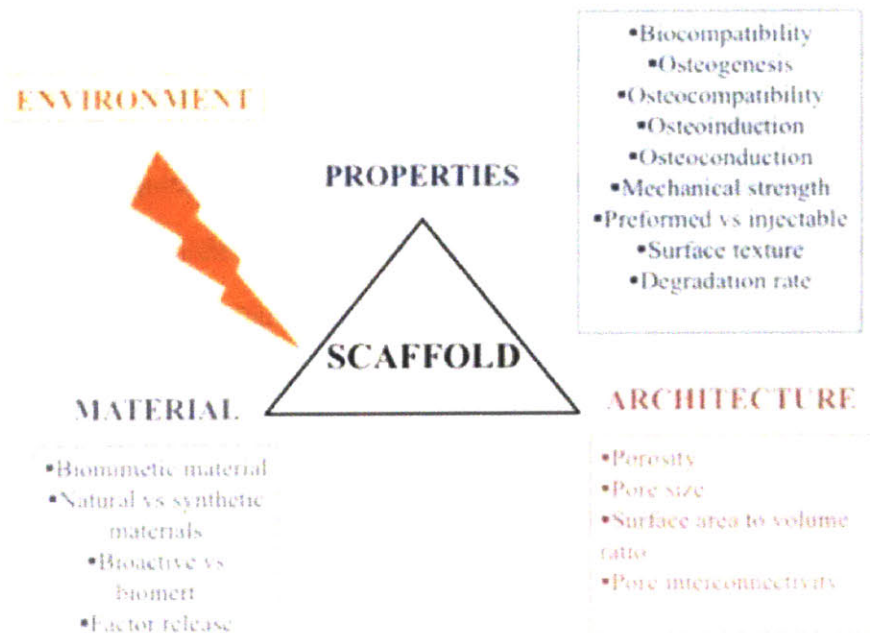
Figure 1.2. Numerous design and engineering considerations affect the success of scaffolds for bone tissue engineering applications. *Source: Szpalski et al., 2010* [26].

1.1.2. Understanding MSC behavior for scaffold design

The challenge to more rationally designing scaffolds to affect MSC behavior is that we actually understand very little regarding the MSCs behaviors of proliferation, survival, migration, and differentiation. We do know that a variety of cues present in the MSC *in vivo* environment affect their behaviors, such as growth factors and the extracellular matrix (ECM). The majority of studies investigating the effects of these cues on MSC behavior have focused on survival, proliferation, and differentiation, with a notable lack of similar studies of MSC migration [34,35]. This is an unfortunate imbalance as MSC migration is just as critical of a process in injury repair.

All of this motivates the particular goals of this thesis:

1) To rigorously and quantitatively characterize the effects of growth factor and ECM cues on MSC migration, and

2) To investigate the underlying biochemical mechanisms affecting MSC migration to inform the rational design of bone tissue engineering scaffolds for injury repair.

The remainder of this chapter will explore the background of growth factor and ECM cues for MSC behavior, particularly MSC migration. Where available, we cite work involving MSCs. However, as MSC migration studies are relatively sparse, we borrow also from our knowledge of other migratory adherent mesencymal cells such as fibroblasts.

1.2. Effects of growth factor and ECM cues on MSC behavior

1.2.1. MSCs and growth factors

As previously mentioned, one of the key challenges to using MSCs for clinical therapeutics is the rarity of the cells, making their *ex vivo* expansion critical for clinical success. Growth factors, well-characterized for their effects on increasing cell proliferation are prime candidates to aid in MSC expansion. Several growth factor families have been identified to increase MSC survival and proliferation including transforming growth factor beta (TGFβ), fibroblast growth factor (FGF), vascular endothelial growth factor (VEGF), insulin growth factor (IGF), platelet-derived growth factor (PDGF), hepatocyte growth factor (HGF), and epidermal growth factor (EGF) [34,36].

TGFβ, and TGFβ super family growth factors such as bone morphogenetic protein (BMP), increase MSC proliferation and promotes them down a chondrogenic lineage [37,38]. FGF is shown to increase MSC proliferation and migration, but also biases them toward a chondrogenic lineage [39-41]. PDGF and HGF are known to increase MSC proliferation, survival, and migration [42-44], perhaps through MAPK and ERK signaling [44,45]. EGF has been similarly shown to increase MSC proliferation, but one key difference with EGF is that it does so without triggering differentiation down any specific lineage [45,46]. This latter point is especially crucial for bone scaffolds populated with progenitor cells *ex vivo* as differentiated cells within the scaffold limits the injury sites for which it would be helpful. Furthermore, growth factors triggering differentiation may interfere with MSC expansion, compromising the goal of achieving sufficient cell numbers for therapeutic applications.

Some of these same growth factors mentioned previously have been studied for their effects on MSC migration. Boyden Chamber migration assays identified several growth factors that promote MSC migration, including fibroblast growth factor-2 (FGF-2), platelet-derived growth factor (PDGF), vascular endothelial growth factor (VEGF), insulin growth factor (IGF), and epidermal growth factor (EGF) [42,45,47-49].

21

1.2.2. MSCs and the ECM

The extracellular matrix is a heterogeneous mix of structural fibers such as collagen and laminin with cell adhesion proteins such as fibronectin and vitronectin mixed with biomolecules such as proteoglycans and growth factors that all play a role in mediating cell behaviors such as adhesion, migration, proliferation, survival and differentiation [50]. Cells within ECMs in turn influence their surrounding environment through secretion of proteases, enzymes, and even ECM components to modulate and remodel the matrix [51]. Further complicating the effects of ECM on cell behavior, ECM and growth factors mutually influence each other through sequestration and/or increased cell activities from ECM- and growth factors-stimulated signaling [52].

Besides surrounding cells in three dimensions and providing architecture and mechanical properties to influence cell behavior, the ECM also presents a set of molecular signals that can stimulate cellular response via integrins, a family of heterodimeric adhesion receptors [53]. The role of integrin-binding on MSC differentiation has been investigated in several recent studies, including results showing that osteogenic differentiation depends on focal adhesion kinase (FAK) activation and that vitronectin and collagen I regulate osteogenesis via integrin-signaling [54,55]. Kundu and Putnam showed that MSCs express the integrin heterodimers that mediate cell adhesion to ECM proteins such as collagen I, fibronectin, and vitronectin [54]. Specifically, the $\alpha_v\beta_3$ and $\alpha_5\beta_1$ heterodimers, the ones that bind vitronectin and fibronectin respectively, were significantly higher expressed in MSCs. The inhibition of these heterodimers in MSCs reduced the cells' ability to differente [54].

The effects of ECM on MSC migration is even less studied than growth factors. Ode et al. recently investigated the effects of 13 ECM components on MSC migration via Boyden Chamber assay and found that fibronectin and collagen increased MSC migration, while other components such as vitronectin and laminin had no significant effects on migration [56]. In a separate study,

also using a transmembrane assay, fibronectin, vitronectin, and collagen I were all shown to induce MSC migration [57].

1.2.3. Effects of growth factor and ECM cues on MSC migration

MSC migration is a critical component in the success of MSC-based therapeutics both for *in vivo* bone injury repair as well as *ex vivo* scaffold culturing of MSCs. *In vivo*, the homing of MSCs to injury sites has been likened to leukocyte rolling, adhesion, and migration [58,59], but the specific biophysical and biochemical are largely unknown. Similarly, *ex vivo* expansion of MSCs requires them to be motile in order to infiltrate and populate the entire scaffold as opposed to densely-populated but isolated clusters. Without migration, MSCs survival and proliferation may be largely ineffective as stationary cells would strictly confined injury repair to a small area as opposed to the entire tissue or even surrounding tissue.

To this end, the few migration studies that do exist in literature are actually quite limited in the applicability of the results to bone tissue engineering. Firstly, transmembrane assays measure directed migration which is not necessarily relevant for *ex vivo* scaffold expansion of MSCs. Scaffolds are usually designed to have identical properties throughout so as to reach the end goal of a consistent scaffold full of expanded MSCs for integration to the target tissue. Secondly, while Boyden chambers are good candidates for initial screening of factors affecting directed migration, they offer little insight into the temporal- and substratum-dependent morphology of migrating cells. These limitations motivate techniques to study MSC migration so that the biophysical processes involved in cell motility can be observed.

1.3. Cell migration

Cell migration is an integrated process that involves protrusion at the cell front, translocation of the cell body, and finally detachment at the cell rear [60]. There is a multitude of protein signals shown to be involved in this process, induced by external cues such as growth factors and integrin-binding via adhesion to substratum. All of these signals must be precisely organized spatially and temporally so as to coordinate the multi-step process of cell migration

1.3.1. Biophysical processes of cell migration

Cell migration on 2D substrates is a highly regulated process consisting of a series of biophysical steps (Figure 1.3). Cells first undergo morphological polarization through active membrane extensions before deciding on a direction to travel and forming lamellipodia and filopodia at the leading edge. The membrane extensions at the leading edge attach to the substratum and stabilize into focal adhesions, after which contractile forces move the cell body forward. After translocation of the cell body, efficient mechanisms to release the adhesions at the rear of the cell complete the process of cell migration [60]. These steps of polarization, adhesion, translocation, and detachment are all mediated by protein signals, many of which are signals downstream of EGF receptor (EGFR) and integrin signaling.

During cell polarization, phospholipase C-γ (PLCγ) and Rho family proteins, downstream from EGFR signaling, localize to the leading edge of migration cells and aid in the actin reorganization necessary to establish polarity and to subsequently form lamellipodia and filapodia [61-63]. The differences between the front of the cell versus the back during polarization may be further regulated by integrins, vesicular transport, and phosphoinositide 3-kinases (PI3Ks) [62]. Adhesion contacts between the cell and substratum are also formed near the leading edge of migrating cells, and persist through the translocation and detachment steps [60]. These contacts increase in affinity with the activation of protein kinase-C (PKC), recruiting a multitude of proteins

to the site including focal adhesion kinase (FAK), suggesting a role for PKC in the maturation of

these contact points as they form in the lamellipodia and as the rest of the cell moves over them.

PKC itself is downstream of EGFR signaling, a result of PIP_2 hydrolysis into IP_3 and DAG by PLCγ

[64].
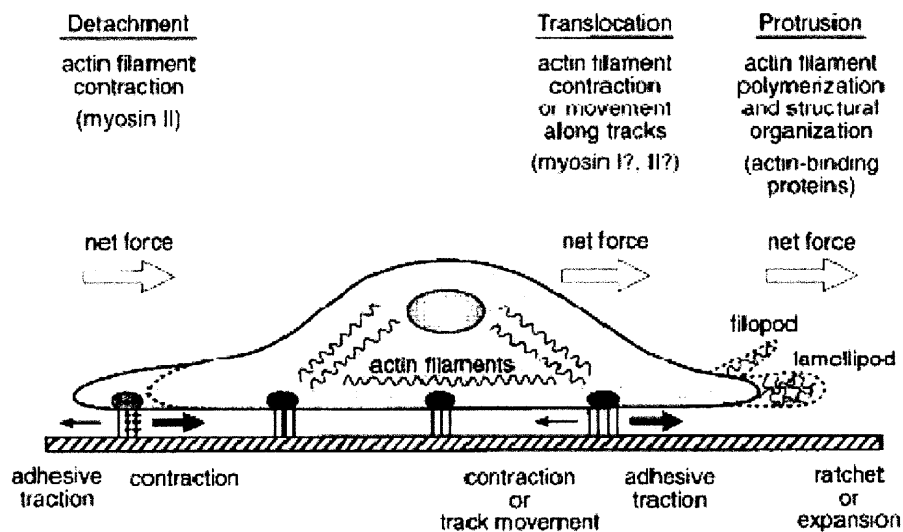
Figure 1.3. Cell migration across a 2D adhesive substratum is an integrated process involving many biophysical processes and forces. *Source: Lauffenburger & Horwitz, 1996* [60].

Following adhesion at the front of the cell, translocation of the cell body is the result of

contractile forces generated across the cell from the leading edge. The adhesion contacts the cell

makes act as the anchor points for traction as the contractile forces pull the rest of the cell to "catch

up" with the protrusions at the leading edge. This force generation is accomplished by the

interaction of myosin II with the actin filaments that attach to these adhesion contacts [65]. Several

protein signals affect myosin II activity, including myosin light chain (MLC), myosin light-chain

kinase (MLCK), Rho kinase (ROCK), extracellular-regulated kinase (ERK), and even PLCγ [62].

There is a tradeoff of MLC activity for cell migration. On the one hand, myosin needs to be activated in order to exert a tensile force for traction on the substratum at adhesion sites. On the other hand, too much traction would make it impossible for the cell rear to detach. The balance between traction and contractility depends on substratum adhesivity, thus migration was theorized to be biphasically correlated with the strength of the attachment at adhesion contacts [66], and later shown experimentally in EGF-induced fibroblast migration [67]. Several reports show that MLC phosphorylation is regulated by ERK and Rho, with ERK inducing cell contraction by phosphorylating MLCk which in turn phosphorylates MLC [68].

Finally, adhesions need to be disassembled and detached in the rear of the cell. This process is regulated by a number of proteases that cleave adhesion points to allow for cell detachment. The calpain family of proteases is heavily involved in disassembly of adhesions at the cell rear and is affected by both integrin- and growth factor-mediate signaling [69]. ERK is thought to be a regulator of cell adhesions by phosphorylating calpains as well as FAK [70], the latter a crucial player in integrin-associated adhesion formation.

1.3.2. Modulation of cell migration for disease therapy

Cell migration is well-studied for its pivotal role in the disease progression of a variety of fields such as wound repair and cancer. In cancer, dysregulation of the motility process leads to the migration of otherwise senescent and non-motile cells, leading to tumor invasion and metastasis. It follows then that inhibiting migration in the appropriate tumor and tissue environments would have beneficial therapeutic effects [71,72]. Similar to bone tissue repair, cell migration in cancer therapy applications is relatively less studied than proliferation and cell death, but nevertheless presents a viable approach for disease therapy [73].

Several studies have recently demonstrated the potential of modulating cell motility for cancer therapy with great success [74-76]. Many of these modulators are small molecule drugs that

downregulate governing pathways for motility, thereby inhibiting cell migration to stop metastatsis and disease progression. In the case of injury and wound repair however, the desired effect of therapeutics is usually to increase cell motility, proliferation, survival, and other responses to promote tissue repair. Thus, the approach of downregulating signals to attenuate cell motility, unlike in cancer therapy, is not productive. Rather, we aim to identify signals whose downregulation would lead to upregulation of cell motility response. The opposing nature of these inhibition targets with the intended results of promoting cell response suggests that while theoretically possible, these therapeutic targets are likely to be non-intuitive and difficult to find by trial and error.

This further underscores the importance of comprehensive understanding of the biochemical effectors of the biophysical processes of MSC migration. This understanding may illuminate key signaling pathways that are non-intuitive for the rational and systematic discovery of protein targets that are therapeutically beneficial.

1.4. Thesis motivation, objectives, and outline of chapters

1.4.1. Motivation and objectives

Given that EGF promotes MSC migration, survival, and proliferation without biasing the cells down any particular lineage, it emerges as an optimal growth factor to study for the expansion of MSCs *ex vivo*. Furthermore, EGF-induced fibroblast migration is a well-characterized process both in response and in protein signal effectors, thus providing a good platform from which to study MSC migration.

While our understanding of MSCs has advanced in the last decade or so, MSC migration has remained relatively unstudied, leading to significant voids in our knowledge of MSC biology and limiting the therapeutic potential of bone stromal cells.

All of this motivates the particular goals of this thesis:

1) To rigorously and quantitatively characterize the effects of EGF and ECM cues on MSC migration, and

2) To investigate the underlying signaling pathways affecting MSC migration

3) To use MSC migration response and signaling knowledge to inform the rational design of bone tissue engineering scaffolds for injury repair.

1.4.2. Outline

This thesis will progress along five chapters, building on each other to ultimately show a two-layer combined biomaterials and small molecule solution for controlling MSC migration.

Chapter 2 discusses decision tree signal-response modeling to generate testable hypotheses for protein signal modifications that would change fibroblast cell migration response. This lays the groundwork and proof of concept for a similar approach to studying MSC migration.

Chapter 3 presents a model system for biomaterial scaffolds on which to grow MSCs and explores the combined effects of substrate conditions and soluble EGF on MSC migration.

Chapter 4 builds on the findings in Chapter 3 to present a biomaterials surface with immobilized EGF (tethered EGF, or tEGF). MSC migration response on these surfaces which combine ECM and growth factor presentation is thoroughly and quantitatively explored. Furthermore, MSC signaling as mediated by these cues is measured and characterized.

Chapter 5 borrows the methods presented in Chapter 2 to study fibroblasts and reports on their applications to MSC migration signal and response. We develop decision tree models to analyze MSC migration and generate non-intuitive hypotheses and report on experimental test results. This chapter presents the key findings of this thesis.

Finally, **Chapter 6** summarizes the findings and conclusions of this thesis while suggesting future areas of study.

2. Chapter 2 - Decision tree prediction of fibroblast migration

2.1. Introduction

A major goal of systems biology and proteomics is to find computational techniques to accurately model the relationships between protein signaling pathways and cellular functional responses such as cell migration and protein signaling [77-80]. Data driven models attempt to reach this goal by finding correlative "cause-effect" relationships between proteins, genes, and functional cell responses and have productively elucidated the signaling pathways of migration in a variety of contexts [81]. Cells can be stimulated via various cues, resulting in protein signaling changes as well as changes in behavioral response. The central hypothesis is that the cues modulate the particular responses through the changing proteins pathways, and the ultimate cellular responses depend on the states of protein activities. In this paradigm, the "cause-effect" relationship becomes a defining function that "calculates" the cellular response given a particular

set of protein signal inputs: $f(X)$ = **y** where **X** is the matrix of input signals, **y** the vector of responses, and $f(X)$ the model.

Work prior to this thesis explored decision tree modeling of cell "signal-response" relationships using EGF-induced fibroblast migration as a model system [82]. The cues used to affect fibroblast migration were four different surface fibronectin adhesivnesses with or without additional stimulation of EGF for eight total combinations of cue conditions. Maheshwari et al. reported fibroblast migration on these eight conditions extensively and showed fibroblast migration speed to be biphasic versus fibronectin adhesivness. That is, speeds were highest at intermediate levels of fibronectin where there was an optimal balance of adhesiveness for attachment at the cell front without hindering detachment at the cell rear [67]. The intracellular phosphorylation levels of five signaling proteins five minutes after EGF-stimulation were measured using the same eight conditions in fibroblasts. These five signaling proteins were EGFR, ERK, MLC, PLCγ, PKCδ, chosen for their roles in driving the major biophysical processes of cell migration as well as being downstream of EGF-induced signaling. PLCγ is a major signal in lamellipodial protrusion, while PKCδ and MLC affect cell contractility and force transmission, and m-calpains regulate detachment of cell membranes at the rear of the cells [61]. After rigorous quantitative treatments of the signaling dataset to significantly expand the information content, several thousand decision trees were generated using discretized signals as predictors of the measured migration speeds.

Figure 2.1 shows the best-performing decision tree from that previous study [82]. The protein signals form the nodes of the trees as predictors (round), and the migration speed categories are the leaves (square). Numbers attached to each branch indicate the split categories of the parent node. Percentages below each leaf represent the fraction of cases of that migration category explained by this leaf classification. For example, if ERK is 0, the migration category is 0,

and 90% of "0" migration falls into this leaf. By contrast, if ERK is 1, and MLC is 2, the migration

category is 1 with 42% of cells migration at "1" classifying to this particular leaf.

One primary goal of data-driven computational models for cell "signal-response" is the

ability for these models to generate otherwise non-intuitive hypotheses under signaling landscapes

altered from that of prior experiments. The decision tree in Figure 2.1 presents one such *a priori*

prediction. When ERK is 1, and MLC is 2, the tree predicts that cell migration speed will be "1," or

intermediate. However, if ERK is 1, and MLC is also 1, the tree predicts that cell migration speed

will be "2," or high. Thus, the hypothesis is that decreasing MLC signaling will increase fibroblast
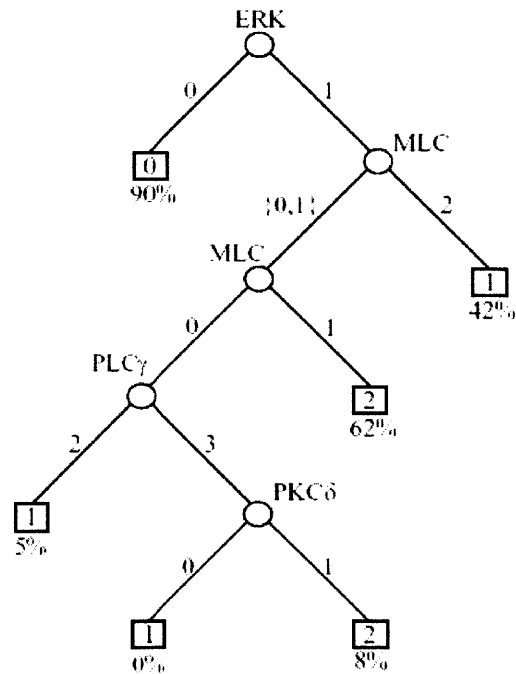
Figure 2.1. Decision tree model classifies
fibroblast migration speed using signaling
proteins measured 5 minutes after EGF
treatment. *Source: Hautaniemi et al., 2005*
[82].

migration speed. This prediction is not readily apparent from evaluating the original protein signals dataset, only after the classification of migration speeds using these protein predictors.

In Chapter 2, we discuss using decision tree models to predict non-intuitive response results from signaling changes and the associated experimental proof of this hypothesis.

2.2. Methods and materials

2.2.1. Cells and cell culture

We used NR6 mouse fibroblasts derived from a 3T3 lineage that are devoid of endogenous EGFR and overexpressed with human EGFR. These cells are called NR6 wild types (NR6 WT) and provide a good model to study EGFR mediated signaling events as well as cellular biophysical processes like migration.

NR6 WT cells were maintained in modified Eagle's medium-α containing (MEMα) 7.5% fetal bovine serum (FBS) and 1% of each of the following: penicillin/streptomycin, L-Glutamine, non-essential amino acids and sodium pyruvate. The medium contained 350 μg/ml of Geneticin as a selection agent for human EGFR. Quiescent and assay media containing 0.5% dialyzed FBS, instead of 7.5% full FBS, was used during experiments to minimize the effects of serum growth factors. All cell culture reagents were purchased from Gibco, a division of Life Technologies (Carlsbad, CA).

2.2.2. Fibronectin adsorption to migration surfaces

Sterile stock fibronectin (Sigma, St. Louis, MO) was diluted in 1x phosphate-buffered saline (PBS) to the desired concentrations of 0.3, 1, 3, and 10 ug/mL. 2 mL of diluted fibronectin was added to each DeltaT dish with 0.17 mm thick glass bottom (Bioptechs) and incubated at room temperature for two hours. The dishes were washed once with PBS and then incubated with 1% bovine serum albumin (BSA) (Sigma, St. Louis, MO) for one hour at room temperature to block surfaces with no protein adsorbed. After blocking, the dishes were washed three times with PBS, and used immediately.

2.2.3. Single cell tracking for cell speed analysis

NR6 WT cells were seeded at low densities, 6,000 cells per dish, on fibronectin-adsorbed substrata in 2 mL of quiescent media containing 0.5% dialyzed serum. The cells were allowed to

attach for 16 hours, after which the medium was replaced with 3.2 mL of fresh low-serum media, which we will call assay media. In migration versus fibronectin validation studies, the replacement assay media contained 10 nM EGF. In MLC inhibition studies, the assay media contained 0, 2, 4, or 10 µM ML-7 (MLCK inhibitor). 10 nM EGF was added 45 minutes after ML-7 exposure.

The plates of treated cells were sealed with a coverglass lid lined with vacuum grease and placed in a heated stage insert for a Ludl 99S008 motorized stage on a Zeiss Axiovert 35 microscope. Three to five random fields of cells, with five to ten cells per field, were chosen to be tracked for up to 20 hours, during which an image was captured per field every 10 minutes using OpenLab automated software and a Hamamatsu camera. Following the experiment, the images were stitched into QuickTime movies and exported as TIFF stacks for data analysis.

2.2.4. Image analysis for migration quantification

Based on previous reports that cell speeds steadily increase after EGF stimulation until reaching a plateau 4-6 hours afterwards [67], we calculated speeds for cells at time points after 6 hours to ensure we were capturing the steady-state migration response. The TIFF stacks for each field were analyzed with Visible software (Reify Corporation, Cambridge, MA) to extract individual cell coordinates for cell speed. Figure 2.2 shows part of a screenshot from Visible tracking a single cell. The algorithm evaluates time-adjacent images to determine pixel-by-pixel differences and assume that unchanging pixels are background whereas clusters of changing pixels are classified as moving cells. Visible identified the cluster of pixels associated with the cell in Figure 2.2, determined its boundaries, and also ascertained the center of motion of the cell (light blue crosshair). This "center of motion" gives the coordinates of this cell at this particular time point and is determined both by the location of the clusters of pixels as well as by the instantaneous velocity vectors associated with each pixel.
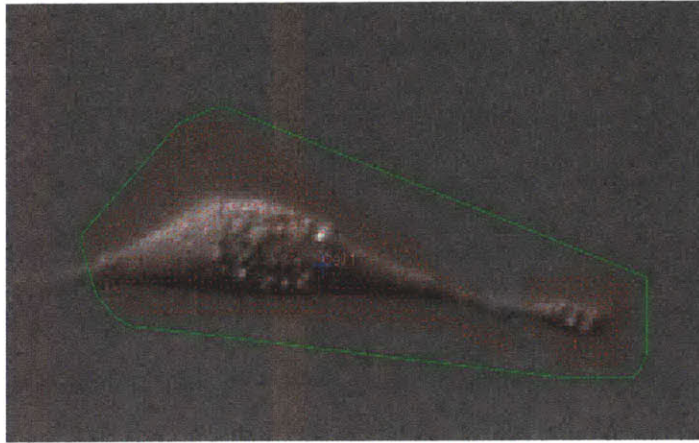
Figure 2.2. Visible tracking of cell migration. Red arrows on each pixel indicate instantaneous migration direction. Green boundary outlines the cell based on its moving pixels. Light blue crosshairs give the coordinates of the cell.

2.3. Results and discussion

2.3.1. Fibronectin-dependent fibroblast migration

Because Visible was a new addition to the lab, we first validated that it would indeed track

fibroblasts and yield comparable results as previous studies. We tracked 15-20 fibroblasts for 10

hours, analyzed the speeds as dependent on time, and plotted the average speeds between 6 and 10

hours against 4 fibronectin concentrations: 0.1, 0.3, 1, 3 ug/ml (Figure 2.3). The speeds as analyzed

by Visible also showed a biphasic correlation between cell motility and fibronectin concentration

where an intermediate level of fibronectin and optimal and yields the fastest cell motility.
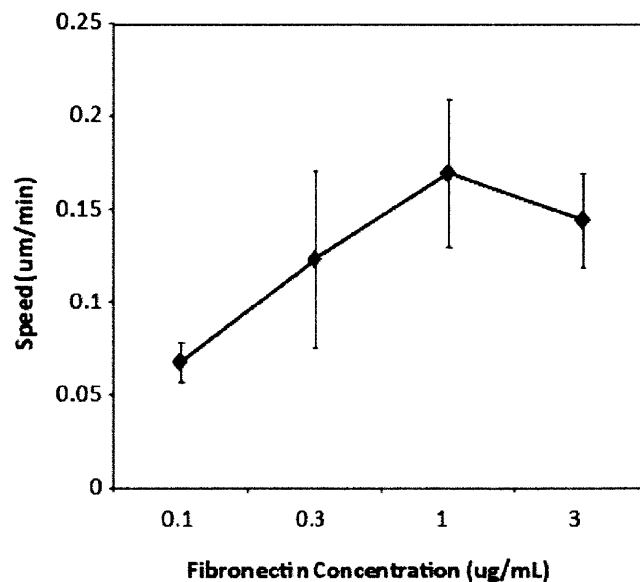
Figure 2.3. Fibroblast migration speed versus
fibronectin concentration. 15-20 cells were tracked
per condition, and error bars show +/- SEM.

2.3.2. Predicting testable hypotheses from decision tree model

We generated two new decision tree models in addition to the one in Figure 2.1 which uses signaling data five minutes after EGF stimulation. The two new trees were from protein signal datasets collected at 1 hour and 16 hours. As cellular response is dictated by the temporal activation and involvement of tens and hundreds of signaling proteins, it is important to ascertain the upstream signaling time points most predictive of longer-term responses. Using 1000 independent validation data sets, quantitatively generated from actual experimental data and associated statistical noise, the best-performing trees using the 1 h, 16 h, and 5 minute protein signals were compared. The 5-minute tree (Figure 2.1) correctly accounted for about 70% of the validation data sets. The 1-hour tree (Figure 2.4) accounted for over 75% of the validation data. The 16-hour tree, in contrast explained less than 60% of the validation data sets. Of the three trees, we chose to focus on the 1-hour tree since it performed the best with the validation data sets.

As seen in the decision tree model in Figure 2.4, after EGFR activation, MLC-mediated contractility was the most crucial ingredient in mediating maximal motility in fibroblasts. According to the predictions from this 1-hour tree, the cells move with highest speeds when EGFR is activated and MLC phosphorylation is low. In our training set, 68% of the situations in which cells move with high speed can be explained with this classification alone (red box, Figure 2.4). In other words, lowering MLC activation and resultant contractility to a subtotal level apparently leads to enhanced cell motility whereas total MLC inhibition can abrogate cell motility. While the effects of total MLC inhibition on cell motility have been intuitive and published by Iwabu *et al* [83], the biphasic dependence of cell migration (speed) upon subtotal inhibition of MLC is non-intuitive and novel. Moreover, it is an especially significant prediction for targeted therapeutics because it indicates that subtotal versus total abrogation of a key signaling pathway node can have drastically opposite cell responses.
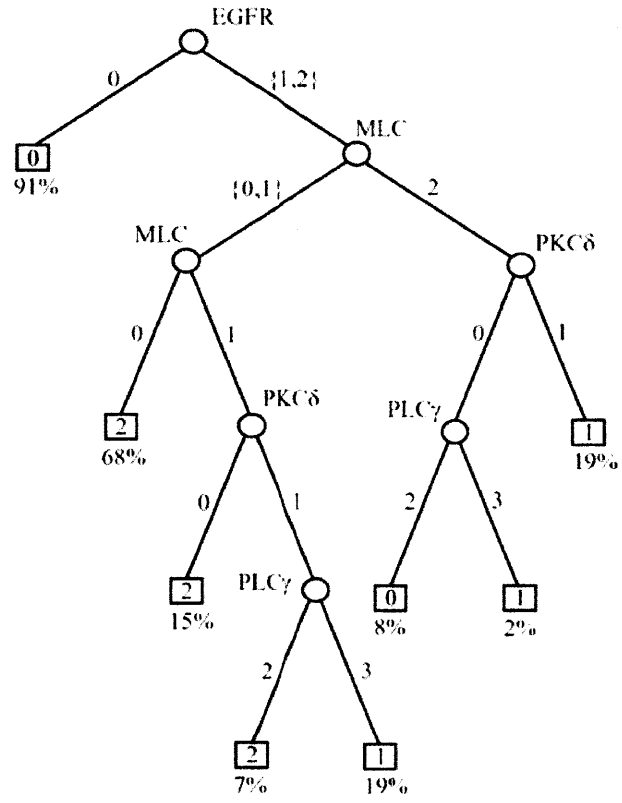
Figure 2.4. Decision tree generated from discretized 1-hour signaling data. Circles denote signaling proteins, whereas squares give the migration speed categories. Percentages below each leaf indicate the fraction of that class explained by that branch.

2.3.3. Subtotal inhibition of MLC activation increases cell speed

To test the model predictions under such conditions, we employed a well-characterized MLCK inhibitor, ML-7, to measure fibroblast migration speed under the same extracellular conditions (four fibronectin concentrations +/- EGF). This downstream inhibitor was chosen (as opposed to PKCδ inhibitor Rottlerin) because it is MLC kinase-specific. This means that resulting cellular responses can be attributed directly and specifically to MLC inhibition, whereas using PKCδ inhibitor Rottlerin would lead to unclear results since PKCδ is involved in diverse cellular responses in addition to motility [84]. In addition, fibronectin ligandation can activate MLC-based contractility, a process that is likely independent of PKCδ. Evidence of these differences between MLC inhibition versus PKCδ is further consistent with our decision tree model wherein MLC lies hierarchically above PKCδ.

Immunoblotting analysis of activated MLC (with EGF treatment) showed that phosphorylated MLC levels decreased with increasing ML-7 concentration on fibronectin (Figure 2.5). Three fibronectin concentrations were tested, 0.1, 1, and 3 ug/mL. MLC activity is completely abrogated at concentrations above 15 uM of ML-7.
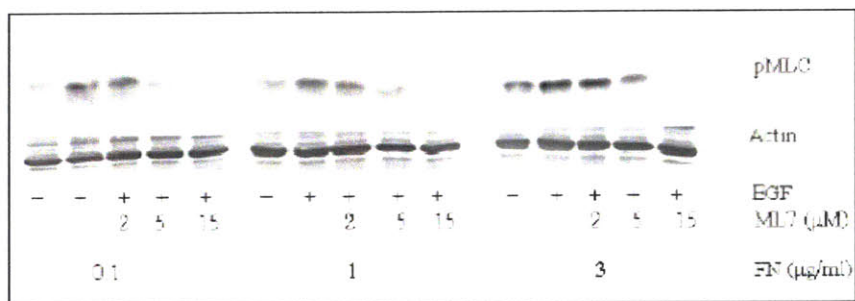
Figure 2.5. Effects of MLCK inhibitor, ML-7 on MLC activity on fibroblasts stimulated with EGF.

41

Using ML-7 to inhibit fibroblast MLC, we tracked cell migration on 3 ug/mL fibronectin

using 2 and 10 uM of ML-7. Average speeds of 15-20 cells per experimental condition are shown in

Figure 2.6. With no MLCK inhibition, fibroblasts migrated at an average speed of 0.077 um/min, or

4.62 um/hour. 2 uM of ML-7 slightly increased average speed to 0.090 um/min, and 10 uM of ML-7

significantly increased average cell speed to 0.14 um/min ($p < 0.05$). This outcome is consistent

with the decision tree model-predicted hypothesis that decreasing MLC would increase cell motility

speed. Further inhibition of MLC to the ranges of total MLC abrogration however, would

predictably lead to an ultimate decrease of cell speeds. This was confirmed in experiments with

MDA-MB-231 cancer cells [85].
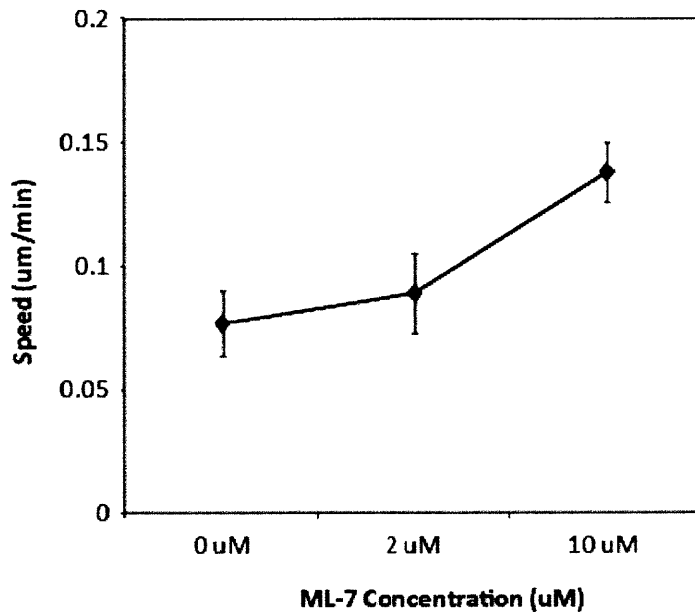
Figure 2.6. Inhibition of MLC increased fibroblast migration speeds in a dose dependent manner. Error bars show +/- SEM.

2.4. Summary and conclusions

This chapter demonstrated the success of decision tree analysis to identify the crucial effectors of cell motility depending upon a set of extracellular cues. The model was based on the quantitative measurements of five signaling proteins known to mediate EGF-stimulated cell motility. The analysis non-intuitively predicted that abrogating MLC, one of these five proteins would increase EGF-stimulated cell motility, a testable hypothesis that was proven with inhibition of fibroblast MLC activity using ML-7, a potent and specific inhibitor of MLCK, providing several proofs of concept:

1. Early signaling events in cells, even 5 minutes or an hour after treatment, are predictive of biological processes several hours later.

2. Decision trees is can predict cell behavior based on signaling datasets, even when the datasets are small, and the information contents low.

3. From a methodology standpoint, Visible was shown to be a reliable and quantitative analysis method for single-cell migration.

3. Chapter 3 – Effects of soluble EGF on MSC migration response and signaling

3.1. Introduction

Stromal cell migration is integral to the *in vivo* injury repair process as well as to *ex vivo* expansion in synthetic scaffolds for bone tissue engineering. In particular, epidermal growth factor (EGF) is a promising MSC stimulant in that it promotes MSC proliferation, survival, and migration without biasing the cells down any particular lineage. This chapter explores the effects of soluble EGF on MSC migration using an immortalized cell line of bone-marrow derived "mesenchymal" stem cells grown on a comb co-polymer surface as a model system for *ex vivo* MSC expansion on biomimetic scaffolds.

3.1.1. Immortalized MSC cell line

Studies involving multipotental stromal cells (MSCs) can involve primary cells of various origins such as human, murine, and porcine [86-88]. Though primary cells more closely model actual *in vivo* conditions, the heterogeneity of a cell population in addition to differences in cell phenotype from donor-to-donor make the results from MSCs difficult to interpret. Differential effects of integrins and growth factors would be hard to isolate from naturally occurring differences within the heterogeneous population that represent biological variation, often considered as random "noise" in experimental data [89]. Furthermore, primary cells undergo aging and

senescence-associated growth arrest after multiple population doublings, presenting hardships for long-term passaging during *in vitro* cell culture [90]. There is also evidence indicating that higher-passage MSCs have significantly reduced abilities to undergo differentiation, further limiting the applicability of these primary cells for our studies [91,92].

Using MSCs immortalized with human telomerase reverse transcriptase gene (hTERT-MSC) avoids many of these issues associated with primary cells while providing a relatively homogeneous population of MSCs with the same differentiation potentials as primary cells [93-95]. These cells retain their fibroblastic morphologies and proliferative properties even after 150-200 population doublings, and are able to differentiate into osteoblasts, chondrocytes, adipocytes, and endothelial cells even at these high passage numbers [94-96]. These characteristics suggest that hTERT-MSCs can serve as relevant models for studying mesenchymal stem cells while reducing the difficulties associated with culturing primary cells.

3.1.2. PMMA-g-PEO co-polymer surfaces

Many design considerations influence the materials composition of a suitable surface to present biomolecules to mediate cellular functions such as MSC migration. To excise exact control over surface properties achieved through functionalizing a base material, the base itself must be intrinsically resistant to protein adsorption and cell adhesion. In this manner, all interactions between cells and the surface would be the result of biomolecules used to modify the base with no contributions from cells interacting with any proteins that are non-specifically adsorbed to the surface such as serum proteins and/or cell secretions.

Toward this end, a poly(methyl methacrylate)-grafted-poly(ethylene oxide) (PMMA-g-PEO) co-polymer was developed in the Griffith lab [97,98]. The polymer consists of a hydrophobic PMMA backbone with hydrophilic PEO sidechains that extend away from the PMMA at the polymer-water interface, forming a comb-like structure leading to the polymer sometimes being referred to as a

comb polymer. The ends of the PEO sidechains provide sites for peptide coupling to further modify

the surface such as the tethering of EGF to present immobilized growth factors to cells (Figure 3.1).
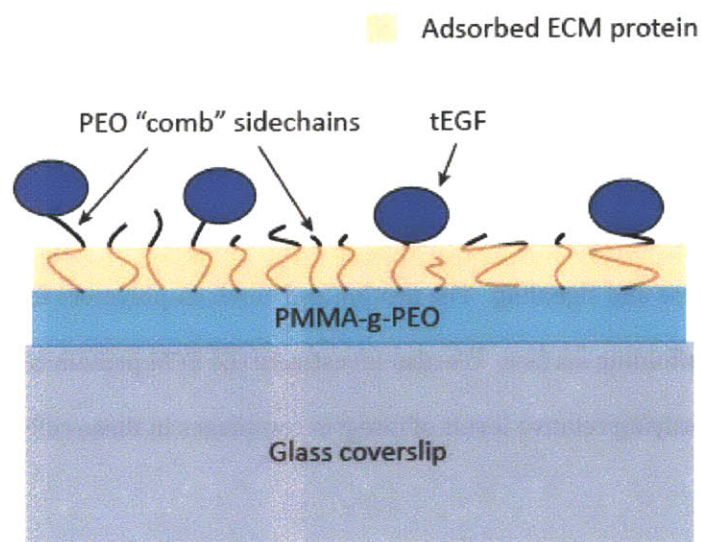
Figure 3.1. PMMA-g-PEO polymer base surfaces can be modified to provide multiple cues to affect hTERT-MSC migration. Figure not drawn to scale.

The design of this comb co-polymer allows for systematically varied densities of the PEO

sidechains, depending on the ratios of PMMA to PEO during synthesis. The various densities

subsequently give rise to differences in protein adsorption capabilities based on differences in

molecular interactions between protein molecules and the PEO brush [99]. Briefly, if proteins

molecules are assumed to be impenetrable spheres, there is some PEO comb density threshold

below which the protein "spheres" can settle between comb extensions to adsorb onto the polymer

surface. Above this density threshold, the PEO sidechains are too closely packed to allow proteins

to settle in between without sacrificing enormous amounts of energy inevitably necessary for

sidechain compression [99].

Following this principle, we synthesized two co-polymers, Comb 1 and Comb 2, differing in their percentage by weight (wt%) of PEO and thus in the density of their comb PEO sidechains [100]. Comb 1 contains 22 wt% PEO and allows for non-specific protein adsorption and thus is adherent for cells. Comb 2 contains 30 wt% PEO, does not enable protein adsorption and thus is cell-resistant.

3.1.3. Chapter scope

In this Chapter, we explore the effects of fibronectin-adsorbed PMMA-g-PEO surfaces on MSC migration response and signaling. The protein and comb co-polymers together present a model biomimetic scaffolding surface. We also investigate the ECM protein-binding potential of hTERT-MSCs by quantifying relative levels of integrin complexes in these cells.

3.2. Materials and methods

3.2.1. Cells and cell culture

An immortalized cell line of MSCs, hTERT-MSCs, was a generous gift from the Wells lab

(University of Pittsburgh) at passage 7. Cells were routinely cultured in Dulbecco's Minimum

Essential Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1 mM non-essential

amino acids, 2 mM L-glutamine, 1 mM sodium pyruvate, and 100 i.u./mL penicillin and 200 ug/mL

streptomycin. To maintain culture consistency, lot-matched FBS was purchased from Gemini Bio-

Products (West Sacramento, CA) in bulk. All other culture media components were purchased from

Gibco (Carlsbad, CA). Cells were cultured at 37 degrees Celcius with 5% CO_2, split 1:10 when they

reached 90% confluence to maintain culture (approximately 2 population doubles), and used until

approximately passage 25.

For experiments investigating effects of soluble EGF stimulation of MSCs, quiescent and

assay media containing 0.5% dialyzed FBS (Gibco), instead of 7.5% full FBS, was used during

experiments to minimize the effects of serum growth factors.

3.2.2. Biomaterials surface preparations

3.2.2.1.   Glass coverslip preparation for polymer coating

Circular glass coverslips 18mm in diameter were washed and silanized to prepare them for

polymer thin film spin coating. Silanization produces a hydrophobic surface that aids the polymer-

coating of the glass.

Coverslips were washed for four hours at room temperature under constant agitation while

submerged in 2% ChemSolve (VWR, West Chester, PA), a sodium-hydroxide-based cleaning agent,

diluted in double distilled water. After washing, coverslips were rinsed three times with water and

then agitated vigorously for 30 seconds at room temperature in a solution of 2% Siliclad (Gelest,

Morrisville, PA), active ingredient octadecylsilane, dissolved in double distilled water. Coverslips

were immediately rinsed three times in double distilled water and individually laid out on trays lined with aluminum foil and dried in a glass oven for 20 minutes. After drying, the coverslips were allowed to return to room temperature and stored in a humidity-free vacuum oven for up to 2 months until polymer spin coating.

3.2.2.2.   *Polymer synthesis and polymer spin coating*

Synthesis of poly(methyl methacrylate)-graft-poly(ethylene glycol), PMMA-g-PEO, co-polymers was done by Linda Stockdale following previously described methods [101]. The copolymer was diluted in toluene to a concentration of 20 mg/mL and put onto the surfaces of glass coverslips via thin film spin coating. 20-35 uL of the mixture were dropped onto each coverslips and spin coated for 30 seconds, forming a layer of polymer 8-12 um thick. The polymer-coated coverslips were placed in a humidity-free vacuum oven overnight to allow for the evaporation of any residual solvents before using.

3.2.2.3.   *Fibronectin adsorption*

Stock fibronectin was diluted in 1x PBS to the desired concentrations of 0.3, 1, 3, 10, and 20 ug/mL. The co-polymer surfaces were overturned onto 150 uL droplets of diluted ECM protein solutions and incubated at room temperature for 2 hours in a humidity chamber to allow for non-specific protein adsorption. Adsorbed surfaces were washed with 500 uL PBS each and overturned onto 150 uL droplets of 1% bovine serum albumin (BSA) (Sigma, St. Louis, MO) and blocked for 1 hour at room temperature. After washing three times with 500 uL PBS each time per coverslip, the surfaces were sterilized under ultraviolet (UV) light for 30 minutes while submerged in PBS and used immediately after sterilization.

3.2.3. Single-cell migration tracking and analysis

3.2.3.1.   Experimental setup for single-cell tracking

To capture single-cell migration, MSCs were sparsely seeded onto comb co-polymer and fibronectin surfaces and tracked for 6 hours on a microscope with an automated stage.

Prepared surfaces were glued to the bottom of 0.17 mm thick DeltaT dishes (Bioptechs, Butler, PA) using 2-3 small dots of 5-minute epoxy (VWR, West Chester, PA) near the edges of the coverslips. Cells cultured to 90% confluence in tissue culture plates were trypsinized and counted. 10,000 cells were seeded onto each 18mm surface (~4000 cells/cm^2) in low-serum media containing 0.5% dialyzed FBS and serum-starved for 16 hours in 37°C, 5% CO_2. After 16 hours, the cells were gently washed with warm PBS and the media replaced with 3.2 mL of assay media. Assay media contained 0.5% dialyzed FBS and 10 nM of murine EGF, a saturating amount.

The plates were then sealed with coverglass lids lined with vacuum grease and warmed in a 37°C incubator for 20 minutes, displacing excess assay media. The sealed dish of cells was then placed onto a heated stage insert for a Ludl 99S008 motorized stage on a Zeiss Axiovert 35 microscope. 8 random fields of 5-10 cells each were selected, and an image taken of each field every 10 minutes for 20 hours. The coverglass lid seals the dish to maintain the media pH at 7.4, and automated temperature control maintains the media temperature at 37°C for the duration of the experiment.

3.2.3.2.   Image analysis of time-lapse images

Each set of images for a field was strung together into a time-lapse movie and analyzed with Visible software (Reify Corporation, Cambridge, MA) to extract individual cell coordinates for cell speed and directionality calculations. Visible evaluates each movie to determine pixel-by-pixel differences between adjacent frames. Figure 3.2 shows one raw image (Figure 3.2A) together with its Visible-analyzed counterpart (Figure 3.2B). The analysis algorithm assumes unchanging pixels

to be the background and clusters of changing pixels to be moving cells. The minimum cluster size can be set so that Visible excludes small groups of moving pixels that are usually not cells and are instead particles or even optical artifacts. In Figure 3.2, four false positives were identified, boxed in black. For this particular tracked field-of-view, the minimum cluster size threshold can be raised to eliminate these false positives. Additionally, dividing cells and colliding cells would also cluster together as contiguous groups of changing pixels and would not be filtered by the cluster size requirement (white box, Figure 3.2). Thus, each analysis is manually edited to determine segments associated with pixel clusters that are not singly migrating cells.

For each cluster of changing pixels, Visible ascertains one pixel as the "center" of motion and outputs the coordinates of that pixel. This determination is made from a combination of cluster size as well as the instantaneous velocity vectors associated with each pixel, as determined from the change the algorithm detects in that pixel from frame to frame.
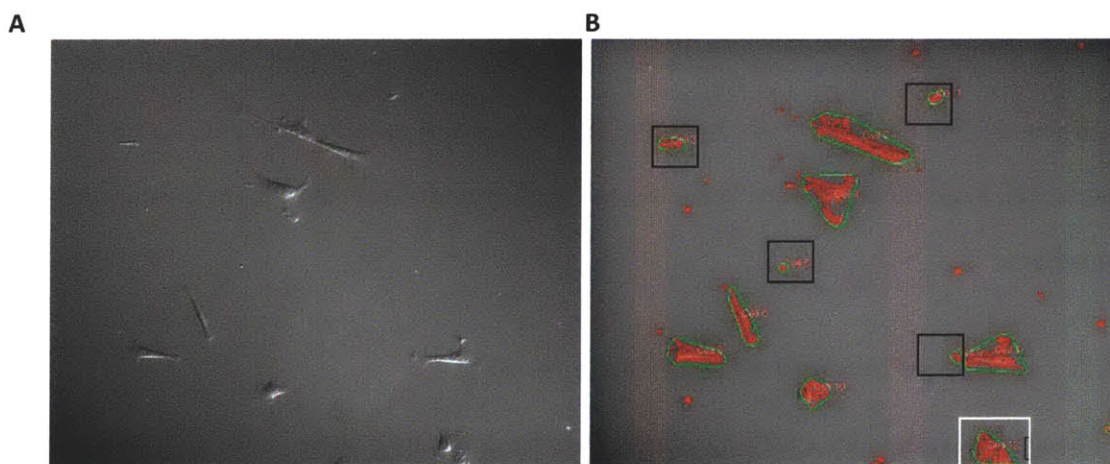
Figure 3.2. Visible analysis of cell migration from 2D images. (A) Raw image taken with 10x DIC objective. (B) Pixel-clustered images following Visible analysis showing clusters of pixels "in motion." The analysis correctly identifies all cells in the frame, but also identifies four false positives (black boxes). White box shows example of a dividing cell whose coordinates were excluded from the final data set.

We export these coordinates to an Excel file and together with the manual evaluation to determine clusters that should be excluded (dividing cells, colliding cells, cells partially out-of-frame, non-cells, etc.), exclude the coordinates from these clusters. The resulting files contain only coordinates for singly migration cells entirely contained in the frame. Furthermore, to minimize skewing due to small sampling size, we select only those cells with at least 1 hour of continuous coordinates (7 time points). These coordinates data were then imported into Matlab and analyzed via the Persistent Random Walk model using overlapping time intervals (see appendix for Persistent Random Walk Model Matlab code) [102].

3.2.4. EGF-stimulated phosphoprotein quantification

3.2.4.1.   *Lysate collection for signaling measurements*

Co-polymer surfaces adsorbed with varying amounts of fibronectin were deposited into 12-well plates, 1 coverslip per well, and covered with 1.5 mL of PBS and sterilized under UV light for 30 minutes in a sterile hood. Plates of hTERT-MSCs cultured to 90% confluence were trypsinized and counted. 100,000 hTERT-MSCs suspended in 1 mL assay media were seeded onto each co-copolymer coverslip and the cells serum-starved at 37°C, 5% CO_2 for 16 hours. After 16 hours, cells are gently washed with warm PBS and saturating amounts of soluble murine EGF (10 nM) were added. Cell lysates were collected 5, 15, 30, 60 minutes after EGF treatment, and an untreated plate was used for control.

Lysate collection was done at each time point as follows. A clean 12-well plate was filled with 1 mL of ice cold PBS per well. The cell plate was placed on ice and the media aspirated, replaced by 1 mL of cold PBS. The coverslips in each well were then removed and transferred to the clean plate of cold PBS. Because seeded cells settle on both the prepared surfaces as well as the tissue culture plastic of the 12-well plates holding the coverslips, this transfer assures that only those cells on the coverslip surfaces are lysed.

After aspirating the PBS, 25 uL of lysis buffer was pipeted onto each surface, and a cell scraper was used to gently remove cell lysates from the surfaces. Fresh lysis buffer was prepared immediately prior to each experiment consisting of lysis buffer base (Bio-Rad Laboratories, Hercules, CA), 1x Factors A and B (Bio-Rad), and 1 mM phenylmethanesulfonyl fluoride (PMSF) (Mallinckrodt Baker, Mansfield, MA). Cell lysates after scraping were transferred to 1.5 mL eppendorf tubes and centrifuged at 13,000 rpm for 10 minutes at 4°C. The supernatant after centrifuging was transferred to a new tube, lysate pellets discarded, and stored at -80°C until needed. A small volume was removed before storage to determine protein concentration via bicinchoninic acid assay following manufacturer protocols (BCA) (Pierce, Rockford, IL).

3.2.4.2. *Luminex xMAP platform to measure phosphoprotein signals*

Phosphoprotein signals were measured via the Bio-Plex suspension array system, a Luminex xMAP platform manufactured by Bio-Rad Laboratories (Hercules, CA). Luminex is a high throughput multiplexible quantitative bead-based antibody assay for protein detection. Briefly, small beads (5.6 um) are internally dyed with two spectrally distinct fluorophores in specific ratios, permitting identifiable spectral addresses to be assigned to a given set of beads giving it a unique spectral address. Particular bead subsets are conjugated to defined capture antibodies, enabling one spectral address per phosphoprotein to be measured. The captured proteins of interest are then quantified with fluorescently-labeled detection antibodies specific to the phosphorylation sites of interest. The Bio-Plex system reads single beads using a two-laser system that determines the bead's spectral address as well as quantifies reporter molecule florescence. Because each phosphoprotein is associated with a different bead spectral address, beads for different proteins can be combined and "multiplexed" within the same wells.

pERK (Thr202/Tyr204, Thr185/Tyr187), pAkt (Ser473), pEGFR (Tyr), pSrc (Tyr416) phosphoprotein bead kits and Phosphoprotein Assay kits were purchased from Bio-Rad

Laboratories (Hercules, CA) and manufacturer's protocol was followed. Briefly, a filter-bottom 96-well plate was washed twice with 100 uL Wash Buffer, and 50 uL of multiplexed beads solution added to each well. After two more washes of 100 uL Wash Buffer per well, with the beads retained in the wells by the filter bottoms, 50 uL of cell lysate diluted to 200 ug/mL in Lysis Buffer and Assay Buffer was added to each well and incubated at room temperature overnight with constant agitation. A blank well (lysis buffer only) and a negative and positive control for each phosphoprotein per set of experiments were run on each plate alongside the samples. The wells were then washed and incubated with biotin-conjugated detection antibody for 30 minutes at room temperature under constant agitation, followed by 10 minutes incubation with streptavidin-PE. A minimum of 25 beads were read to quantify each phosphoprotein per condition.

3.2.5. Flow cytommetry for integrin characterization

Fluorescent Activated Cell Sorting (FACS) was used to characterize and quantify the amount of integrins $\alpha v\beta3$, $\alpha5\beta1$, $\alpha2\beta1$ in MSCs. Cultured cells were detached with trypsin/EDTA (Sigma) and washed with full serum media. The cells were then centrifuged and washed twice with 1x PBS containing 0.1% Tween (PBS-T), after which the cells were transferred to 15 mL conical tubes pre-blocked with 1% BSA to minimize cell loss via non-specific attachment to the tube. These cells were centrifuged to form a pellet, the supernatant aspirated, and 500 uL of PBS added. After vigorous resuspending, 500 uL of cold methanol was added to each tube to fix the cells, and the tube vortexed on medium speed. The cells were incubated at -20°C for 30 minutes and labeled with antibodies specific for the integrins.

Fixed cells were vigorously suspended and transferred to FACS sample tubes. The cells were washed three times each with 1 mL PBS per tube, pelleting by centrifuging at 2,000 rpm each time. Antibodies for $\alpha v\beta3$, $\alpha5\beta1$, and $\alpha2\beta1$ were diluted 1:100 in PBS-T and 500 uL of antibody solution added to each tube of cells and incubated at room temperature for 1 hour or overnight at

4°C. The cells were washed twice with PBS, pelleting each time via centrifugation. 1:250 dilution in PBS-T of secondary antibodies against the species of the primary antibody was then added to each tube of cells (500 uL per tube) and incubated at room temperature for 1 hour covered in foil. The cells were washed twice with PBS, resuspended in 500 uL of PBS, and 10,000 cells were analyzed on a flow cytometer.

3.3. Results & Discussion

3.3.1. Persistent random walk modeling

The Persistent Random Walk model (PRW) of cell migration is a well-reported mathematical model to describe non-directed, "random," cell migration on two-dimensional surfaces [103]. Over short periods of time (usually less than 30 minutes), a cell typically moves in a relatively straight line, whereas over longer periods, cell migration paths resemble the random motion of inert particles characterized by many directional changes [104]. PRW presents one way to characterize and quantify these directional changes in cell migration.

The model relates the average cell speed (S), persistence time (P), and mean square displacement ($<D^2>$) in the following relationship, where T is the time interval over which the square displacement is measured:

$$\langle D^2 \rangle = 2S^2 P^2 \left(\frac{T}{P} - 1 + e^{-\frac{T}{P}} \right).$$

For each cell, we calculated the average cell speed over the entire migration path to use for S in the PRW model. The migration path was then sampled systematically for all possible time intervals lengths, over which the mean square displacement $<D^2>$ for all non-overlapping intervals of that time interval length for the entire migration path was empirically calculated from the cell coordinate data. This leaves the P, the persistence time, as the only unknown in the above equation. We mathematically fit for the value of P using ordinary least squares regression given a vector of T and a vector $<D^2>$, and the average speed S for each cell. The cell's Persistence Time is defined as the characteristic time over which the cell maintains the same direction of migration.

3.3.2. EGF-stimulated MSC migration response

3.3.2.1.   *MSC migration speed and directional persistence*

MSCs were serum starved on fibronectin adsorbed co-polymer surfaces for 16 hours and then stimulated with saturating levels (10 nM) of soluble EGF, after which the cells were tracked for up to 20 hours on an automated microscope stage. Cell coordinates were determined via Visible analysis, and average speed and persistence time calculated using Matlab. Figure 3.3 plots average cell speeds (red) and persistence times (blue). EGF-induced MSC migration was biphasic on fibronectin-adsorbed co-polymer surfaces whereby speeds were highest on intermediate levels of fibronectin (Fn = 1 and 3 ug/mL) and lower at the two extremes of low (0.3 ug/mL) and high (10 and 20 ug/mL) fibronectin. This is consistent with previously reported fibroblast migration on fibronectin surfaces of varying adhesiveness [67]. At intermediate substratum adhesiveness, there
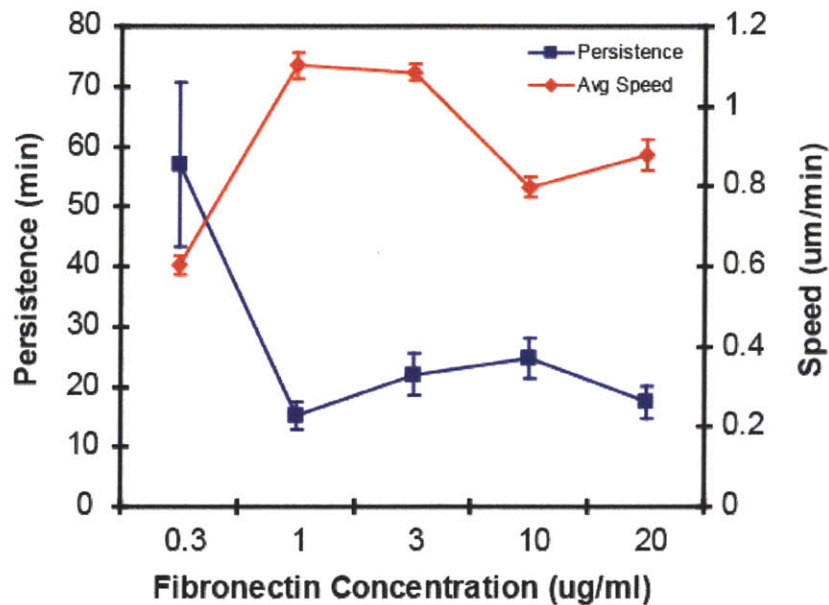
Figure 3.3. MSC Speed (red) and Persistence Time (blue) as a function of fibronectin concentration. Right axis shows speed (um/min), and left axis shows Persistence Time (min).

is an optimal balance between the cell's ability to attach at the front and its ability to detach at the rear. If adhesiveness is too low, the cell is unable to "anchor" at the front to initiate migration. In the case of high adhesiveness, while the cell can attach in the front, it is unable to detach in the rear to productively migrate. Additionally, fibroblasts are reported to spread more on higher levels of fibronectin, also reducing its motility [105]. This phenomenon appears to also be true for MSC migration on fibronectin-adsorbed polymer surfaces.

MSC Persistence Time (PT), on the other hand, appears to be inverse biphasic versus substratum adhesiveness. Persistence Time is highest at 0.3 ug/mL of fibronectin, significantly decreased at 1 ug/mL, followed by a slight increase through 3 and 10 ug/mL, and finally dips again at 20 ug/mL fibronectin. This reciprocal effect of EGF on cell motility speed and directional persistence is also consistent with previously reported studies of fibroblast migration [106]. The biophysical process attributed to this observation is the rate of membrane protrusion, which increases with EGF stimulation. As these protrusions increase resulting in net cell displacement, it makes sense that speed would also increase. These increased protrusions however, also increase the chances that a cell would change direction due to more active and rapid membrane activity [106].

3.3.2.2.   Quantifying effects of EGF on MSC migration

To determine the effects of EGF on MSC migration on these polymer surfaces, we investigated the effects of sub-saturating levels of EGF. Choosing the peak of the migration curve from Figure 3.3, we tracked MSC migration on 3 ug/mL fibronectin stimulated with 0 and 1 nM of EGF. Lowering EGF treatment from 10 nM to 1 nM had little effect on MSC migration speed (red line, Figure 3.4), whereas migration speed with no EGF stimulation was slightly decreased from saturating and sub-saturating levels of EGF. MSC directional persistence however was more significantly affected by EGF than speed. 10 nM of EGF significantly decreased migration

directional persistence from no EGF. This work and others recently conducted in our lab indicate

that directional persistence is likely actually separately regulated from cell speed, suggesting that

perhaps the signaling pathways governing speed are separate and different from those governing

directional persistence [107,108].
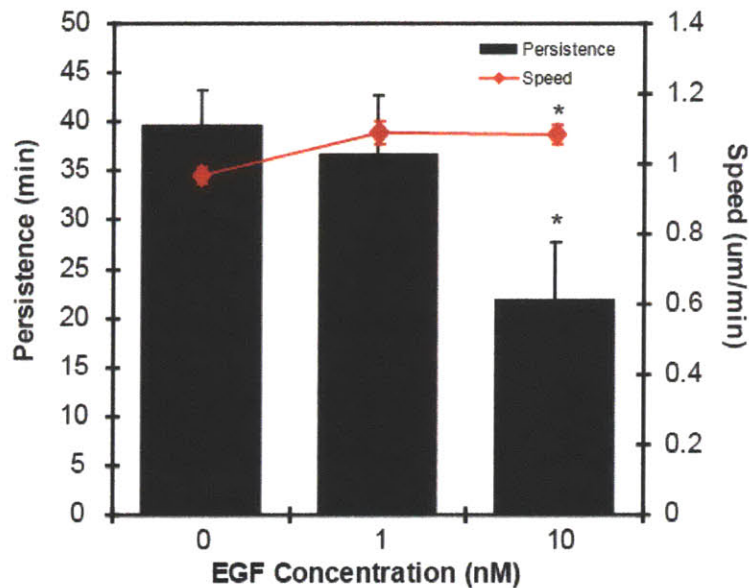
Figure 3.4. Effects of EGF stimulation on MSC migration speed (bar) and persistence (red line). All surfaces were adsorbed with 3 ug/mL fibronectin. Error bars show +/- SEM. * denotes statistical significance from zero EGF, the control condition (p < 0.05).

Putting this all together, at the intermediate fibronectin concentration of 3 ug/mL,

increasing concentrations of EGF increased cell speed (red, Figure 3.5), but decreased cell

persistence (blue, Figure 3.5). The effect on persistence was much more pronounced, with

saturating EGF causing a two-fold decrease in persistence, from 40 minutes with no EGF to 20

minutes with 10 nM of EGF. Speed, on the other hand increased from 0.95 um/min to 1.15 um/min.

Though this difference is statistically significant ($p < 0.05$), the effects are not as dramatic as directional persistence.

Because of the reciprocal effects of EGF on cell speed versus directional persistence, we calculated the distance traveled by a cell before changing direction ($S \times P$) known as the cell's Mean Free Path (MFP) which can be conceptualized as the cell's ability to "cover ground." Figure 3.6 shows that while 1 nM of EGF did not have a significant effect on MSC MFP, saturating levels of EGF at 10 nM significantly decreased the MFP. Conceptually as a cell response, this means that EGF decreases the distance that MSCs travel before undergoing directional change.
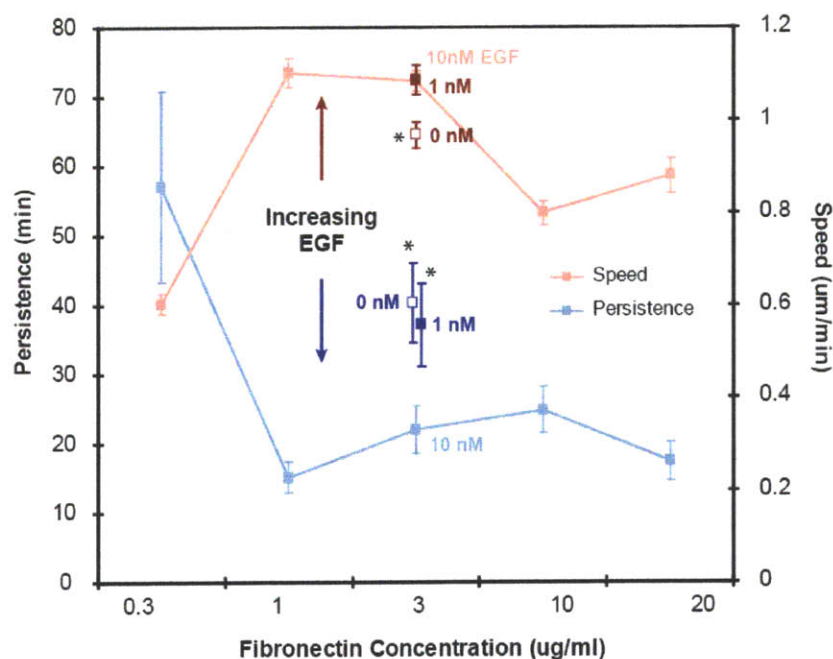
Figure 3.5 Effects of surface adhesiveness on EGF-induced MSC speed and directional persistence. Average speeds are in red/pink, and average persistence times are in blue. EGF dose-response was investigated for 3 ug/mL fibronectin, with EGF concentrations labeled. Error bars show +/- SEM. * denotes statistical significance against 10 nM EGF condition ($p < 0.05$).
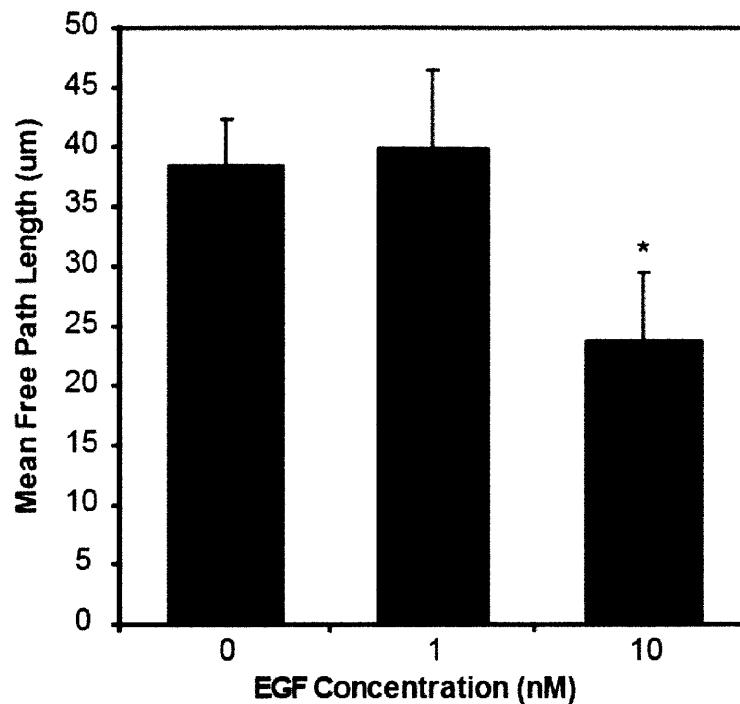
Figure 3.6 Mean Free Path as a function of EGF stimulation of MSC migration on 3 ug/mL fibronectin-adsorbed comb copolymer surfaces. Error bars show +/- SEM. * denotes statistical significance from the no EGF condition ($p < 0.05$).

This reciprocal effect on speed and directional persistence by soluble EGF, as well as the difference in efficacy, suggests that the signals downstream of EGF that govern cellular speed could be different than those that govern persistence, and that speed and persistence may be two separate responses. Depending on the context and application, it may be more desirable for cells to have increased persistence while maintaining speed, or perhaps increased speed but decreased persistence, having these two responses be governed by separate pathways would afford cells the ability to react appropriately depending on the application.

3.3.3. MSC phosphoprotein signaling

To begin exploring the underlying protein signaling affecting these migration responses in MSCs, we measured the levels of four phosphoproteins: EGFR, ERK, Akt, and Src. After 16 hours of serum-starvation, we treated these MSCs with saturating levels of soluble murine EGF and collected cell lysates at 5, 15, 30, and 60 minutes post treatment. Non-EGF-treated cells on all fibronectin conditions were lysed to provide a time zero control.

EGFR was phosphorylated immediately after EGF treatment of MSCs, though did not peak until 15 minutes after stimulation (Figure 3.7A). The Comb Only condition (dark blue line) was not treated with EGF at time zero, and accordingly shows no increase in EGFR phosphorylation through the entire time course. The highest levels of EGFR were on 10 ug/mL fibronectin 15 minutes after treatment, and levels were over 60-fold of the control. At this 15-minute peak, EGFR phosphorylation increased linearly with increasing fibronectin adhesiveness. We showed in Figure 3.3 that increasing fibronectin from 0.3 to 1 ug/mL increased cell speed, but speed stayed the same through 3 ug/mL Fn, and decreased at 10 ug/mL. This is in contrast to the monotonic increase of EGFR with increasing Fn. This suggests that EGFR alone cannot explain for the effects of EGF on MSC migration on fibronectin surfaces, further supporting the notion that cell migration is a "complex" function resulting from the combined effects of multiple signaling pathways.

Akt, like EGFR, was also phosphorylated immediately after EGF treatment, with maximal activation on 10 ug/mL of fibronectin at 5 minutes of about 30-fold over control (Figure 3.7C). Interestingly, also like EGFR, Akt activation increased monotonically to fibronectin concentration, with all EGF-treated conditions signaling above the level of non-treated cells. ERK appears to also be activated 5 minutes after EGF treatment, but there was no clear trend versus fibronectin concentrations. In fact, the non-treated condition (dark blue, Figure 3.7B) was not very different than the EGF-treated conditions. At longer time scales (30 and 60 minutes), ERK phosphorylation seems to decrease to below the time zero control levels. Src, in constrast to ERK, Akt, and EGFR,

was not phosphorylated above the levels of the negative control. The changes shown in Figure 3.7D are most likely due to background noise associated with Luminex measurements.

These signaling and response results together suggest that fibronectin is a dominant cue (compared to EGF treatment) in MSC migration speed. Decreasing EGF treatment to sub-saturating levels had very little effect on MSC motility speed, but speed was clearly biphasic versus surface adhesiveness as modulated by adsorbed fibronectin. Our signaling data suggests that fibronectin's effects on MSC speed is likely through EGFR and Akt signaling, and not ERK and Src, since the former two varied monotonically with fibronectin concentration. The differences in the trends of speed versus fibronectin and signaling versus fibronectin suggests that there are signals in addition to EGFR and Akt that also play critical roles and combinatorially affect MSC migration speed.
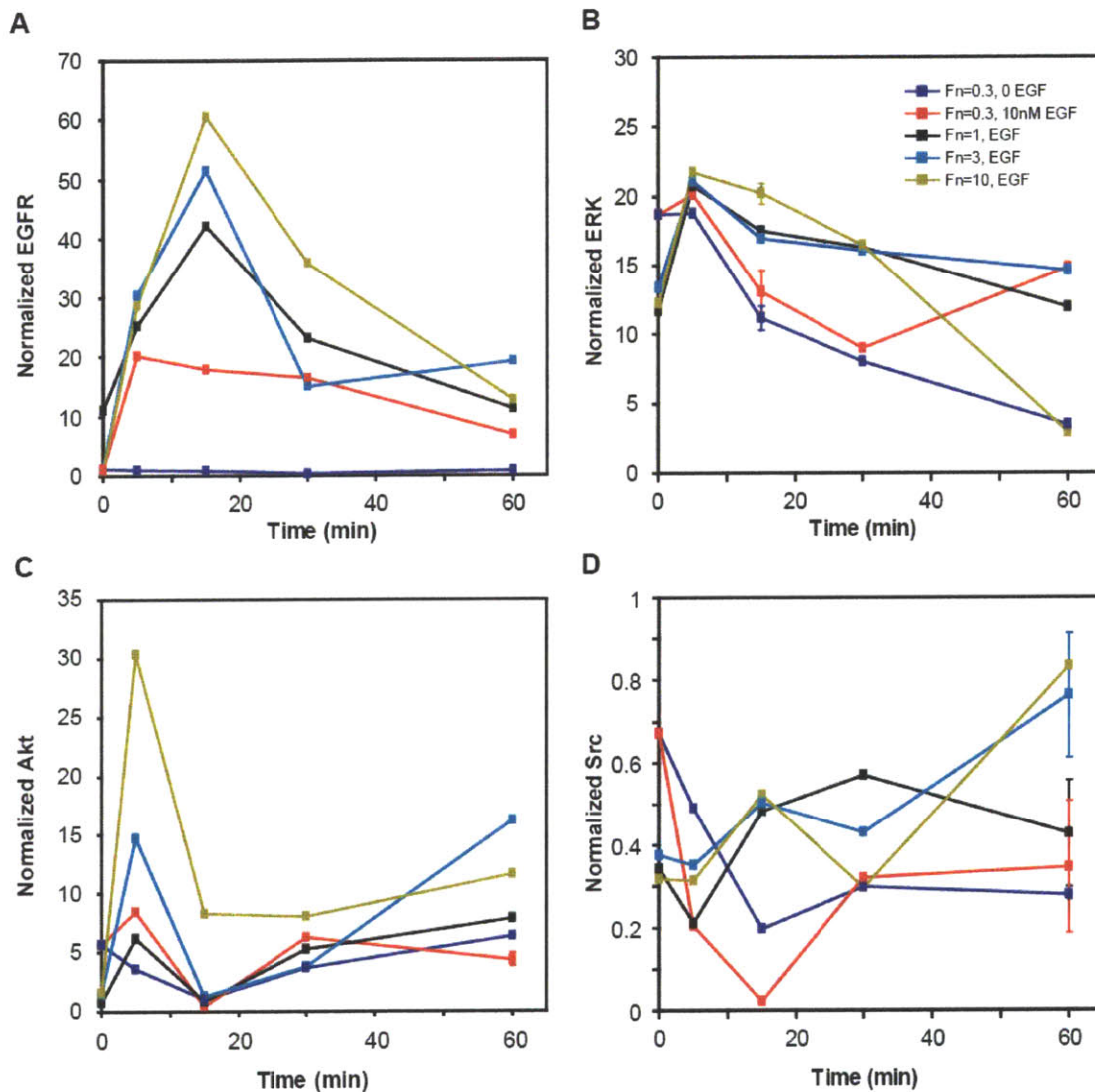
Figure 3.7. EGF-stimulated signaling of MSCs on fibronectin-adsorbed comb copolymer surfaces. Plots show time courses of normalized Bioplex measurements of EGFR (A), ERK (B), Akt (C), and Src (D). Cells were serum-starved for 16 hours and then treated with 10 nM of soluble EGF. Dark blue shows the control, no EGF treatment condition.

3.3.4. Quantification of MSC integrin proteins

The *in vivo* MSC microenvironment contains other extracellular matrix (ECM) proteins besides fibronectin that would also affect the adhesiveness of migration surfaces for MSCs, both in 2D and 3D [109]. However, whether these ECM proteins such as vitronectin and collagen actually play a role depends on whether MSCs have the specific integrin receptors associated with these proteins. Thus, we quantified the relative amounts of $\alpha v\beta 3$, $\alpha 5\beta 1$, and $\alpha 2\beta 1$ integrin complexes in MSCs as these are the major integrin complexes associated with vitronectin, fibronectin, and collagen binding [110].

Analyzing 10,000 cells per condition via flow cytometry, we found that MSCs have highest amounts of $\alpha v\beta 3$, followed by $\alpha 5\beta 1$, and finally $\alpha 2\beta 1$. All three were above control levels, with control being cells only labeled with the secondary antibody. This suggests that in addition to investigating the effects of fibronectin on MSC migration, we should study vitronectin and collagen as well. One shortcoming of this study however is that the results measure total amounts of integrin in each cell, but only those integrin complexes expressed on the cell surface would actually be able to engage ECM protein. Without spatial information, the relevance of higher levels of these complexes is not clear. For the purposes of this preliminary study, we assume that integrins in complex form are mostly present only on the cell surfaces. As well, during the fixation step, we minimize labeling internal integrin complexes by fixing the cells with methanol and not adding Tween to the wash solutions. Methanol fixation leaves cell membrane generally intact, as opposed to formaldehyde fixation. Tween is a chemical known for perforating membrane, which we leave out of our experimental protocol. One shortcoming of this approach is that we don't have an absolute number of integrin complexes, only the amount relative to a control.
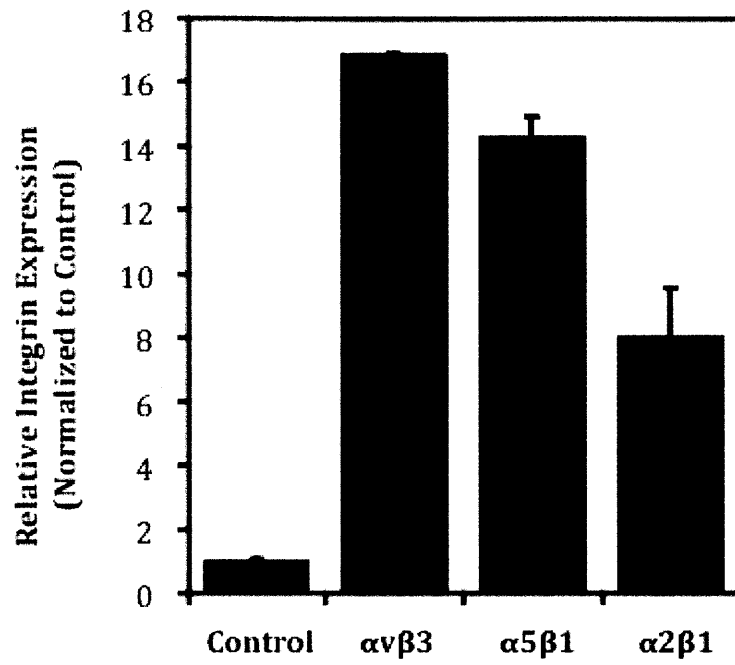
Figure 3.8. Quantification of MSC integrin complexes by flow cytometry. 10,000 cells were analyzed per condition. Error bars show +/- ECM.

3.4. Summary and conclusions

In this Chapter, we laid out the foundational work for the rest of this thesis to study signal-response relationships for MSC migration. Through soluble EGF treatments, we show that hTERT-MSCs migrate at robust speeds on PMMA-g-PEO co-polymer surfaces adsorbed with fibronectin, even following long periods of serum starvation. Migration speeds varied biphasically with surface adhesiveness, as modulated by different concentrations of fibronectin adsorption, whereas directional persistence was inverse biphasically affected by the same range of adhesiveness.

One of the advantages of the PMMA-g-PEO co-polymer is the advantage to tether small molecules such as EGF to the ends of the PEO sidechains, thereby immobilizing the growth factor and effectively taking it "out of solution." This immobilization affords the ability to localize and sustain growth factor signaling as there is presumably no dilution or cell internalization effects. However, a large unknown is how these tethered growth factors would differentially stimulate MSC signaling and how these signaling differences would ultimately affect MSC migration. We explore these questions in Chapter 4.

4. Chapter 4 - MSC migration response and signaling on tEGF biomaterials surfaces

4.1. Introduction

To rationally design controls for MSC behavior, we must understand the underlying

mechanisms that dictate MSC migration. Chapter 3 started to explore EGF effects on MSC migration

and showed that surface fibronectin adhesiveness was a more dominant cue than EGF. However,

we only investigated soluble EGF, and soluble growth factors can be limited in applicability *in vivo*

due to the effects of dilution and receptor internalization and degradation [111,112]. Both of these

challenges are ameliorated with immobilized growth factors, which have been proposed as

functionalizable units of tissue engineering scaffolds for therapy.

4.1.1. Immobilized growth factors for tissue engineering

Growth factor immobilization for bone therapy is not a new concept. Immobilized bone morphogenetic proteins have been shown to induced alkaline phosphatase activity and enhance bone formation [113-115], and immobilized TGFβ induced collagen synthesis [116,117]. In our own lab, tethering EGF (tEGF) to biomaterials surfaces provided a survival advantage to MSCs in the presence of proinflammatory cytokines and also increased cell attachment and spreading without driving MSCs down specific differentiation pathways [118]. Moreover, in the presence of osteogenic stimuli however, tEGF increased the efficiency and number of osteoid colonies, suggesting a productive application for tEGF in bone tissue engineering [119].
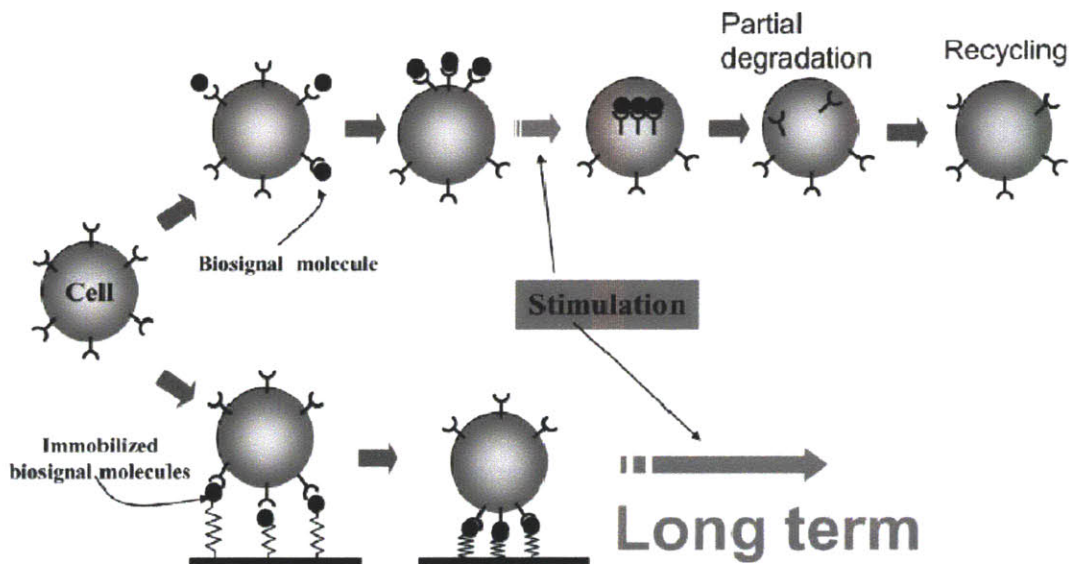
Figure 4.1. Differential receptor trafficking of soluble versus immobilized ligands. Tethered biosignal molecules are hypothesized to prevent and/or delay receptor internalization and degradation leading to more sustained signaling. *Source: Ito, 2008* [120].

Some of these effects are attributed to the differential receptor trafficking mediated by immobilized growth factors compared to their soluble counterparts [120]. Soluble ligands, once bound to their receptors on the cell surface, activate signaling pathways which in turn leads to receptor internalization and degradation. While some internalized receptors are recycled back to the cell surface, this receptor trafficking process is one of the cells' intrinsic negative feedback loops to prevent aberrant growth factor signaling [112]. Immobilized growth factors presumably prevent receptor internalization upon binding, and as a result, are hypothesized to sustain signaling longer than soluble growth factors [120]. Our lab's recent work supports this theory, showing that tEGF sustained EGFR signaling led to survival advantages for MSCs when antagonized as well as increased osteogenic differentiation when MSCs were presented with osteogenic stimuli [118,119].

Additionally, tethered growth factors may increase ligand-binding avidity given that binding of one immobilized growth factor to a cell-surface receptor would bring other receptors in close proximity to additional tethered growth factors on the substrate surface, facilitating further ligand-receptor interactions [121]. Because EGF receptors homodimerize, cross phosphorylate and then initiate downstream signaling, this increased avidity by tethered growth factors may increase signaling by increasing the likelihood of homodimer formations. This increase may be through the proximity of cell surface receptors to additional tEGF on the substrates, but could also be increased by the proximity of bound tEGF-EGFR complexes that facilitate homodimerization.

4.1.2. Chapter scope

The effects of receptor immobilization on MSC migration have not been studied, motivating the primary objectives of this thesis. In Chapter 4, we present hTERT-MSCs with tEGF and ECM cues to investigate their effects on cell migration speed and directional persistence. To start understanding the underlying mechanisms of cell migration, we also measure the phosphorylation levels of four protein signals involved in the EGF and/or integrin-signaling pathways.

71

4.2. Materials and methods

4.2.1. Cell culture and materials

hTERT-MSCs were cultured and maintained as described in Chapter 3. Cells were cultured at 37 degrees Celcius with 5% CO_2, split 1:10 when they reached 90% confluence (approximately 2 population doubles), and used until approximately passage 25.

4.2.2. Biomaterials surface preparations

4.2.2.1.   Polymer synthesis and polymer spin coating

PMMA-g-PEO Comb 1 co-polymer synthesis was conducted as described in Chapter 3. To allow for EGF tethering, we also synthesized a second comb co-polymer with its PEO sidechains activated with 4-nitrophenyl choloroformate. The two comb polymers differ their percentage by weight (wt%) of PEO and thus in the density of their comb PEO sidechains. Comb 1 contains 22 wt% PEO and allows for non-specific protein adsorption and thus is adherent for cells. Comb 2 contains 30 wt% PEO, does not enable protein adsorption and thus is cell-resistant. Comb 2's PEO sidechains are also activated with 4-nitrophenyl chloroformate (4-NPC) to present a reactive, hydrolysable leaving group for tethering EGF. We use Comb 1 and Comb 2 in a 60:40 ratio to achieve optimal tEGF presentation density while still allowing for protein and cell adhesion [101].

The comb polymers mixtures were diluted in toluene to a concentration of 20 mg/mL and put onto the surfaces of glass coverslips via thin film spin coating as described in Chapter 3. If not immediately used after spin coating, the coverslips were stored in the humidity-free vaccum oven for up to one month to minimize NPC hydrolysis.

4.2.2.2.　EGF-tethering and ECM adsorption

Murine EGF was tethered to the NPC-activated PEO sidechains via a primary amine. We chose murine EGF as opposed to human EGF because the only reactive amine is at the N-terminus, whereas human EGF contains a lysine residue mid-sequence with a reactive primary amine [122]. Thus, human EGF could potentially tether to the ends of PEO in two different orientations, whereas murine EGF theoretically only has one possible tethering site. Additionally, tethering at the terminal amine would present tEGF fully available for ligand interaction, producing as consistent of a substrate as possible. It should be noted that both human and murine EGF contain the other three amino acid residues besides lysine that also have primary amines in their R' groups: asparagine, glutamine, and arginine. However, these amino acids have bulkier R' groups than lysine, and there are also other groups nearby that reduce the primary amine's "willingness" to react. EGFR signaling experiments in hTERT-MSCs were conducted to evaluate the signaling capabilities of murine versus human EGF. Data from our lab showed that soluble versions of human and murine EGFs activated EGFR comparably.

The amine-NPC reaction was carried out at room temperature to tether murine EGF following previously described methods [101,118,123]. Stock murine EGF (Peprotech, Rocky Hill, NJ) was diluted to 25 mg/mL in 100 mM pH 9 phosphate buffer for tEGF surfaces, and plain 100 mM pH 9 phosphate buffer was used for the control surfaces. 150 uL of diluted EGF or phosphate buffer were dropped onto lab parafilm, and one glass coverslip was overturned onto each droplet, polymer-coated side down. The coverslips were incubated in a humidity chamber at room temperature for 20 hours, after which each coverslip was washed with 500 uL of 1x phosphate-buffered saline (PBS). Any remaining free NPC on the surfaces were blocked on 150 uL droplets of 100 mM pH 9 Tris buffer for two hours at room temperature. The surfaces were then washed three times with 500 uL of 1x PBS each time per surface and either used immediately or stored in PBS at 4°C for a maximum of two days.

Extracellular matrix proteins were adsorbed in varying concentrations to the comb co-polymer surfaces after tethering. Stock fibronectin (Sigma, St. Louis, MO), vitronectin (Millipore, Billerica, MA), and collagen (rat tail collagen I, Invitrogen, Carlsbad, CA) were diluted in 1x PBS to the desired concentrations of 0.3, 1, 3, and 10 ug/mL. The tethered surfaces were overturned onto 150 uL droplets of diluted ECM protein solutions and incubated at room temperature for 2 hours in a humidity chamber to allow for non-specific protein adsorption. Adsorbed surfaces were washed with 500 uL PBS each and overturned onto 150 uL droplets of 1% bovine serum albumin (BSA) (Sigma, St. Louis, MO) and blocked for 1 hour at room temperature. After washing three times with 500 uL PBS each time per coverslip, the surfaces were sterilized under ultraviolet (UV) light for 30 minutes while submerged in PBS and used immediately after sterilization.

4.2.3. Single-cell migration tracking and analysis

hTERT-MSCs between passages 8-20 were used for migration experiments to minimize cell-to-cell morphological changes resulting from high population doubling numbers. Prepared surfaces were glued to the bottom of 0.17 mm thick DeltaT dishes (Bioptechs, Butler, PA) as described in Chapter 3, and 10,000 cells seeded onto each 18mm surface (~4000 cells/cm^2) in full-serum media. The cells were allowed to attach, spread, and acclimate in 37°C, 5% CO$_2$ for 18 hours. A small amount, 5-10 mL, of identical media was simultaneously added to a T-flask and placed into the same incubator to be used as assay media. For EGFR inhibition experiments, 10 uM of AG1478, a specific EGFR kinase inhibitor, was also added to the seeding and assay media. We used full-serum media for all of our experiments because zero- and low-serum conditions seemed detrimental to MSC viability on these polymer surfaces.

After 18 hours, each DeltaT dish was prepped and sealed as described in Chapter 3 and placed onto a heated stage insert for a Ludl 99S008 motorized stage on a Zeiss Axiovert 35 microscope. 15 random fields of 5-10 cells each were selected, and an image taken of each field

every 10 minutes for 7 hours for a total of 43 images per field. The coverglass lid seals the dish to maintain the media pH at 7.4, and automated temperature control maintains the media temperature at 37°C for the duration of the 7-hour experiment.

Each set of 43 images for each field was strung together into a time-lapse movie and analyzed with Visible software (Reify Corporation, Cambridge, MA), exporting coordinate data as described in Chapter 3 and manually inspected to remove pixel-cluster segments not associated with singly migrating cells. To minimize skewing due to small sampling size, we select only those cells with more than 1 hour of continuous coordinates (7 time points). These coordinates were imported into Matlab and analyzed via the Persistent Random Walk model using overlapping time intervals (see Appendix for Matlab code).

4.2.4. Phosphoprotein quantification

4.2.4.1.   Lysate collection for signaling measurements

Because the co-polymer surfaces present tEGF, EGFR signaling can be activated as soon as cells settle onto the surfaces, muddling tEGF-induced EGFR signaling with cell attachment and spreading events. To separate signaling associated with attachment from tEGF-induced signals, MSCs were seeded in the presence of AG1478 (EMD Chemicals, Gibbstown, NJ), a potent and specific inhibitor of EGFR. Surface-prepped coverslips were deposited into 12-well plates and sterilized (see Chapter 3), and 100,000 hTERT-MSCs suspended in 1 mL assay media with 10 uM AG1478 were seeded onto each coverslip.

Cells were allowed to attach and spread for 5 hours in the presence of 10 uM AG1478 at 37°C, 5% CO_2, after which the cells are washed once with warm PBS and assay media added to each well. Cell lysates were collected 5, 15, 30, 60, and 180 minutes after inhibitor washout. A control condition without AG1478 washout was lysed to represent time point 0, and was lysed between the

30- and 60-minute steps. Lysate collection at each time point and protein determination was conducted as described in Chapter 3.

4.2.4.2. *Luminex xMAP platform to measure pEGFR, pAkt, pERK*

We used the Luminex xMAP system to measure levels of phosphorylated ERK (Thr202/Tyr204, Thr185/Tyr187), Akt (Ser473), and EGFR (Tyr), following Bio-Rad manufacturer's protocols and as briefly described in Chapter 3.

4.2.4.3. *ELISA for pFAK and total FAK quantification*

Phospho-FAK at Tyrosine 378 and total FAK levels were measured using commercial star enzyme-linked immunosorben assay (ELISA) kits (Millipore, Billerica, MA) following all manufacturer protocols. Briefly, 10 ug of cell lysate protein or predetermined amounts of pFAK standard were diluted in ELISA Diluent to a total volume of 50 uL, which was then loaded into each wells of a 96-well plate with the well bottoms pre-coated with capture antibody. 50 uL of phospho-FAK antibody or total FAK antibody diluted in antibody diluent buffer were also added to each well and the plate incubated at room temperature with constant agitation for three hours. The wells were washed four times each with 300 uL of 1x Wash Buffer each time, and then incubated with 100 uL of HRP-conjugated antibody against the phosphospecific antibody for 30 minutes at room temperature with constant agitation. After washing, 100 uL of TMB enzymatically activated detection reagent was added to each well and incubated at room temperature, protected from light, while monitoring the color development. After appropriate color development (usually 5-10 minutes), 100 uL of stop solution were added per well and the absorbance read at 450 nm. A standard curve was generated, and samples' absorbance compared to the standard to determine amount of phosphorylated FAK.

4.3. Results & Discussion

4.3.1. Persistent random walk modeling

As in Chapter 3, we mathematically fit for each cell's directional persistence via the

Persistent Random Walk model (PRW) using the cell's Speed and Mean Square Displacement for all

available time intervals. There are two ways to sample square displacements over all possible time

intervals lengths: overlapping versus non-overlapping. As an example, in the hypothetical cell path
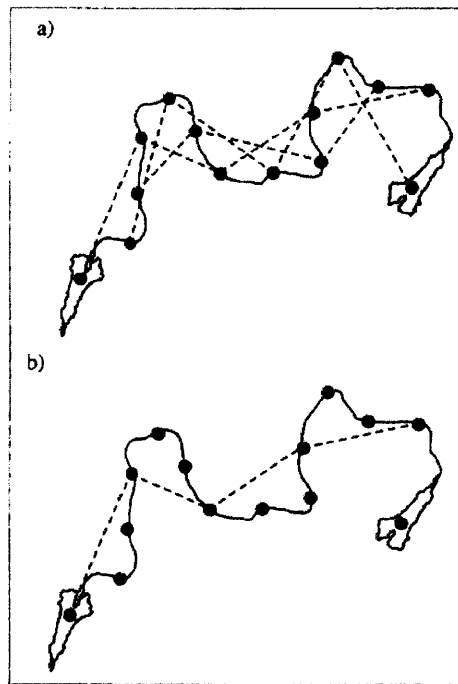
Figure 4.2. Hypothetical cell migration
path with 14 observed points and sampled
with overlapping time intervals (A) and
non-overlapping time intervals (B), both of
length three. *Source: Dickinson et al., 1993*
[102].

77

in Figure 4.2, 14 cell positions (n = 14) were recorded. There are then 13 (n – 1) possible values of T, time interval length for this cell. The dotted lines connect positions for when T equals three, or time interval lengths that are 3 unit-times. The overlapping and non-overlapping ways to sample square displacements for each time interval are shown in Figure 4.2A. There are 11 (n – 3) total overlapping intervals over which to average D^2 for $T = 3$. Figure 4.2B, on the other hand, shows 4 ($FLOOR \left[\frac{n}{3}\right]$) non-overlapping $T = 3$ intervals.

Overlapping time intervals maximize the number of square displacements being averaged, minimizing the skew of outliers, but samplings are not statistically independently. Non-overlapping intervals ensure sample independence, but suffer with small sample sizes. In fact, for time interval lengths where $T > \frac{n}{2}$, the sample size is only 1 with non-overlapping intervals, defined as from the start (coordinate 1) to coordinate number T+1. Overlapping intervals decrease in sample size with increasing T as well, but the sample size is only 1 when T equal n-1. To determine which sampling interval to use, we looked at several characteristic cell paths to compare the performances of overlapping versus non-overlapping time intervals.

For very persistently migrating cells (Figure 4.3), the two sampling methods produced similar values for Persistence Time (PT) with high r^2 values for each. For this particular cell, 41 coordinates were tracked, and the cell moved in a relatively straight line for the entirety of its migration path (Figure 4.3A). The dots in this figure represent a measured cell coordinate, and the dots were connected to visualize the cell path. The fitted parameter, persistence time (P) was 103 minutes using both non-overlapping (Figure 4.3B) and overlapping intervals (Figure 4.3C), and both r^2 values were in the mid-90s. In these two plots, "x" is the actual $<D^2>$ calculated for each time interval length, and the red line is the PRW model. Both sampling methods show good PRW model fit to data.
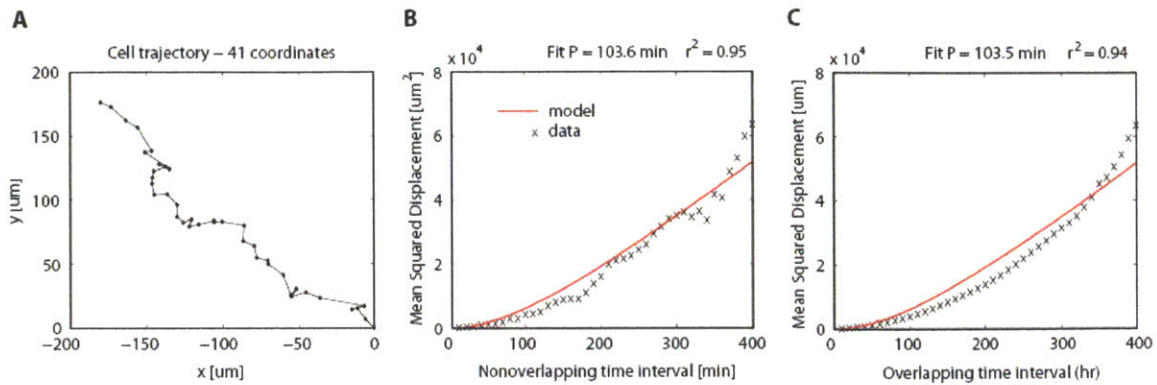
Figure 4.3. Persistent Random Walk model to fit Persistence Time for a cell with high directional persistence. (A) Raw cell path with 41 total recorded coordinates. (B) Model fit using non-overlapping time intervals. (C) Model fit using overlapping time intervals.

However, for cells that travel in less persistent paths, non-overlapping time intervals resulted in poor PRW model fit to data (Figure 4.5). For these more circular migration paths, indices that are large time interval lengths apart may actually be quite close to each other in linear distance, and thus the associated square displacement would be small. Because the non-overlapping interval sample size decays rapidly for increasing values of T, small square displacements at high values of T disproportionately affect the mean. For $T > \frac{n}{2}$, there is only one square displacement which is then itself the mean. This leads to wildly varying $<D^2>$ versus T curves (Figure 4.5B), instead of the monotonically increasing "ideal" curves seen with more persistent cells. Because PRW-predicted $<D^2>$ increases monotonically versus T, the PRW model is a poor fit for data from circularly migrating, and we see low r^2 values (Figure 4.5B). If we use overlapping intervals on the other hand (Figure 4.5C), these differences are better smoothed out with larger sample sizes of square displacements, particularly at high values of T. For this particular cell of 39 coordinates, r^2 is much improved, from 0.32 to 0.76. Using a different sampling method also led to a different fitted P, and PT decreased 15% from 12.3 minutes to 10.4 minutes with overlapping intervals versus non-overlapping intervals. Even with overlapping interval
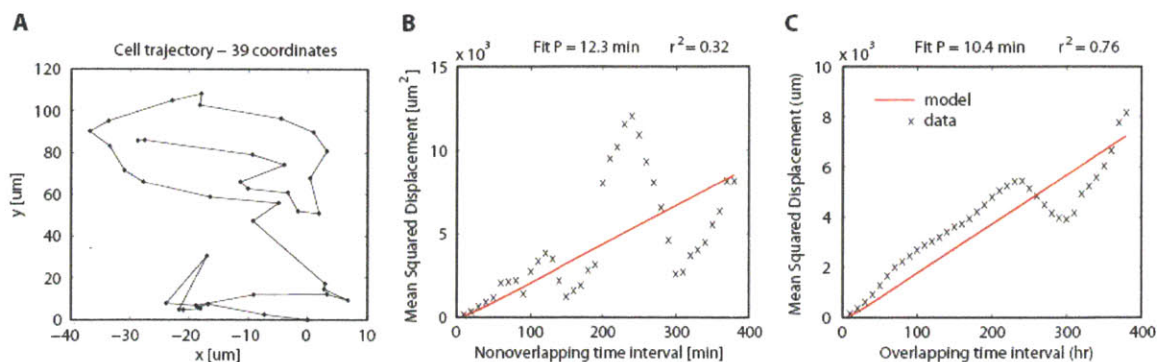
Figure 4.5. Persistent Random Walk model fit to fit Persistence Time for a cell migrating in a circular path and lower directional persistence. (A) Raw cell path with 39 total recorded coordinates. (B) Model fit using non-overlapping time intervals. (C) Model fit with overlapping intervals.
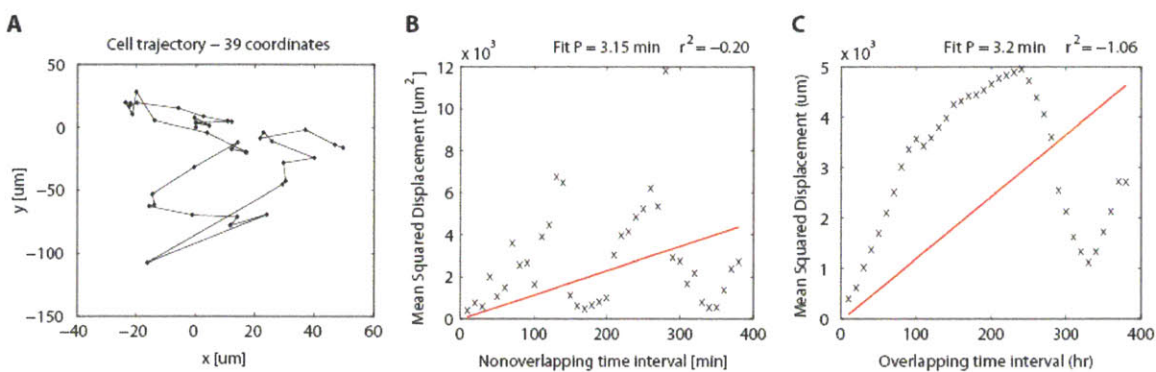
Figure 4.4. A cell with worse PRW model fit using overlapping intervals. (A) Raw cell path with 39 total recorded coordinates. (B) Model fit using non-overlapping time intervals. (C) Model fit with overlapping intervals.

sampling, sample sizes still decrease with increasing T, so we still see a dip in the measured $<D^2>$ for this circularly-migrating cell (Figure 4.5C), but the fit is overall much improved from the non-overlapping situation.

There are also instances where r^2 actually decreases significantly using overlapping intervals instead of non-overlapping ones (Figure 4.4). Because the residuals on overlapping intervals are correlated, if the general PRW model fit is extremely poor, using overlapping intervals

actually results in even greater systemic poor fit for all data points along the $<D^2>$ versus T curve.

This phenomenon makes sense when look at how we calculate the r^2 values.

The formula for r^2 is

$$r^2 = 1 - \frac{SS_{err}}{SS_{tot}},$$

where SS_{err} is the residual sum of squares defined as

$$SS_{err} = \sum_i (y_i - f_i)^2,$$

where y_i is the observed data, and f_i is the PRW model-calculated value, and SS_{tot} is the total

sum of squares defined as

$$SS_{tot} = \sum_i (y_i - \bar{y}_i)^2,$$

where y_i is the observed value, f_i the modeled value, and $\bar{y}_i$ the mean of all the observed

values. The calculated r^2 is negative when the ratio $\frac{SS_{err}}{SS_{tot}}$ is greater than 1, meaning that the

"variability" of the data from the model is greater than the "variability" of the data from its observed

mean. In the cases where overlapping interval sampling decreases r^2 is the model fit is so poor such

that the correlated residuals result in a very large SS_{err}.

To evaluate the general performance of non-overlapping versus overlapping intervals, we

examining the PRW model fit for a set of 55 cells. Figure 4.6A plots the change in r^2 using

overlapping intervals versus non-overlapping intervals for each of the 55 cells. Only 9 cells (16%)

showed a decrease in the "goodness" of the PRW model fit. Of these, only 3 cells, highlighted in red,

showed r^2 decreases of more than 0.2. These cells with large r^2 decreases exemplified the situation

described above where there is poor systemic fit of model to data, and overlapping intervals

magnified that poor fit due to correlated residuals.
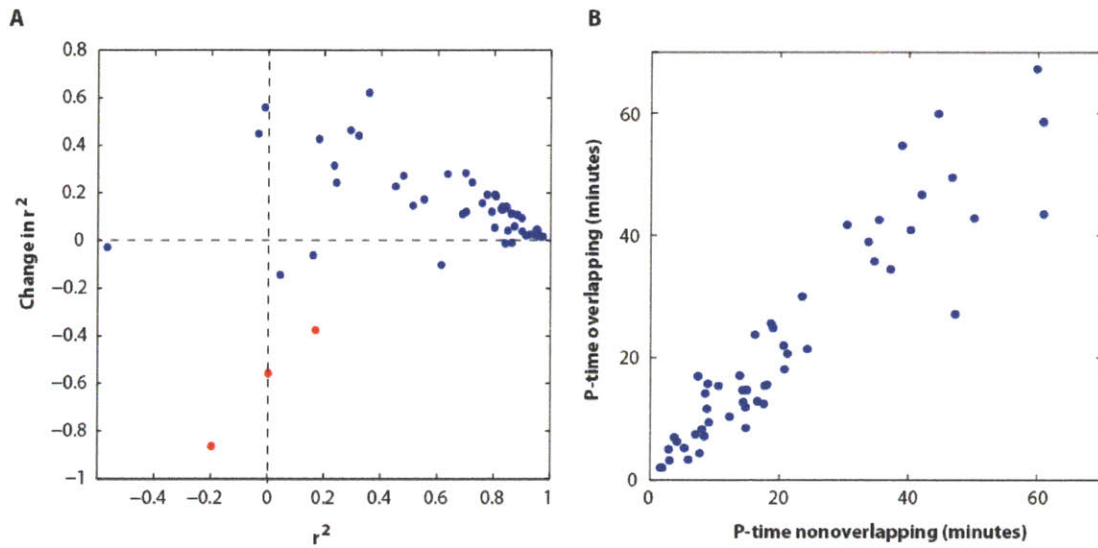
Figure 4.6. Comparison of model fit and persistence time between overlapping versus non-overlapping intervals for 55 cells. (A) The change of r^2 versus original r^2 of non-overlapping

We also looked at how persistence time values themselves might differ between using overlapping versus non-overlapping intervals (Figure 4.6B). Plotting overlapping P_O versus non-overlapping P_{NO}, the data points do line up along the diagonal indicating that similar trends in persistence time are observed from cell to cell using the two methods, though P_O and P_{NO} are not usually equal. The average change in P from non-overlapping to overlapping was 7.6 minutes (25.5% of P_{NO}). Plotting P_O and P_{NO} versus r^2 (Figure 4.7), we see a right shift from the blue data points (P_{NO}) to the red data points (P_O) indicating that as a population, these cells' PT are largely unchanged despite better r^2 values across the board. Given these characterizations, we prioritized PRW model fit and analyzed all migration paths using overlapping intervals.
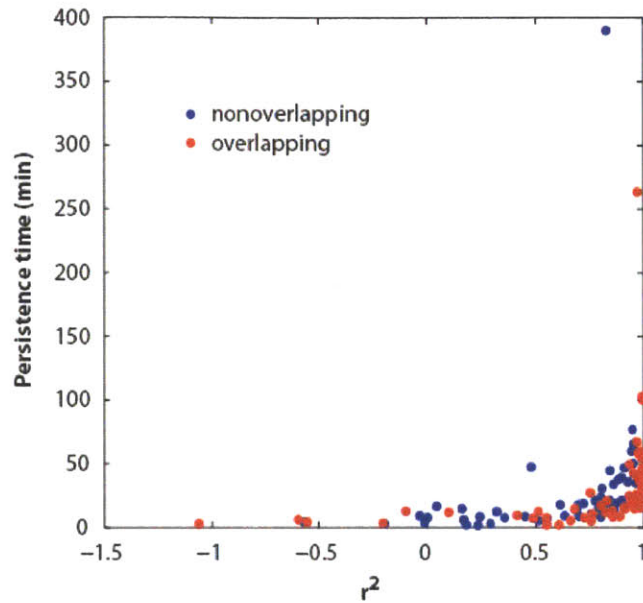
Figure 4.7. Calculated persistence time versus model fit, comparing non-overlapping and overlapping intervals.

4.3.2. Migration response on fibronectin-adsorbed surfaces

We carried out all of our tEGF experiments in full-serum media (10% FBS) because preliminary experiments showed that hTERT-MSC attachment to the PMMA-g-PEO surfaces was not consistent in low-serum media containing 0.5% FBS. Because of our previous experience with fibronectin effects on fibroblast migration, we extensively investigated MSC migration on fibronectin (Fn) adsorbed co-polymer surfaces. Migration speed versus surface fibronectin adhesiveness was biphasic on both tEGF and control surfaces (Figure 4.8). We saw peak speeds on the intermediate 3 ug/mL Fn (Fn 3) surfaces for both tEGF and control surfaces, at 63.8 um/hour and 61.7 um/hour respectively. This peak was significantly higher than all other fibronectin adhesiveness on tEGF surfaces and significant from Fn 1 and Fn 10 on the control surfaces. On 1, 3, and 10 ug/mL fibronectin, the average speeds on tEGF surfaces were higher than those on the

control surfaces, though these differences were not statistically significant. For each condition, the

speed is an average of 100 to 200 individual cell speeds across 2-3 biological replicates.
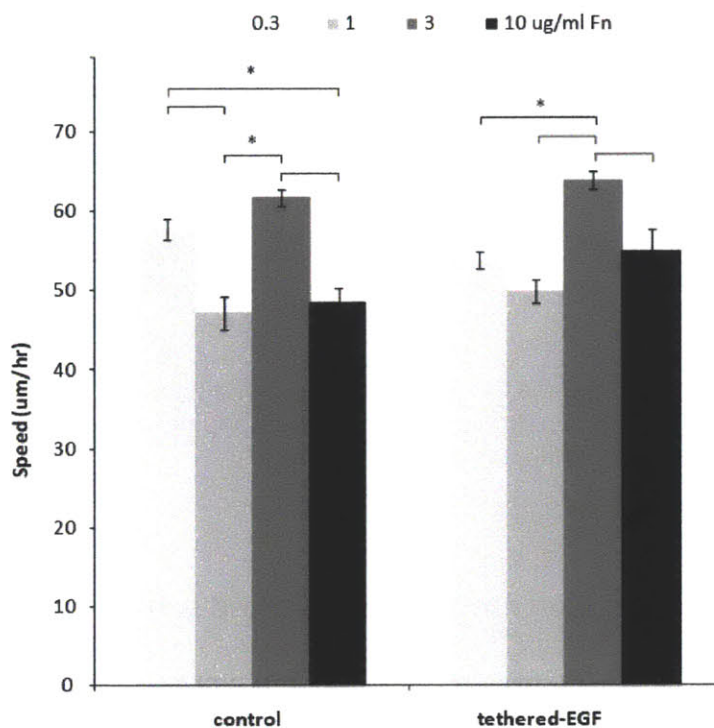
Figure 4.8. MSC migration speed on tEGF and control fibronectin-adsorbed copolymer surfaces. Error bars show +/- SEM. * denotes statistical significance between bracketed conditions ($p < 0.01$).

Fitting for cell persistence via the Persistent Random Walk model, differences between

control and tEGF surfaces emerge (Figure 4.9). On the control surfaces, persistence varied inverse

biphasically with surface adhesiveness, whereas the trend seemed to be biphasic like migration

speed on the tEGF surfaces. The peak persistence, however, was at 1 ug/mL Fn as opposed to the

speed peak at 3 ug/mL, though this peak was not statistically significant. The variance of

persistence was much higher than that of speed, however, and only the persistence time differences on control surfaces were statistically significant from each other. The persistence time on 10 ug/mL of fibronectin was higher than that on 0.3 and 1 ug/mL.
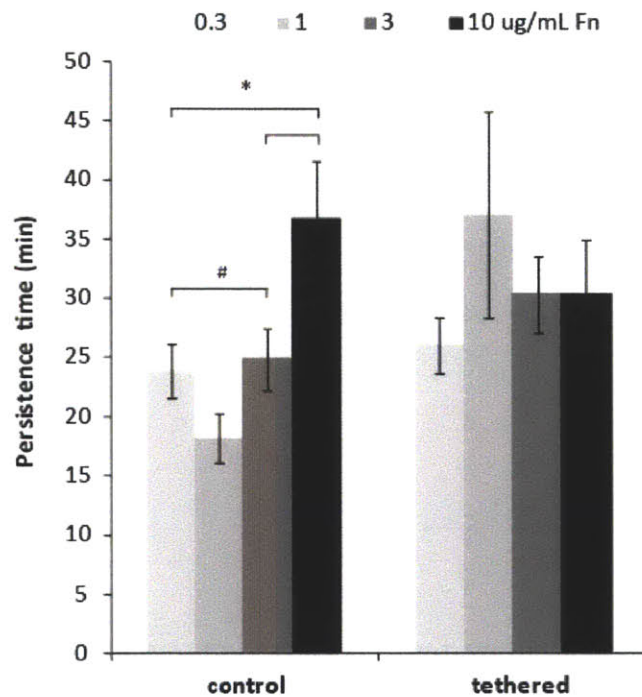
Figure 4.9. MSC migration persistence time versus fibronectin adsorption concentration, on control and tethered-EGF surfaces. Error bars show +/- SEM. Statistical significance between bracketed conditions is denoted by * (p < 0.01) and # (p < 0.05).

To determine whether this high variance in persistence time condition is due to widely varying shapes of cell migration paths, or if it is actually a side effect of fitting for persistence time using the PRW model, we directly calculated the "directionality" of the cells from their migration paths' raw coordinates. We defined the parameter Directional Index (DI) as the ratio of the cell's total migration path length over its total displacement. The lower bound of DI is one, achieved with

a perfectly persistent cell that migrates in a straight line and thus having exactly equal values for its total path length and displacement. Above one, the higher the cell's DI, the more "meandering" of a migration path it has taken, and thus the more directional changes it has undergone. Relating back to Persistence Time from the PRW model, DI and PT should be inversely related.

We plotted average cell DI versus fibronectin adsorption concentration on control and tEGF surfaces (Figure 4.10A), and showed that DI decreases with increasing fibronectin on both control and tethered surfaces. This trend is consistent with PRW persistence time increasing on control surfaces, but does not agree with the biphasic relationship seen on tEGF surfaces (Figure 4.9). The variance of DI was similar to that of persistence time though, indicating that cell directionality is indeed a much more heterogeneous parameter than speed. Interestingly, though persistence time
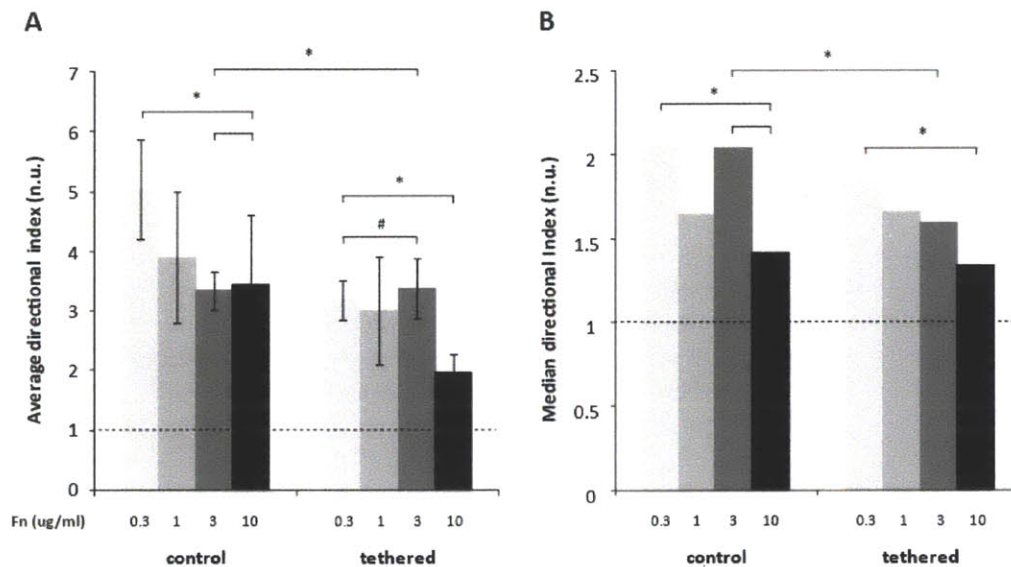
Figure 4.10. Directional index of MSC migration on fibronectin, control and tEGF surfaces. (A) Average DI. (B) Median DI. Error bars are +/- SEM. Statistical significance between bracketed conditions is denoted with * ($p < 0.01$) and # ($p < 0.05$). Dotted lines indicate DI's lower bound of one.

between the different fibronectin adhesiveness on tEGF surfaces were not statistically significant, the differences in cell DI were. Specifically, cell DI on 10 ug/mL Fn with tEGF was significantly lower than Fn 0.3. Furthermore, tEGF decreased average DI on all four fibronectin surfaces, though only the differences on Fn 3 were statistically significant. This overall effect was also seen in the PTs of Figure 4.9 with the exception of Fn 10, though none of the differences were statistically significant.

Because the variance of DI was so large in all of the conditions, we also plotted the median DI for each condition. For the most part, the medians trend similarly to the averages, with the exception of Fn 3 control (Figure 4.10B). The average DI on Fn 3 control was lower than both Fn 1 and Fn 10 control surfaces, but its median was higher. The overall similarities between the median and average DI trends suggest that although the standard deviations are large for the average DIs for each conditions, these averages are still telling and representative of the differences in cell migration persistence response.

We then compared the differences in cell migration DI versus PT from the PRW model to evaluate model appropriateness. On control surfaces, median and average DI both showed that cell persistence increases with increasing Fn, and PT essentially showed the same trend though there were some disagreements for Fn 3 (Figure 4.9, Figure 4.10). On the tethered surfaces, median and average DIs decreased with increasing fibronectin adhesiveness, showing the same trend of increasing cell persistence. However, PT on tEGF surfaces is biphasic, peaking at Fn 1, though the error bars are very large for Fn 1. Fn 3 and Fn 10 do have higher PTs than Fn 0.3, so potentially PT also shows the same trend of increasing persistence with increasing surface adhesiveness.

Hypothesizing that extreme outliers are skewing the population averages, especially those of PT, we plotted the averages of the 25-75th percentiles of cell PT for each condition as well as the medians. Indeed, plotting both the median and 25-75th percentiles showed that cell persistence generally increases with increasing surface adhesiveness (Figure 4.11). Despite error bars being

much smaller in the 25-75th percentiles plot compared to the full averages because we are excluding the outlier "tails," differences on the tEGF surfaces were still not statistically significant (Figure 4.11B).
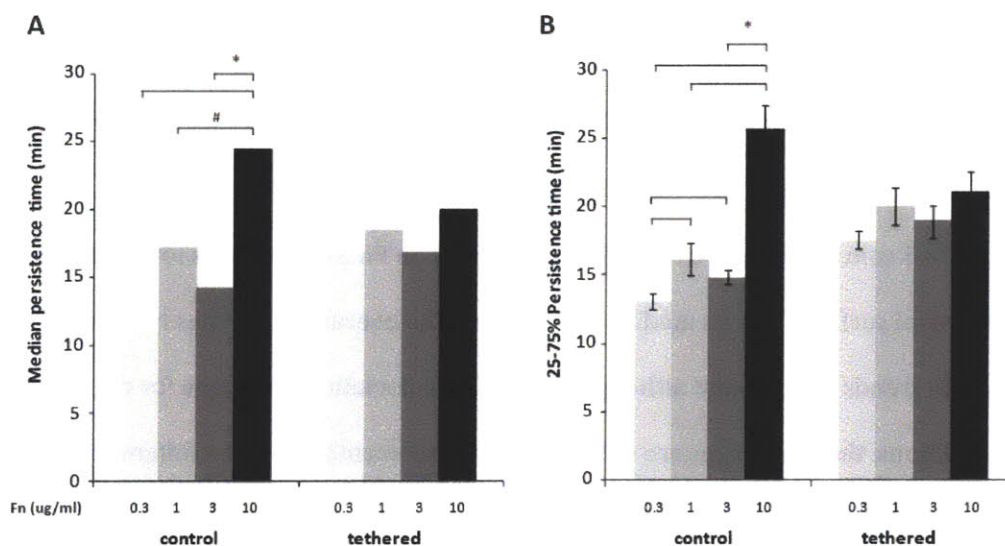
Figure 4.11. MSC migration directionality on fibronectin surfaces with and without tEGF. Data shown are median persistence time (A) and the 25-75 percentile of average persistence time. Error bars represent +/- SEM. Statistical significance between bracketed conditions is denoted by * ($p < 0.01$) and # ($p < 0.05$).

To investigate the effects of EGFR phosphorylation, we tracked cell migration while inhibiting EGFR activity with 10 uM AG1478. Because migration speed peaked on Fn 3 on both control and tEGF surfaces (Figure 4.8), we tracked cells seeded on 3 ug/mL with and without tEGF in the presence of 10 uM AG1478. The EGFR inhibitor AG1478 significantly decreased cell speed on both control and tEGF surfaces (Figure 4.12) from 61.7 um/hour to 51.9 um/hour (15.9%) on the control surfaces, and from 63.6 um/hour to 54.3 um/hour (14.6%) on the tethered surfaces. The speeds with inhibitor were still higher on the tethered surfaces than control, though this difference was not statistically significant.

We also looked at affects of AG1478 on cell migration median DI and the average 25-75th percentile persistence time. Tethered EGF significantly increased cell Persistence Time from control conditions (Figure 4.13), but the EGFR inhibitor significantly decreased cell persistence on both control and tEGF conditions. The decrease in persistence was reflected in both the PRW persistence time and in the median cell DI (Figure 4.13).
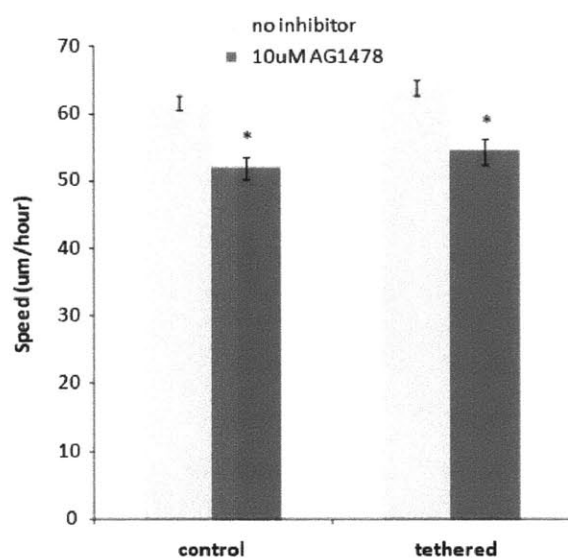
Figure 4.12. Effects of AG1478, an EGFR inhibitor, on MSC migration speed on 3 ug/mL fibronectin. Error bars show +/- SEM. * denotes statistical significance from the no inhibitor condition ($p < 0.01$).

The fact that our EGFR inhibitor would decrease speed and persistence on tEGF and control surfaces is baffling since there is no direct EGF stimulation on control surfaces. Because we run these experiments in full serum conditions, serum growth factors could be influencing cell behavior. Furthermore, this could be due to autocrine EGF signaling whereby the cell secretes its own EGF which then re-engages growth factor receptors to activate signaling. Autocrine signaling

can help sustain signaling downstream of growth factors and further mediate cell migration[124]

There is evidence of autocrine FGF and TGFβ, as well as autocrine EGF in primary MSCs, making this

a plausible hypothesis [119,125,126].
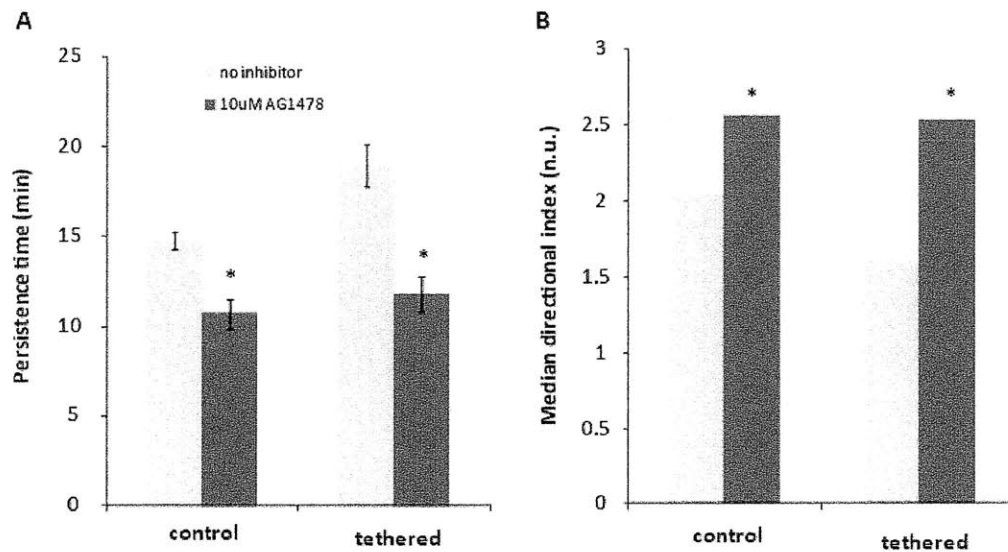
Figure 4.13. Effects of AG1478, an EGFR inhibitor, on MSC migration directionality on 3 ug/mL fibronectin surfaces with and without tEGF. Data shown are the 25-75 percentiles of average persistence time (A) and median directional index (B). Error bars show +/- SEM. * denotes statistical significance from the no inhibitor condition ($p < 0.01$).

4.3.3. Effects of extracellular matrix proteins on migration response

In vivo, the MSC environment includes a variety of extracellular matrix proteins besides

fibronectin, such as vitronectin and collagen. Thus, we also tracked hTERT-MSC migration on

vitronectin- and collagen-adsorbed tEGF surfaces. Speed on vitronectin-adsorbed surfaces was

biphasic versus surface adhesiveness, but collagen-adsorbed surfaces showed no obvious trend

(Figure 4.14). Speed on vitronectin-adsorbed surfaces was significantly higher at 1 ug/mL Vn, the

peak of the Vn biphasic curve. This is in comparison to the fibronectin curve peaking at 3 ug/mL of

ECM. The average cell speed on Vn 3 was 73.4 um/hr, compared to the next fastest condition of Vn

0.3 with average cell speed of 50.7 ug/hr. Migration speed on collagen, on the other hand, was
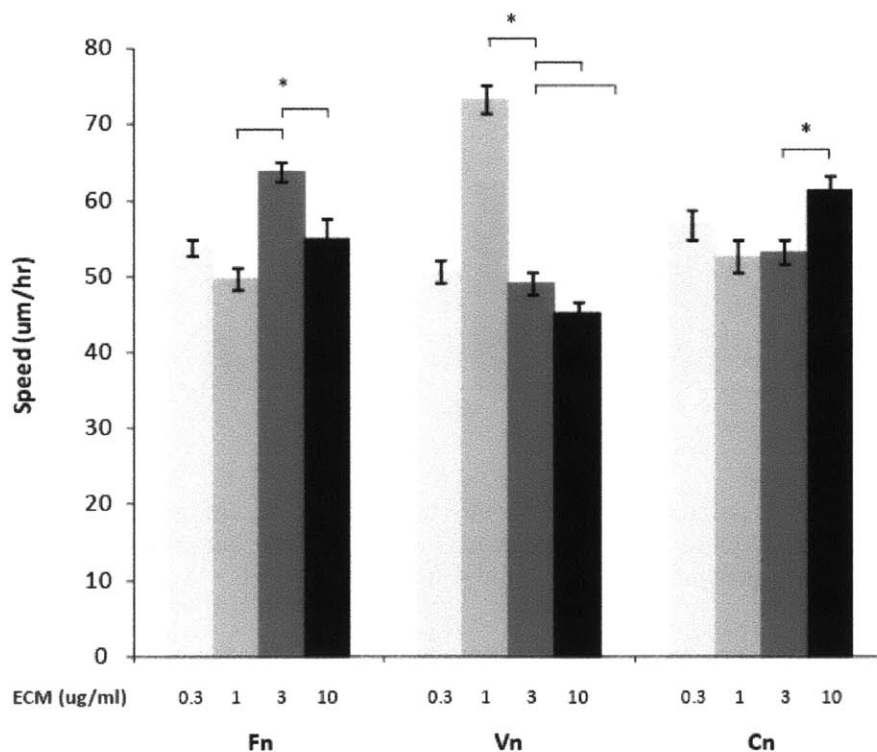
Figure 4.14. Effects of ECM proteins fibronectin, vitronectin, and collagen on MSC migration speed. Error bars show +/- SEM. * denotes statistical significance between bracketed conditions (p < 0.01).

significantly higher at 10 ug/mL with an average cell speed of 61.5 um/hr. Though the differences were not statistically significant, average cell speed was lowest on Cn 1 and Cn 3, and highest on Cn 0.3 and Cn 10. This could be either an inverse biphasic curve of speed versus collagen adhesiveness, or perhaps Cn10 is the peak of a biphasic curve that extends to higher unmeasured conditions of collagen, such as 20 ug/mL collagen. Overall, cells migrated at more consistent speeds across the four adhesiveness of collagen than across the same adhesiveness of fibronectin and vitronectin.
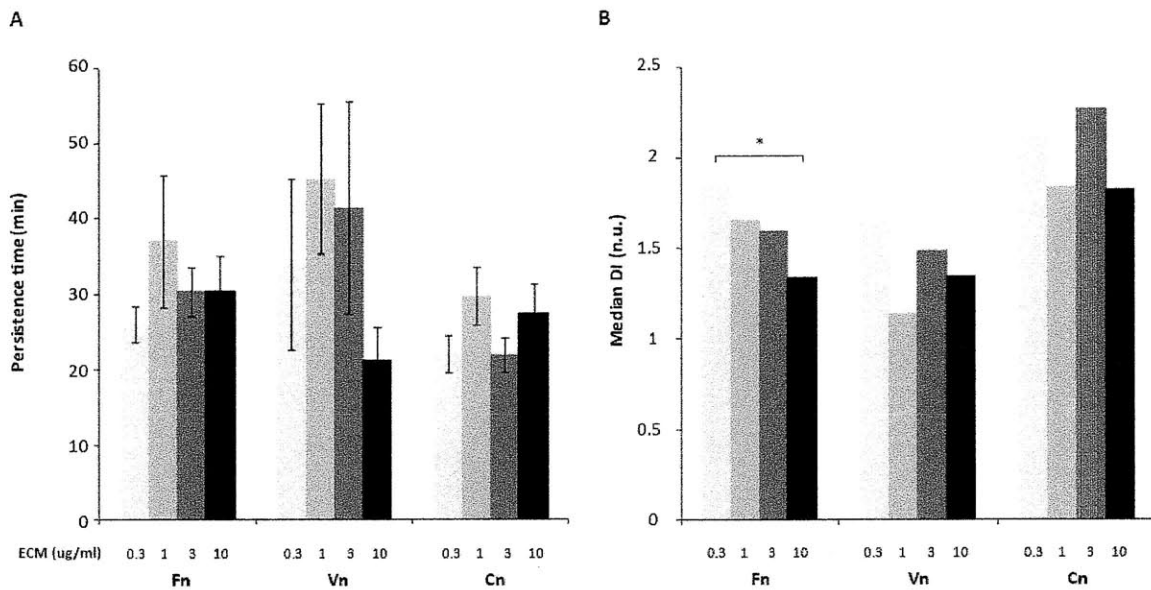
Figure 4.15. Effects of ECM proteins fibronectin, vitronectin, and collagen on MSC migration persistence time (A) and median directional index (B). Error bars show +/- SEM. * denotes statistical significance between bracketed conditions (p< 0.01).

Persistence versus surface adhesiveness for the various extracellular matrix proteins showed trends similar to speed (Figure 4.15). On fibronectin and vitronectin surfaces, persistence varied biphasically to surface adhesiveness, with the highest average persistence times at

intermediate levels of fibronectin and vitronectin (Fn 1 and Vn 1) (Figure 4.15A). On collagen surfaces, persistence time was highest on Cn 1 and Cn 10 surfaces, and lower on Cn 0.3 and Cn 3, though the differences were not statistically significant. Vitronectin conditions produced the most widely varying values of persistence time: both the highest and lowest overall persistence times across all the conditions were on vitronectin surfaces (Vn 1 was highest, and Vn 10 was lowest (Figure 4.15A). Collagen surfaces, as a whole, led to the lowest persistence times compared to fibronectin and vitronectin. The average persistence time across all four collagen surfaces was 25.3 minutes, compared to 31.0 minutes average on fibronectin and 35.5 minutes on vitronectin. The high variance in persistence on vitronectin surfaces may be due to our having tracked fewer cells on vitronectin surfaces ($n \approx 40$) than fibronectin ($n \approx 150$) and collagen ($n \approx 100$) cells.

These general trends were observed in the median DI values as well (Figure 4.15B). We have already addressed the different parameters of persistence on fibronectin surfaces. For vitronectin, the average median DI across all four conditions was the lowest of the three surfaces at 1.41, indicating that cells were the most persistent on vitronectin surfaces. Average median DI was lowest on collagen surfaces (2.02), and fibronectin was in between the two (1.62).

4.3.4. Mean free path as a comprehensive measure of cell migration

We have now examined two separate parameters for cell migration: speed and persistence. We have looked at persistence using Persistence Time, a fitted parameter from the Persistent Random Walk model and using the cell's Directional Index (DI), a parameter calculated directly from each cells' migration path coordinates. Of the two, we chose to use the PRW persistence time as the parameter for cell persistence.

Both cell speed and persistence affects the ultimate migration behavior of cells [127]. In our application of using MSCs for tissue engineering applications, the cells should optimally fill out a synthetic scaffold as efficiently as possible, which requires the cells to migration fast and

persistently. A fast cell with low persistence would migrate in a meandering path, whereas a slow

cell regardless of its persistence would take a long time to reach all the areas of a synthetic scaffold.

Because both cell speed and persistence are important, we examined a parameter that encapsulates

both.

Cell Mean Free Path (MFP) is calculated by multiplying a cell's speed with its persistence

time. The resulting MFP is a measure of distance, the length of each "free path" segment, i.e. the

average distance the cell travels before changing direction. This directly relates to how much

"ground" a cell is able to cover: the longer the cell's MFP, the more far-reaching its migration path

would be, determined by a combination of the cell's speed and persistence. With this parameter,

we can now compare the "effectiveness" of cells' migration paths across conditions taking into

consideration both the cells' speed and persistence.
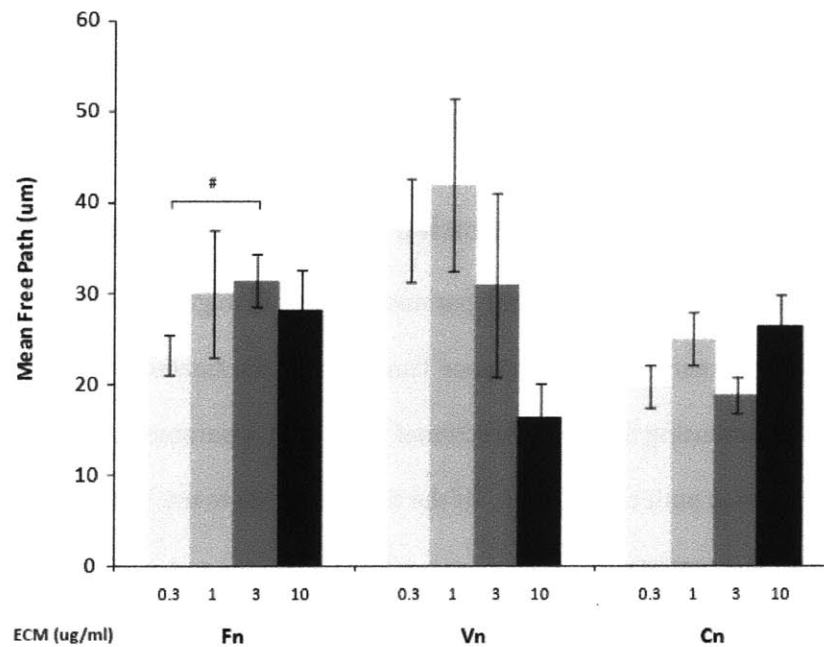
Figure 4.16. Effects of ECM proteins fibronectin, vitronectin, and collagen on MSC Mean Free Path (MFP), representative of a cell's ability to "cover ground." Error bars show +/- SEM. # denotes statistical significance between bracketed conditions (p < 0.05).

94

We calculated MFPs for each of the ECM conditions on tEGF surfaces. For each cell per condition, we multiplied its speed and persistence to determine the cell's MFP. Then for each condition, all of the individual cell MFPs were averaged to calculate the mean value. This method more accurately captures cell-to-cell variability, as opposed to multiplying the bulk averages of speed and persistence together. As seen in Figure 4.16, the highest variability of MFP was observed on vitronectin surfaces, just like cell persistence. Vn 3 had the highest MFP across all conditions, and Vn 10 had the lowest MFP across all conditions. Overall, the average MFP across all four vitronectin conditions was 29.3 um, which was still the highest of the three ECM surfaces. Collagen surfaces had the lowest average MFP at 23.2 um, and fibronectin was in between with an average of 28.1 um.

4.3.5. Cues for signaling measurements

We used two kinds of cues to affect cell migration response: tEGF and ECM proteins. The latter was achieved through non-specific adsorption of three ECM proteins—fibronectin, vitronectin, and collagen—at four concentrations spanning the range of low to high surface adhesiveness: 0.3, 1, 3, and 10 ug/mL ECM protein. The ECM proteins may influence migration behavior through two avenues: integrin-mediated signaling, and integrin-mediated adhesivness. Our underlying premise is that we can account for both effects, along with those from tEGF, by characterizing signaling activities. We thus presented these cues to cells to measure cell signaling so that we can ultimately correlate signaling with migration responses to determine how signaling affects response and the signals that govern MSC migration.

Table 4.1 displays the matrix of cue combinations for which we measured MSC phosphoprotein signaling with three cues: tEGF, ECM protein, and EGFR inhibitor. For fibronectin and collagen surfaces, we measured signaling on both tEGF and control substrates at all four concentrations (0.3, 1, 3, 10 ug/mL). For vitronectin surfaces however, we measured signaling on

95

tEGF surfaces for all four adhesiveness but only measured the 1 ug/mL vitronectin adhesiveness on

the control surface. Additionally, because the EGFR inhibitor AG1478 affected cell migration

response on both control and tEGF surfaces, we also measured signaling in the presence of this

inhibitor (10 uM AG1478). The Luminex xMap platform provides a high throughput assay to

measure activity of multiple phosphoproteins.

	Control	Tethered-EGF	tEGF+ AG1478
Fn0.3	•	•	•
Fn1	•	•	•
Fn3	•	•	•
Fn10	•	•	•
Vn0.3		•	
Vn1	•	•	
Vn3		•	
Vn10		•	
Cn0.3	•	•	
Cn1	•	•	
Cn3	•	•	
Cn10	•	•	

Table 4.1. Conditions matrix for MSC signaling. AG1478 is a
specific and potent EGFR inhibitor.

96

4.3.6. Phosphoprotein signaling on biomaterials surfaces

4.3.6.1.   *Data normalization to compare results between all surfaces*

We measured the phosphorylation levels of four phosphoproteins: EGFR, ERK, Akt, and FAK at 0, 5, 15, 30, 60, and 180 minutes after tEGF stimulation. Our original set of proteins to measure was much larger (Figure 4.17), including others also reported to be important in cell migration such as ILK-1 (integrin-linked kinase-1), Src, paxillin, MLC (myosin light chain), and ROCK (Rho-
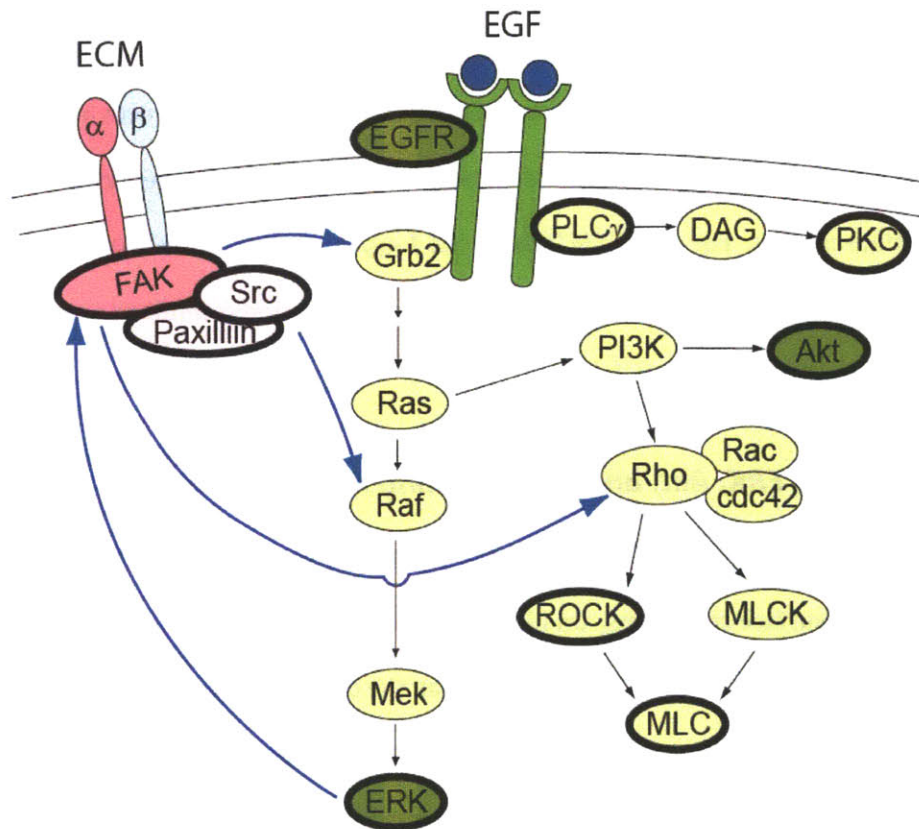
Figure 4.17. Signaling pathways downstream of EGF and ECM that affect migration. Proteins with bolded outline are ones we attempted to measure. Proteins shaded with a darker color are the ones we successfully detected and/or measured discernible differences between treatment conditions.

associated kinase) [128-132]. Most of these proteins were however either undetectable or differences between conditions indiscernible, even with soluble-EGF stimulation, and thus were not included in the final signaling results. The signaling time courses for the EGFR, ERK, Akt, and FAK are shown in Figure 4.18, Figure 4.19, Figure 4.20, and Figure 4.21 respectively. In all of these figures, the tEGF conditions are solid lines with circular time points, and the control conditions are dotted lines with square time points. Color intensity increases with increasing adsorbed ECM concentration.

The data shown are the averages of two or three biological replicates. For each replicate, the raw measured fluorescence for each condition was normalized to the fluorescence at 1 ug/mL ECM on control surfaces, time 0 minutes, to obtain a fold-difference in phosphorylation across the entire experiment for that condition and phosphoprotein. These fold differences are then comparable across experiments and were averaged across three biological replicates to produce the signaling time courses. For each ECM "block," the 1 ug/mL control surface condition is 1 at time 0 minutes, termed "basal" from here on on. All other time points and conditions are represented as a fold-difference over the basal signal.

4.3.6.2.   EGFR, ERK, Akt, and FAK signaling time courses

In Figure 4.18, we show EGFR phosphorylation at 0, 5, 15, 30, 60, and 180 minutes after treatment on all 29 surface conditions. On fibronectin-adsorbed surfaces (Figure 4.18A), EGFR phosphorylation on tEGF surfaces were all higher than the control surfaces. Peak activation on tEGF surfaces occurred 30 minutes after tEGF treatment on Fn 3 and Fn 10 surfaces to a max of about 5.5 times that of the baseline control (Fn 1, control surface, 0 minutes). Fn 0.3 and Fn 1 surfaces, however, actually show a decrease in phosphorylated EGFR after 30 minutes. This decrease in signaling recovers to levels comparable with Fn 3 and Fn 10 after 60 minutes, and this is the signaling peak for Fn 0.3 and Fn 1 at 4-4.5 times the control signaling. All of the non-tEGF

control surfaces, on the other hand, show minimal EGFR signaling. Aside from a doubling of pEGFR on the Fn 3 surfaces one hour after treatment, EGFR phosphorylation levels all stayed at approximately the same low basal level. This EGFR phosphorylation on a control surface may be evidence of the previously hypothesized autocrine signaling.

Another observation from this representative time course is the steady state EGFR signaling level on tEGF versus control surfaces. Even at time 0 in the presence of 10 uM AG1478, EGFR signaling was consistently higher on the tEGF surfaces, about 2.5-3 times that of the control. After acute EGFR activation by tEGF, cells on tEGF surfaces sustained signaling levels approximately 2.5 times that of control even three hours after treatment suggesting that EGFR signaling is basally high on tEGF surfaces.

Figure 4.18B shows EGFR signaling on the same fibronectin-adsorbed surfaces in the presence of 10 uM of the EGFR inhibitor AG1478. Peak EGFR signaling was only 2.5 times that of the control, observed on the Fn 10 tEGF surface 60 minutes after treatment. This is a significant inhibition from the 5.5 times control peak that we observed without AG1478 and also a 30 minute delay from the no inhibitor case (Figure 4.18A). Near-basal EGFR signaling was observed on Fn 0.3 and 3 surfaces, and Fn 1 showed a peak after 30 minutes of twice that of the control. Sustained signaling on tEGF surfaces, observed at both time 0 and three hours, was about 1.5 times that of the control, lower than the 2.5- to 3-fold difference over control seen without the inhibitor.

EGFR signaling on collagen was not transiently stimulated after treatment (Figure 4.18C), showing tEGF surfaces having 2 to 3.5-fold differences over the control at time zero, but this increased signal decreased significantly after 30 minutes. By three hours, the signaling on all surfaces were at similar and at near-basal levels. Signaling on vitronectin tEGF surfaces resembled the AG1478 inhibitor case with major peaks 60 minutes after treatment (Figure 4.18D). We only measured signaling on 1 ug/mL vitronectin adsorbed onto control surfaces. EGFR signaling in this case stayed at basal levels throughout the time course.

99

Figure 4.19 shows ERK phosphorylation on fibronectin (Figure 4.19A), fibronectin with AG1478 (Figure 4.19B), collagen (Figure 4.19C), and vitronectin (Figure 4.19D) surfaces. Phospho-ERK peaked 5 minutes after signaling on all of the surfaces, and the highest fold difference peak over basal conditions was observed on 3 ug/mL collagen with tEGF. Overall across all conditions, the highest fold-differences were seen on collagen surfaces with peaks 8-10 times that of basal levels 5 minutes after treatment followed by sharp decreases in all conditions at 15 minutes. Interesting, there was a late, more gradual signaling of phospho-ERK on the collagen, peaking around one hour after treatment. By three hours, signaling had decreased again with tEGF surfaces sustaining ERK at a higher level than control surfaces.

Higher ERK signaling was observed on the tEGF fibronectin surfaces as well compared to control (Figure 4.19A), both sustained signaling and at the peaks. Peak signaling over basal was about four-fold, and after three hours, ERK on all tEGF fibronectin surfaces returned to basal conditions. Though ERK signaling did still occur on the AG1478-treated cells, the activation was very low (peak at 1.2-fold over basal), and we also do not observe the sharp peak at 5 minutes followed by an immediate signaling decrease. These ERK time courses together suggest that tEGF activates ERK in a sharp and rapid manner (peaks seen at 5-minutes), but that there is a separate activation of ERK that affects all surfaces regardless of EGF. This activation could be from the media change at time zero as well as ECM-associated ERK activation.

Akt signaling seems less varied on these surface conditions than ERK or EGFR (Figure 4.20). Peak signaling for vitronectin and collagen surfaces was around 1.6-fold over the basal and sustained to three hours after treatment. This was the same peak in the AG1478-treated cells, but Akt signaling was not maintained past this peak. On fibronectin surfaces, 0.3 and 10 ug/mL fibronectin control surfaces showed the highest peak of Akt signaling of approximately 2.75-fold over basal levels occurring 30 minutes after treatment. This peak decreased to the same level as the other Fn-adsorbed conditions after 60 minutes to about 2-fold over basal. These signaling

patterns suggest that neither tEGF nor ECM-adsorption have significant effects on Akt signaling since phosphor-Akt levels are very similar across all conditions.

Finally, we measured pFAK levels on all 29 surface conditions (Figure 4.21). Like Akt, there were not clear FAK signaling differences between tEGF and control conditions or between different levels of ECM adsorption. Maximum activation levels were also low across all conditions, with the highest being about 1.8-fold over control. However, two interesting trends emerge from the pFAK data. First, on the fibronectin-adsorbed conditions, FAK signaling shows a slight decrease 5 minutes after treatment, but then monotonically increases over time until the end of our time course at 3 hours post-treatment (Figure 4.21A). This trend was not seen in the AG1478-treated cells. Second, on collagen conditions after three hours, all of the tEGF surfaces have higher FAK signaling than the control surfaces. This same trend is true at time zero, though the differences are harder to see in Figure 4.21C because the data points are much closer together. These together suggest that while tEGF does appear to have a slight effect, FAK signaling is more affected by both the specific ECM protein present on the surface, and in the case of collagen, the adsorbed amount.

4.3.6.3. *Experimental variability between replicates*

The previously-mentioned time courses plot the average fold-differences of two to four biological replicates per condition. Some of the conditions varied substantially between replicates, so the associated error bars were large.

These differences are not surprising given the way that we prepare the surfaces. To conserve cost, we use sub-saturating levels of murine EGF in the tethering step. This results in clusters of EGF that are not uniformly distributed, and the cluster densities are sensitive to a number of factors. First, the specific polymer mix affects the actual availability and density of PEO sidechains. Though we mix Comb 1 and Comb 2 in a 60:40 ratio, this mixture is made fresh for each batch of coverslips to be spin coated. The end tEGF density is highly sensitive to this ratio, the

tethering buffer pH, and the tethering time [101]. We try to minimize these variables as much as possible by preparing coverslips and buffer solutions in batches. Coverslips were usually spin coated in batches of 48, enough for four ECM concentrations with and without tEGF, and six time points for each. All 48 are also EGF-tethered and ECM-adsorbed simultaneously using the same batch of freshly-made buffers and solutions. This ensures the ability to compare normalizations within each set of 48 coverslips, but variations from biological replicate to replicate are high.

Furthermore, because EGFR signaling occurs through cross-phosphorylation by homodimer pairs of ligand-bound EGFR, the tethered EGF on our surfaces must be in clusters close enough to each other to allow for receptor homodimerization. Because the tethering of EGF to PEO sidechains is random, we cannot control the actual locations and distributions of the EGF-tethering, and back-of-the envelope calculations show that the resulting tethered-EGF density using a 25 ug/mL concentration of murine EGF is right on the cusp of a distribution density that allows for receptor dimerization [123]. All of these variables also lead to high replicate-to-replicate signaling variations.
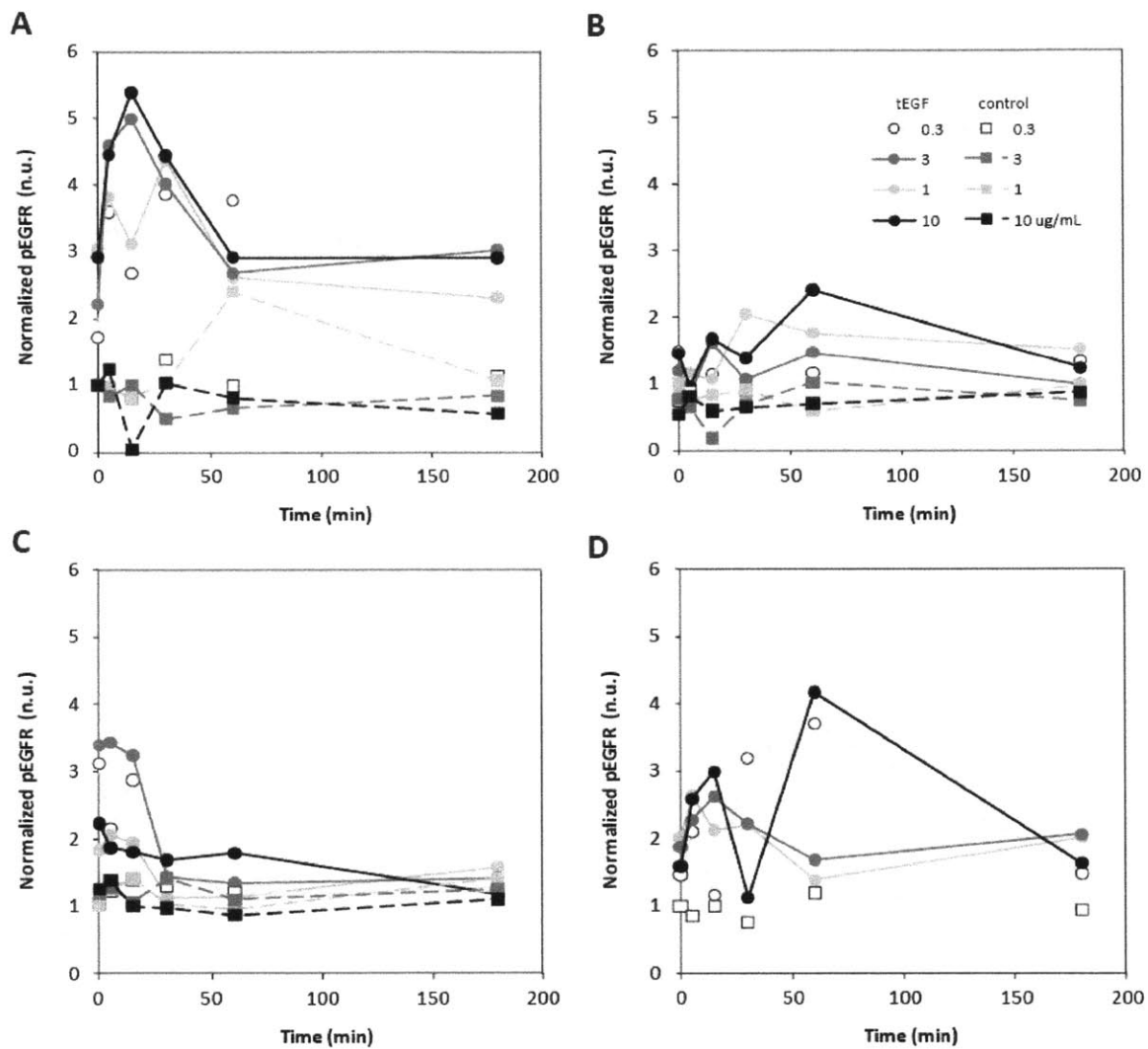
Figure 4.18. EGFR phosphorylation time course on tEGF (circles, solid lines) and control (squares, dotted lines) surfaces adsorbed with 0.3, 1, 3, and 10 ug/mL of (A) fibronectin, (B) fibronectin and EGFR inhibited with 10 uM AG1478, (C) collagen, and (D) vitronectin. Data points represent the averages of three biological replicates.
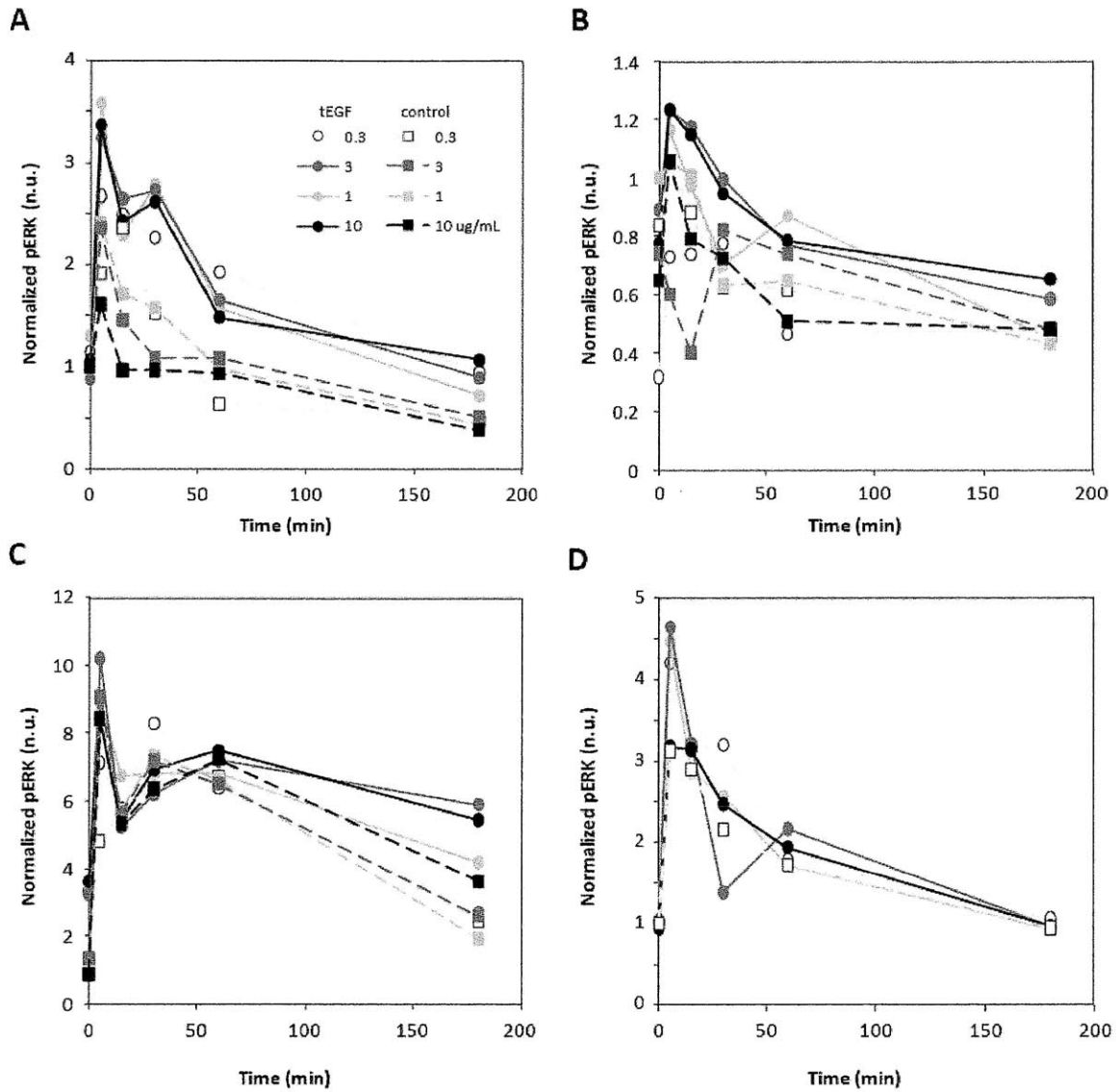
Figure 4.19. ERK phosphorylation time course on tEGF (circles, solid lines) and control (squares, dotted lines) surfaces adsorbed with 0.3, 1, 3, and 10 ug/mL of (A) fibronectin, (B) fibronectin and EGFR inhibited with 10 uM AG1478, (C) collagen, and (D) vitronectin. Data points represent the averages of three biological replicates.
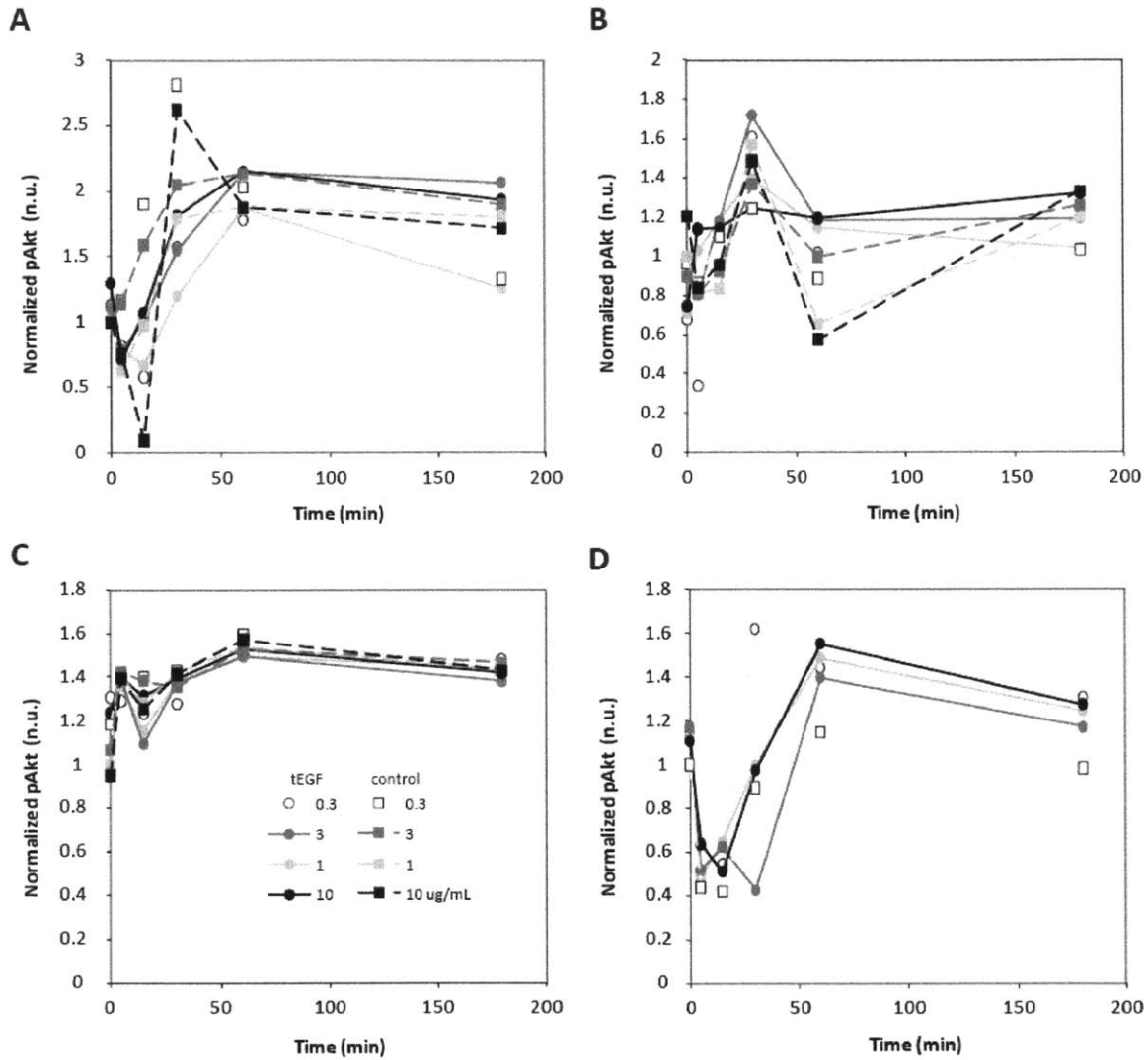
Figure 4.20. Akt phosphorylation time course on tEGF (circles, solid lines) and control (squares, dotted lines) surfaces adsorbed with 0.3, 1, 3, and 10 ug/mL of (A) fibronectin, (B) fibronectin and EGFR inhibited with 10 uM AG1478, (C) collagen, and (D) vitronectin. Data points represent the averages of three biological replicates.

105
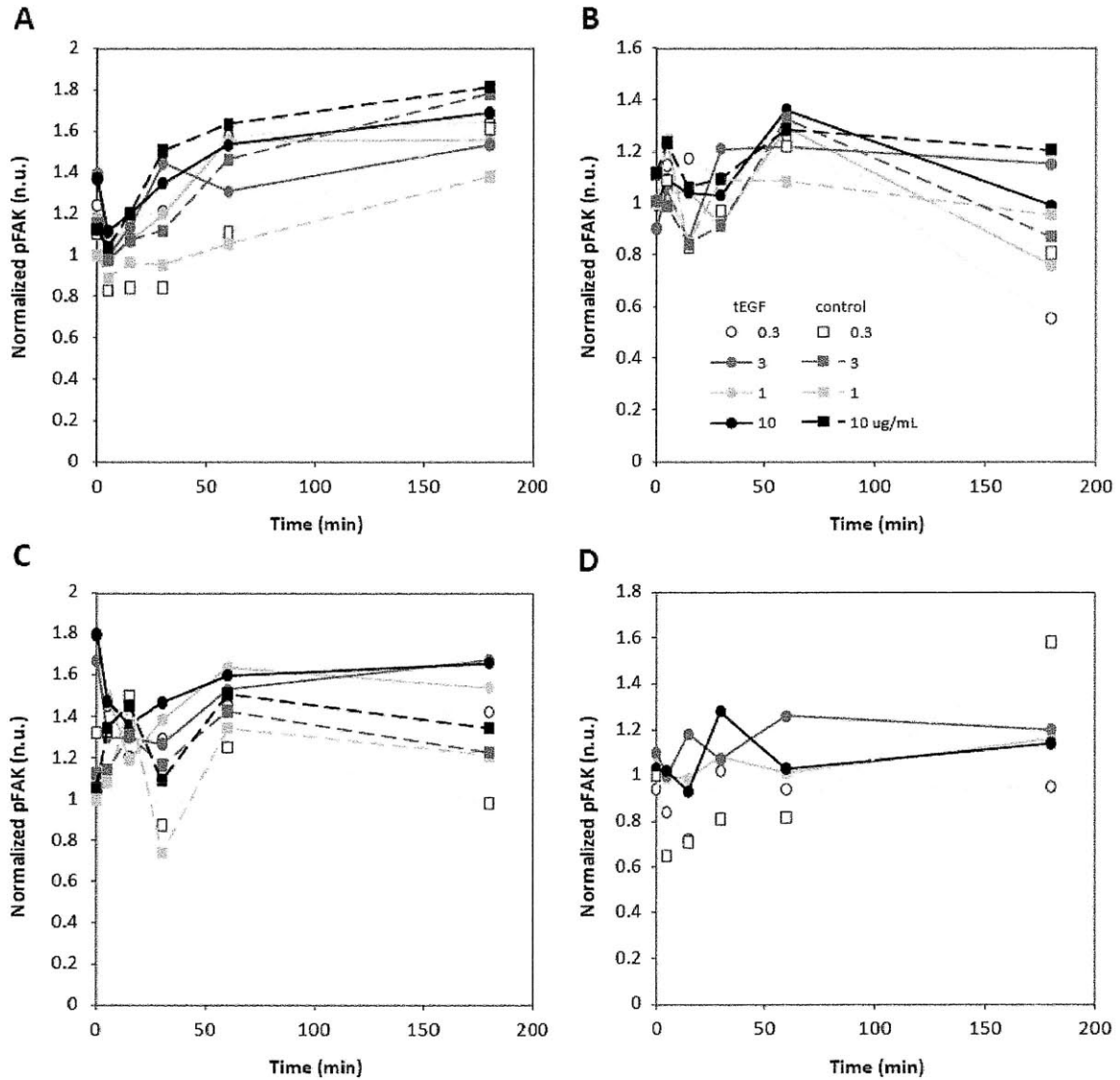
Figure 4.21. FAK phosphorylation time course on tEGF (circles, solid lines) and control (squares, dotted lines) surfaces adsorbed with 0.3, 1, 3, and 10 ug/mL of (A) fibronectin, (B) fibronectin and EGFR inhibited with 10 uM AG1478, (C) collagen, and (D) vitronectin. Each data points represent the averages of three biological replicates.
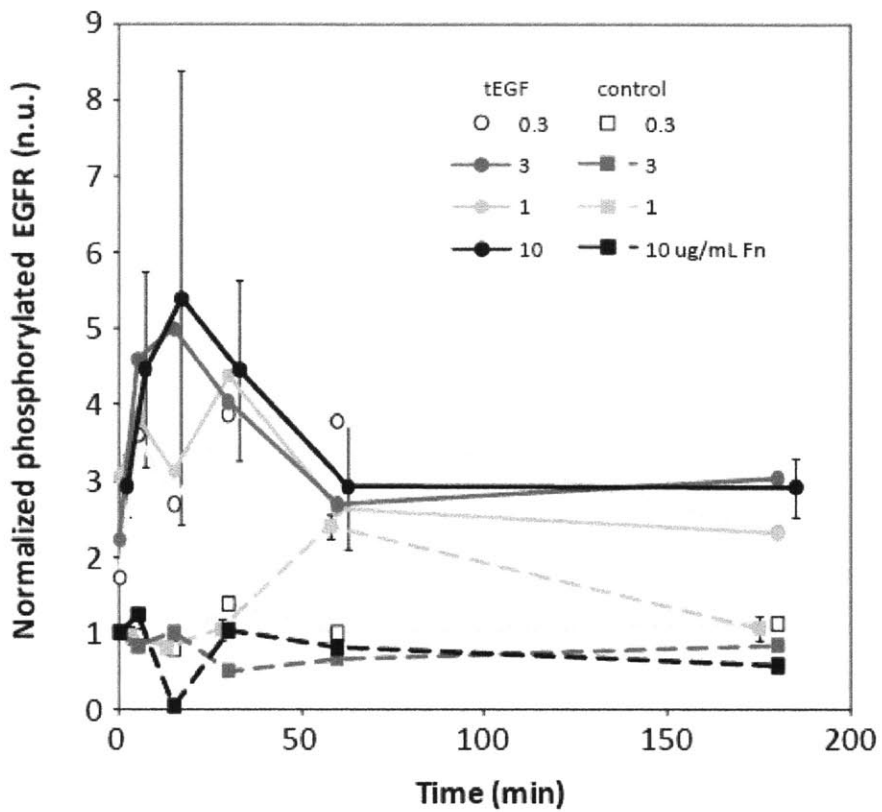
Figure 4.22. Error bars for a representative time course. To avoid clutter, only the error bars for tEGF and control on 1 ug/mL Fn conditions were plotted. Control surfaces show +/- SEM for two biological replicates, and tEGF surfaces show +/- SEM for four biological replicates.

4.3.7. Data processing of signaling measurements

Comparing pEGFR signaling on tEGF to soluble EGF (sEGF) results in Chapter 2, tEGF-stimulated signaling at 5- to 6-fold over control is much lower than the 50- to 60-fold stimulation seen with soluble EGF (Chapter 3). Differences between the 29 surface conditions were even harder to ascertain than the sEGF cases because maximum signal is only 5 to 6 times of the control. As a result, it is more difficult to tell signaling differences between surfaces because even the highest signaling peaks are only 5.5 times that of the control. Furthermore, because of the inherent variability of the tEGF surface preparation, signaling differs significantly between biological replicates. We control for some of this by normalizing within each set of experiments, but the trends are not always the same from replicate to replicate.

To minimize the effects of these variations as much as possible, we time-integrated the signals for each surface condition to compare instead of comparing the signaling levels at individual time points. For each set of 48 coverslips, essentially each block of fibronectin, vitronectin, collagen, and fibronectin plus AG1478 conditions, we normalized and averaged EGFR, ERK, Akt, FAK signaling across biological replicates exactly as described above for EGFR on fibronectin surfaces: 1) normalized raw signaling within each biological replicate to the 1 ug/mL ECM, time 0 value; 2) averaged these fold-differences across biological replicates; 3) normalized these signaling averages to the 1 ug/mL ECM average at time zero. These final normalized values were the ones we then time-integrated. These time-integrated values then represent the normalized average fold-difference over the "basal control" condition of 1 ug/mL ECM protein within each block of same-batch surfaces. Thus, these time-integrated values are then directly comparable across the entire matrix of conditions shown in Table 4.1.

Table 4.2 shows a representative data set of these time-integrated signals, using FAK signaling, across all of the 29 conditions we measured. Signaling was highest on fibronectin and collagen surfaces, as compared to their respective controls. On the collagen surfaces, a consistent

trend of increasing FAK with increasing collagen adhesiveness was seen. On the collagen control

surfaces, time-integrated FAK increased from 204 on Cn 0.3 to 249 on Cn 10. On the tEGF collagen

surfaces, FAK increased from 254 on Cn 0.3 to 284 on Cn 10. This trend also clearly shows that

tEGF increases FAK signaling on collagen, which is not the case on vitronectin and fibronectin

surfaces. The increase in FAK signaling with increasing collagen is consistent with FAK's role as a

critical integrin protein. Increasing surface adhesiveness should lead to increased integrin size and

formation, which would lead to higher FAK signaling. We saw the lowest FAK signaling on the

Fibronectin	Int FAK	Fn + AG1478	Int FAK	Collagen	Int FAK	Vitonectin	Int FAK
Fn 0.3 tEGF	269	Fn 0.3t + AG	178	Cn 0.3 tEGF	254	Vn 0.3 tEGF	168
Fn 1 tEGF	260	Fn 1t + AG	187	Cn 1 tEGF	277	Vn 1 tEGF	192
Fn 3 tEGF	248	Fn 3t + AG	208	Cn 3 tEGF	274	Vn 3 tEGF	216
Fn 10 tEGF	273	Fn 10t + AG	208	Cn 10 tEGF	285	Vn 10 tEGF	196
Fn 0.3 ctrl	218	Fn 0.3 + AG	183	Cn 0.3 ctrol	205	Vn 1 ctrl	191
Fn 1 ctrl	204	Fn 1 + AG	185	Cn 1 ctrl	219	**Average**	**192.6**
Fn 3 ctrl	265	Fn 3 + AG	193	Cn 3 ctrl	234		
Fn 10 ctrl	291	Fn 10 + AG	218	Cn 10 ctrl	249		
Average	**253.5**	**Average**	**195.1**	**Average**	**249.5**		

Table 4.2. Integrated FAK signaling from all conditions. Within each set of surface conditions, signaling at each time point was normalized to that condition's control 1 ug/mL ECM condition. These normalized values were integrated over the entire time course for each condition.

vitronectin conditions and also fibronectin conditions in the presence of the EGFR inhibitor

AG1478. The average time-integrated FAK on all vitronectin surfaces and all fibronectin plus

AG1478 surfaces were 192.6 and 195.1 respectively, compared to values in the mid-200s on both

fibronectin and collagen surfaces. The decrease in FAK on AG1478-inhibited surfaces supports

reported evidence of EGFR and integrin-signaling crosstalk [133], and overall suggests that there

are more and/or larger focal adhesions on fibronectin and collagen surfaces than vitronectin although FAK immunofluorescence staining did not show obvious differences.

4.3.8. MSC signaling on ECM, tEGF surfaces

We measured the signaling of four phosphoproteins at six time points on 29 different surface conditions. In order to better graphically see all of our, we generated a heatmap of all of the integrated signals across all of the conditions (Figure 4.23). Each column of the heatmap was rescaled to be between 1 and 2, with the lowest signaling condition having a value of 1 (black), and the highest signaling condition having a value of 2 (white) with a spectrum of red to orange to yellow in between. We chose 1 as the baseline instead of 0 to show that the lowest integrated signals for each protein does not necessarily mean that there was no signaling, but that that the integrated signal was lowest when compared to the other 28 conditions. For example, Vn 0.3 tEGF shows up as black in the FAK column of the heatmap (Figure 4.23), but Table 4.2 clearly shows that FAK signaling is not zero for this condition. The trend we observed of time-integrated FAK being lower on vitronectin and AG1478 conditions is clearer when seen graphically: these blocks in the heatmap are more red and darker than the fibronectin and collagen blocks which are more yellow and lighter in color.

Integrated Akt trends are similar to FAK: lowest on vitronectin and with AG1478, higher on collagen surfaces, and highest on fibronectin surfaces. ERK, by comparison, is very high on collagen surfaces and low on all other surfaces. EGFR activation is highest on tEGF surfaces. Within the block of fibronectin surfaces in the presence of AG1478, the low activation of EGFR is clearly seen on the tEGF surfaces compared to control. Overall, EGFR time-integrated signals were highest on fibronectin surfaces and lowest on collagen surfaces.
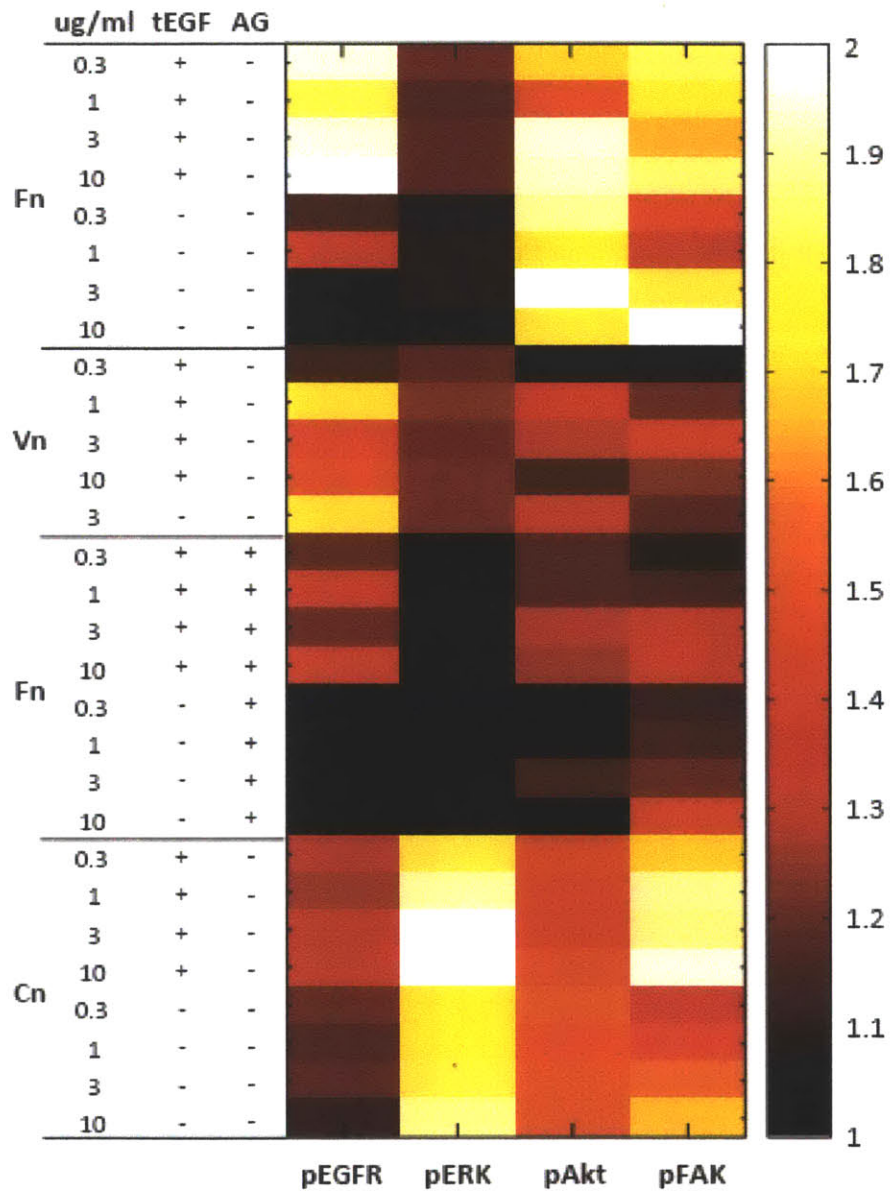
Figure 4.23. Normalized integrated signals on all cues conditions. The integrated values for each signaling protein for all of the measured conditions were rescaled to between 1 and 2 to visualize this heatmap. A black block (heatmap value = 1) means integrated signaling was the lowest at this condition compared to the other 29, and not necessarily that there was no signaling at this condition.

4.3.9. Univariate correlations between signal and response

We measured signaling on more surface conditions than we tracked migration, so we reduced the signaling matrix from the end of Chapter 4 down to the 19 conditions for which we have tracked cell migration data (Figure 4.24A). The removed conditions are the control vitronectin and collagen surfaces and a number of the AG1478-treated cells on fibronectin surfaces. The only three EGFR-inhibitor conditions that we tracked cell migration for were 3 ug/mL Fn, with
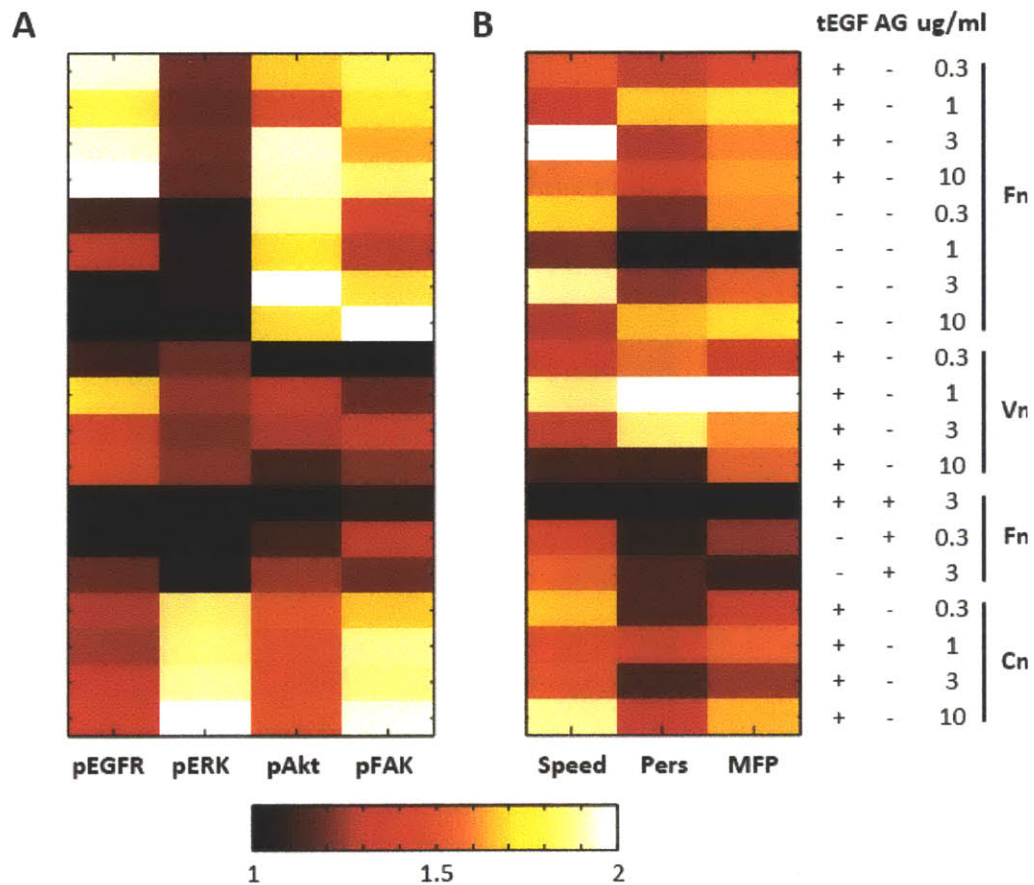
Figure 4.24. Normalized time-integrated protein signals (A) and migration responses (B) rescaled from 1 (lowest, black) to 2 (highest, white) for 19 conditions.

and without tEGF, and 0.3 ug/mL Fn without tEGF. FAK and Akt signaling on Vn surfaces and under AG1478 treatment were clearly lower than on the other surface conditions. Integrated EGFR and ERK signaling were lowest on control fibronectin surfaces and also in AG1478-treated cells. Figure 4.24B is a heat map of the raw responses of Speed, Persistence Time, and MFP for these same 19 conditions, rescaled from 1 to 2.

We first investigated whether any single one of the proteins signals we measured would be predictive of the migration responses. Plotting univariate correlations of the migration responses of Speed, Persistence Time, MFP, and DI versus time-integrated EGFR, we see no clear relationships between Speed, PT, MFP and EGFR signaling (Figure 4.25 A, B, C). There does appear to be an inverse relationship with DI where very high DIs are associated with low EGFR signaling (Figure 4.25D). Since there was no correlation between PT and EGFR, this DI versus EGFR relationship is baffling. In Chapter 4, we explored the differences between DI and PT and found trends of both on ECM-adsorbed comb surfaces to be similar. Plotting DI versus PT explicitly, we also see an inverse trend, suggesting that these two parameters do represent the same cell behavior (Figure 4.26).

The DI versus EGFR and DI versus PT plots resemble each other in that both have three data points on the left that have much higher DIs than all other points. The three conditions associated with these points are 0.3 ug/mL Fn control, 0.3 ug/mL Fn control plus AG1478, and 3 ug/mL Fn tEGF plus AG1478, all conditions with very low EGFR signaling as well as low directional persistence. These may actually be outliers given that cells on these surfaces show lower migration speeds than other surfaces as well, rather than true correlation between cell migration DI and EGFR signaling.

Figure 4.25. Migration responses versus time-integrated EGFR. 19 conditions of Speed (A), Persistence Time (B), MFP (C), and DI (D) were plotted against the conditions' time-integrated EGFR signal.

114
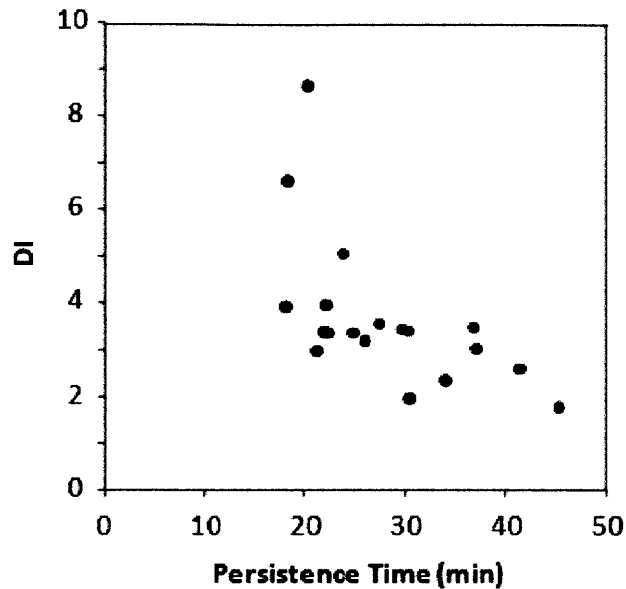
Figure 4.26. MSC DI versus Persistence Time.

We also generated the same four correlation plots for each of the other three phosphoproteins that

we measured: ERK, Akt, and FAK. Two distinct clusters of ERK signaling emerge: low versus high,

but there are otherwise no obvious relationship between MSC migration response and integrated

ERK within each cluster (Figure 4.27). The four data points at high integrated ERK were the four

collagen points.

The Akt and FAK plots also show no obvious relationships between migration response and

these single signals (Figure 4.28, Figure 4.29), suggesting that no one pathway can univariately

predict MSC migration response. Given that cell migration is a complex biophysical process

involving several signaling pathways, this is not altogether surprising. We then explored

computational models such as decision trees to ascertain multivariate relationship between EGFR,

ERK, Akt, FAK protein signaling with MSC migration response.

115
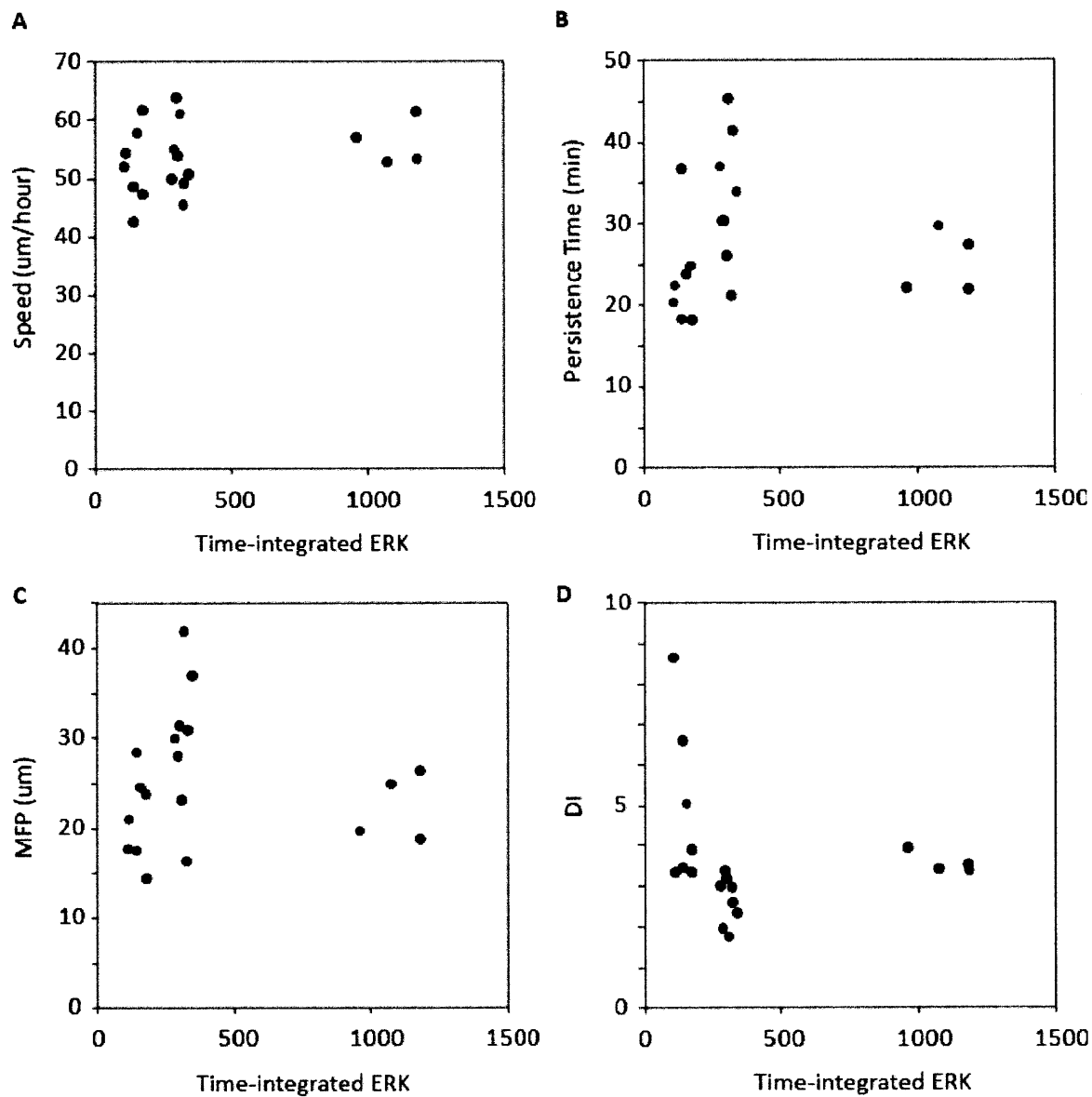
Figure 4.27. Migration responses versus time-integrated ERK. 19 conditions of Speed (A), Persistence Time (B), MFP (C), and DI (D) were plotted against the conditions' time-integrated ERK signal.

116
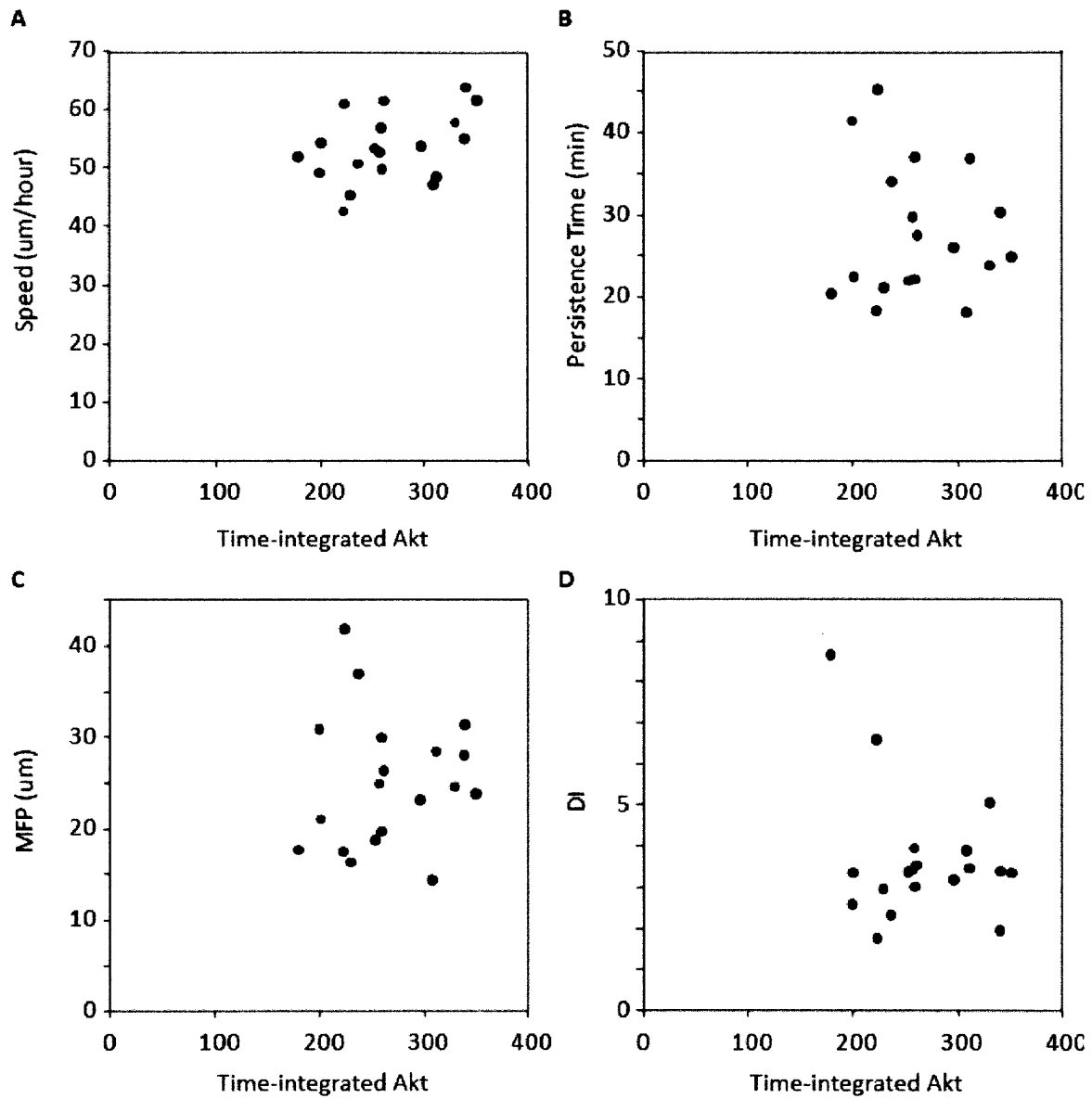
Figure 4.28. Migration responses versus time-integrated Akt. 19 conditions of Speed (A), Persistence Time (B), MFP (C), and DI (D) were plotted against the conditions' time-integrated Akt signal.
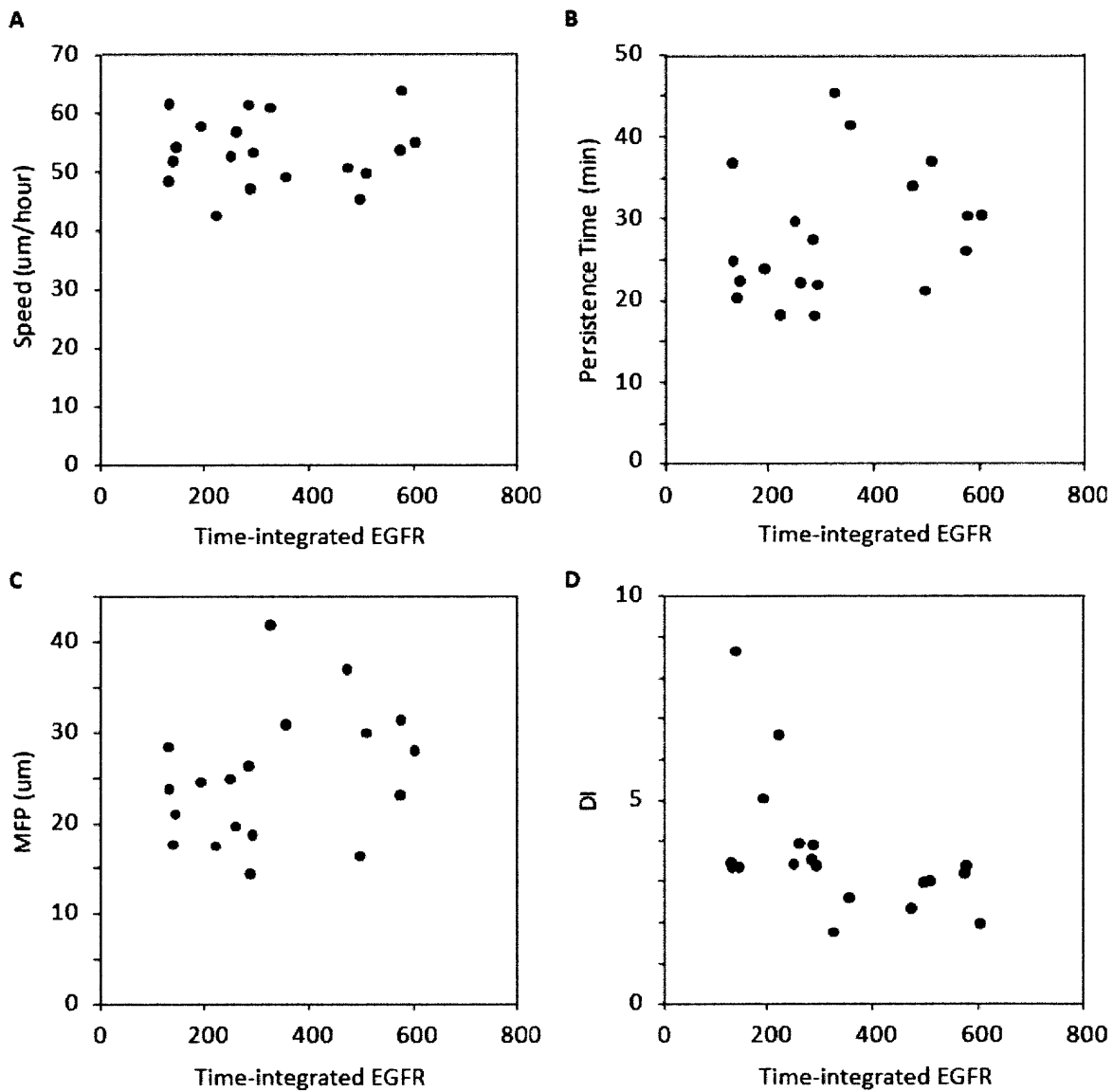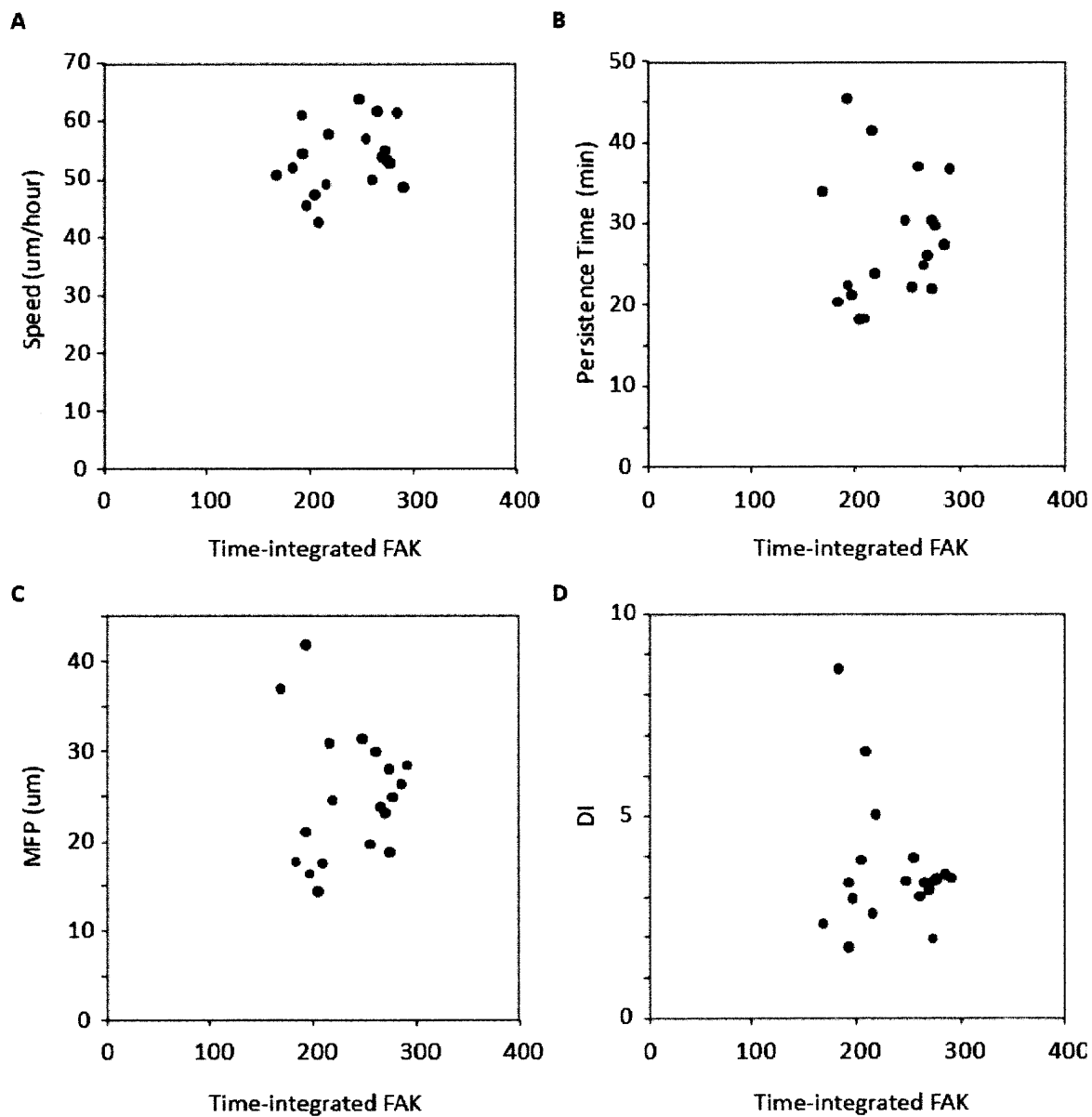
Figure 4.29. Migration responses versus time-integrated FAK. 19 conditions of Speed (A), Persistence Time (B), MFP (C), and DI (D) were plotted against the conditions' time-integrated FAK signal.

118

4.4. Summary and conclusions

In this chapter, we studied the effects of tEGF and ECM adhesiveness on MSC migration response and signaling. Using single cell tracking, we presented a rigorous quantitative approach to investigate MSC migration, augmenting the limited prior knowledge in this field. Single cell tracking of migration allowed us to separate cell migration into its two intrinsically different parameters of speed and directional persistence. The conventional treatment of cell migration as only cell speed is limited in its evaluation MSCs' ability to infiltrate target tissue for wound healing applications. A fast moving cell does not necessarily translate to productive migration if it undergoes frequent directional changes. By calculating cell migration speed and persistence time separately, we were able to produce a composite parameter for MSC migration: cellular mean free path (MFP), defined as the the characteristic distance a cell migrates before changing directions. The higher the cell's MFP, the better its ability to "cover ground" *in vivo* and in populating *ex vivo* scaffolds.

We measured the activities of four phosphoproteins as effectors of MSC migration in response to tEGF and ECM cues: EGFR, ERK, Akt and FAK. Disappointingly, many signals were undetectable, or did not vary discernibly with cue conditions, in hTERT-MSCs. Even so, the four we measured are part of disparate pathways that affect migration with Akt potentially acting as a surrogate for the status of ROCK and MLC as these proteins and Akt are all downstream of PI3K. No single measured protein signal was able to predict MSC migration, suggesting that combinations of protein signaling contribute to overall migration response. This is not surprising given the intricacies of the biophysical processes of cell migration and necessitates a systems-level approach to multi-variately relate signaling pathways to migration.

119

5. Chapter 5 - Decision tree modeling to discover signaling targets for controlling MSC migration

5.1. Introduction

Cues external to a cell affects responses through the amplification, propagation, and activation of various signaling and metabolic pathways within the cell. With the advent of high-throughput methodologies for measuring intracellular pathway states, we are often inundated with complex arrays of data from crosstalking pathways that necessitate systems-level analysis approaches [134]. Computational models that multivariately correlate complex crosstalking pathways with cell responses are critical tools in the process of siphoning information and identifying not only governing pathways, but also non-intuitive interactions and potential targets for drug therapy [135]. In this Chapter, we present decision tree analysis as one such computational tool and show a proof of concept of non-intuitive drug target discovery for the rational design of tissue engineering therapeutics to repair bone injury.

5.1.1. Decision tree modeling

Decision tree-based analysis is a form of supervised learning for the interpretation of observed data with the advantages of being easily interpretable, transparent, flexible, and applicable to a wide array of problems in computational and systems biology [136]. Trees are easy
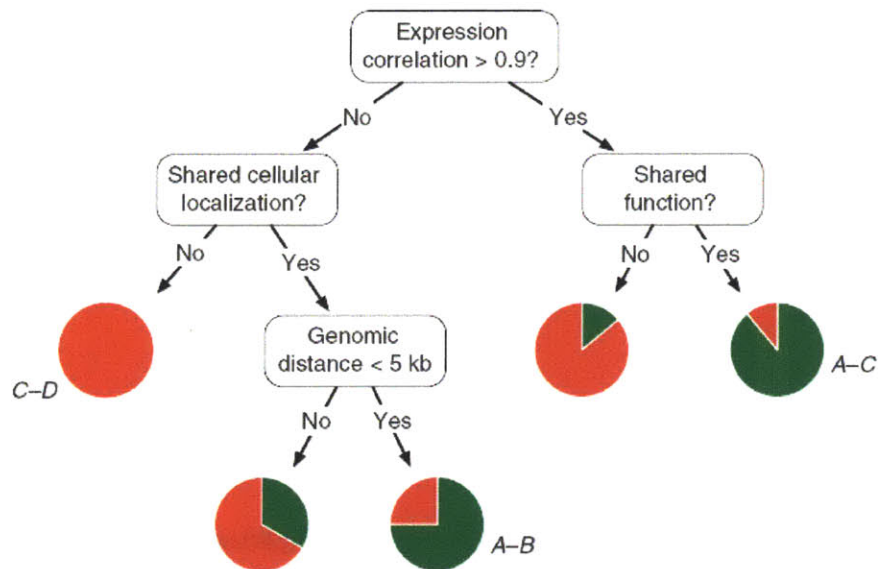
Figure 5.1. Hypothetical prediction of protein-protein interactions via a binary decision tree classification. Pie-charts show the percentage of interactors (green) versus non-interactors, with the leaf "class" determined by the predominant color. *Source: Kingsford & Salzberg, 2008* [137].

to read partly because each node can be viewed as containing a question. In the hypothetical tree of Figure 5.1, the questions in each node have "yes" or "no" answers, with the corresponding branches leading to new nodes with new questions and finally to the result classes in the leaves [137]. Tree are constructed with training data, but after construction, new data can be classified by following the path from the root node at the top of the tree, answering questions along the way at each node until a leaf is reached. The "class" associated with the reached leaf is then the predicted

classification for the unknown data. Provided that the quality of the training data set used for tree construction is high, decision trees provide highly accurate predictions.

In Chapter 2, we described another application of decision trees: identifying potential signaling perturbation targets to controllably modulate fibroblast cell behavior. In a signal-response tree where the tree nodes are signals and the leaves are classes of responses, visual inspection of the tree will 1) identify the hierarchy of signals' contribution to cell response and 2) node splits leading to different response classifications in the leaves [82]. Particularly where these nodes lead to drastically different response classes, each node signal generates a hypothesis regarding the contribution of that node to cell response. As discussed in Chapter 2, decision tree modeling predicted that subtotal inhibition of MLC would lead to an increase in fibroblast migration speed, a non-intuitive prediction that was tested and confirmed experimentally.

5.1.2. Chapter scope

As shown in Chapter 4, univariate correlations between the activities of EGFR, ERK, Akt, and FAK were unable to predict migration responses, suggesting that a combination of all four phosphoproteins is needed to encapsulate the MSC migration responses of Speed, PT, and MFP. Given the proof of concept from Chapter 2, we modeled MSC migration signal and responses using decision tree analysis. The rest of this chapter discusses those results, predictions and hypotheses from tree construction, and finally an experimental test of one hypothesis. Finally, we discuss the implications of this modeling in the context of designing solutions for bone tissue engineering.

5.2. Materials and methods

5.2.1. Cell culture and surface preparation

Cell culture of hTERT-MSCs was done following the same methods as described in Chapter 4. Glass coverslips spin coated with comb co-polymer were also prepared as previously described and further adsorbed with ECM proteins of interest.

5.2.2. ERK inhibition dose response

ERK was inhibited with either U1026 or CI-1040 (PD184352), both purchased from Selleck Chemicals (Houston, TX). Stock U1026 or CI-1040 was reconstituted in DMSO and stored at -20°C until dilution and use.

To test dose-dependent inhibition of ERK by both drugs, 200,000 MSCs were seeded in each well of a 6-well plate and allowed to attach and grow overnight. Cells were treated with 0.01, 0.1, 1, and 10 nM of either U1026 or CI-1040 for one hour prior to adding 10 nM of soluble murine EGF. Cells in each well were lysed with 75 uL of lysis buffer per well consisting of Bioplex lysis buffer lysis buffer base (Bio-Rad Laboratories, Hercules, CA) plus 1x Factors A and B (Bio-Rad), and 1 mM phenylmethanesulfonyl fluoride (PMSF) (Mallinckrodt Baker, Mansfield, MA), collected into eppendorf tubes, and centrifuged for 10 minutes at 13,000 rpms and 4°C. A small volume was removed for protein quantification via BCA Assay, and the rest stored at -80°C until use.

5.2.3. ERK inhibited migration and phosphoprotein measurements

10,000 cells were resuspended in 2 mL of assay media containing 0, 0.3, or 1 nM of U0126 and seeded onto prepared Comb co-polymer surfaces adsorbed with 3 ug/mL of collagen in DeltaT dishes. The cells were allowed to attach, spread, and acclimate at 37°C, 5% CO_2 for 18 hours. A small amount, 5-10 mL, of identical media was simultaneously added to a T-flask and placed into the same incubator to be used as assay media. Following 18 hours, the media in the DeltaT dishes

were replaced with 3.2 mL of assay media, and cell migration tracked using previously described procedures.

For phosphoprotein measurements, 100,000 cells resuspended in 1 mL media containing 10 uM of AG1478 were seeded onto each surface. After 3 hours, almost all cells have attached and spread. The desired concentration of U1026 was then added to treat the MSCs, still in the presence of AG1478. After two hours, the media was replaced with assay media containing only U1026 and no AG1478. Cells were treated for 5 minutes and lysed following the protocols described in Chapters 3 and 4. 10 ug of cell lysate protein was used to measure phosphoprotein levels via ELISA and the Luminex xMAP system following manufacturers' protocols and as described in previous chapters.

5.3. Results and discussion

5.3.1. Decision tree modeling of MSC migration "signal-response"

Binary decision trees were generated using the Matlab function `classregtree`, which takes two inputs, a vector **y** of *n* responses and a matrix **X** of *n* rows by *m* columns. Each column in **X** contains the time-integrated and normalized values of one phosphorotein signal, or one predictor of the responses in **y**. Each row in **X** contains the predictor values associated with a specific surface condition and corresponds to the same row index in **y**, which contains the response of that condition. We have three response vectors: MFP, Speed, PT. Given that PT and DI are related, we only made trees for PT and not both. Our predictor matrix has four columns of phosphoprotein time-integrated signals; EGFR, ERK, Akt, and FAK.

The Matlab function `classregtree` can generate two types of trees: regression or classification trees. A vector **y** of numerical values, such as cell MFPs, defaults `classregtree` to generate a regression tree. However given our relatively small dataset of 19 conditions, regression trees is more likely to over fit our data. Thus, we discretized the predictors and responses and use these discretized classes to generate classification trees.

5.3.1.1.   Data discretization

We used a simple range discretization method to discretize our data to three bins: low (0), medium (1), and high (2). Because our data vectors are not normally distributed, range discretization where the data range of each bin is equal is more appropriate than quantile discretization where each bin receives the same number of data values. It is important to note that "0" corresponds to low signal or low response and not absolute zero. Figure 5.2 shows the time-integrated signals and migration responses from Figure 4.24 after 3-bin discretization.
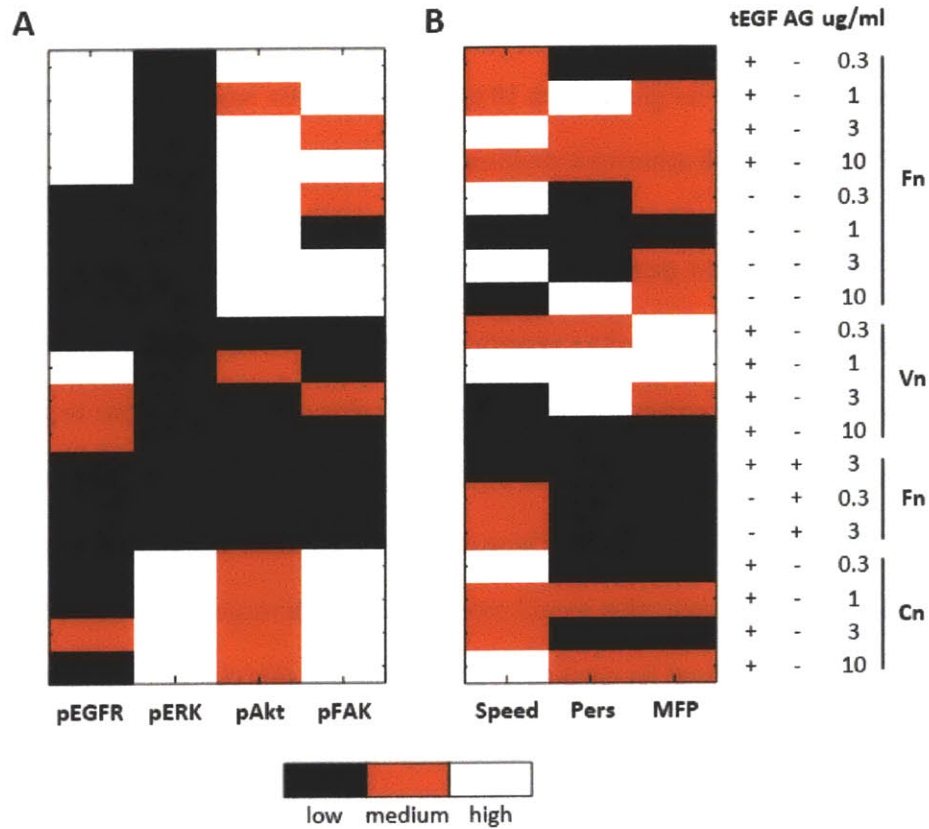
Figure 5.2. Discretized MSC migration signals (A) and responses (B). Three-bin range discretization was used for each column of signals and responses, turning original values to low, medium, and high.

In the signals heat map, the entire block associated with AG1478-treated cells is black, indicating that the signals of all four phosphoproteins in these cells are low. Collagen surfaces had low EGFR signals but high ERK compared to the other conditions. In fact, because ERK signaling was so high on the collagen surfaces, no conditions even had medium levels of ERK in comparison, overwhelmed by the levels of ERK on collagen. From the response heat map, the importance of the MFP parameter, as an all-encompassing parameter encapsulating both the cell speed and PT, is well-highlighted. When cell speed and PT are both low, MFP would naturally be low as well (3 ug/mL Fn with AG1478, tEGF). However, even when speed is high (3 ug/mL Fn, tEGF), a mediocre

127

PT prevents the cell from covering a lot of ground and thus taking advantage of the full potential of its high absolutely speed motility. On the other end of the spectrum, a cell moving slowly does not necessarily mean it cannot cover ground. On 10 ug/mL fibronectin mock surfaces, cell speed is low, but because PT is high, the cell achieves a medium level of MFP.

5.3.1.2.   Decision tree generation

The algorithm employed by `classregtree` is the classic Classification and Regression Tree (CART) system which searches through all possible splits on every predictor to generate the nodes and branches of the tree [138]. We use `classregtree` with its default *Gini* criterion for choosing to make splits. Briefly, when searching through all possible splits on all predictors, *Gini* criterion algorithm looks for splits that would produce subsets of responses, "classes," that carry the same class values prioritizing for class size. This means that *Gini* chooses the split that would produce the largest subset of the same class and separates this class from the others. The separated class is considered "pure," and this splitting process is repeated for the other presumably "unpure" branch until the resulting subsets are homogeneous, identical in predictor values, and/or the `splitmin` value is reached. The alternative criterion to *Gini* is *Twoing*, which first tries to split the data set into two equal-sized partitions while prioritizing the partition elements to be as homogeneous as possible.

The *Gini* criterion for splitting is the default because it often produces the best splitting rules for the best fitted trees [138], but the criterion can result in lopsided trees, especially given our relatively small number data set of 19 conditions. After each split, the separated class is often considered by the algorithm to be "pure enough," leading to trees with a central backbone and single leaves branching off of them. Thus, it is sometimes more informative to use *Twoing* as an alternative criterion.

Using the default *Gini* criterion for splitting, we generated three decision trees, one for each of the responses of Persistence Time, Speed, and Mean Free Path (Figure 5.3A, B, C respectively). The leaves in these trees represent the response, and the nodes represent the phosphoproteins that lead to those responses along the values on the branches. Under each leaf, "n" indicates the number of conditions out of 19 that classified to this leaf, and the percentage shows what percent of that number actually exhibits that classified value. For example, in Figure 5.3A, from the top node, if EGFR is high (right branch), and Akt is high (also right branch), then the tree predicts the cells to have medium persistence. Of the 19 conditions, three classified to this leaf meaning three of the 19 conditions had high EGFR and high Akt. However, of these three, only two (67%) actually had medium persistence.

The Persistence Time tree uses all four phosphoproteins as predictors, whereas the Speed and MFP trees each used only three - Speed leaves out ERK, and MFP leaves out Akt (Figure 5.3B, C). Each tree was constructed in Matlab with the `splitmin` parameter set to 4. This means that impure nodes must have four or more observations to continue to be split to prevent over fitting of the tree. Otherwise, it is conceivable that the tree continues splitting until all leaves contain only one observation, making perfect predictions. Given the *Gini* split criterion, the top node in each tree, the "root predictor" is the hierarchically most important predictor since it best separates out the largest class of responses. Interestingly, PT, Speed and MFP all had different root predictors of EGFR, Akt, and FAK respectively.

The PT tree starts with EGFR, which split the responses into a left group of 14 conditions with low and medium EGFR signaling and a right group of 5 conditions with high EGFR signaling (Figure 5.3A). The right group was further split along Akt signaling (low/medium versus high), while the left group was split via ERK (low versus high). The left branch of the ERK split was further split by FAK (low versus medium/high). Because the medium/ FAK branch still contained four responses, the `splitmin` parameter, these four responses were split again along their EGFR

129

signaling. This split is in contrast to the low and medium EGFR, high ERK leaf. It also contains four responses, of which only two (50%) exhibit the "phenotype" at that leaf of low persistence. This leaf however was not further split because all there was not a predictor that differed to make a potential split.

Figure 5.3C shows the *Gini*-generated MFP decision tree, which has good predictability. The root predictor was FAK (low versus medium/high), and the left branch was further split by EGFR (low versus medium/high), and the right branch was further split by ERK (low versus high). All but the right-most leaf had high prediction accuracies, though the leaf for high MFP only contained one condition.

Figure 5.3B shows the generated decision tree for MSC Speed. Akt is the root predictor, splitting the responses in a left group of six (low Akt) and a right group of 13 (medium/high Akt). The right group was split via FAK (low versus medium/high), and the low branch was further split via predictor EGFR. This is not a particularly well-classified tree. There are two leaves that contain only one response each, and the other two have low percentage accuracy. Because *Gini* did not perform particularly well with this data set, we tried a *Twoing* split tree.

Figure 5.4B shows the Speed decision tree using *Twoing* criterion. Like the tree in Figure 5.3B, ERK is not a predictor node. Because of the nature of the *Twoing* splitting criterion whereby the root predictor is one that can split the data into equal-sized partitions, the partitions themselves are often further split along identical predictors leading to the type of tree seen here. Repeating predictors at sublevels is less common for *Gini* trees because the most ubiquitous predictors are often split first, and subsequent splits "refine" the splits above it in the tree. This *Twoing*-split Speed tree has three levels of predictors: Akt, EGFR, and FAK. The leaves are much more homogeneous than in the *Gini*-generated tree, with all leaves having above 50% accuracy in prediction.

A

pEGFR

pERK — L,M ... H — pAkt

pFAK — L ... H — L (n = 4, 50%)

L,M — H (n = 2, 100%) ... H — M (n = 3, 67%)

L — M,H

L (n = 6, 83%) ... pEGFR

L — M

L (n = 3, 67%) ... H (n = 1, 100%)

Persistence

B

pAkt

L — M,H

L (n = 6, 50%) ... pFAK

L — M,H

pEGFR ... M (n = 11, 45%)

L — M,H

L (n = 1, 100%) ... H (n = 1, 100%)

Speed

C

pFAK

L — M,H

pEGFR ... pERK

L — M,H ... L — H

L (n = 6, 83%) ... H (n = 1, 100%) ... M (n = 8, 88%) ... L (n = 4, 50%)
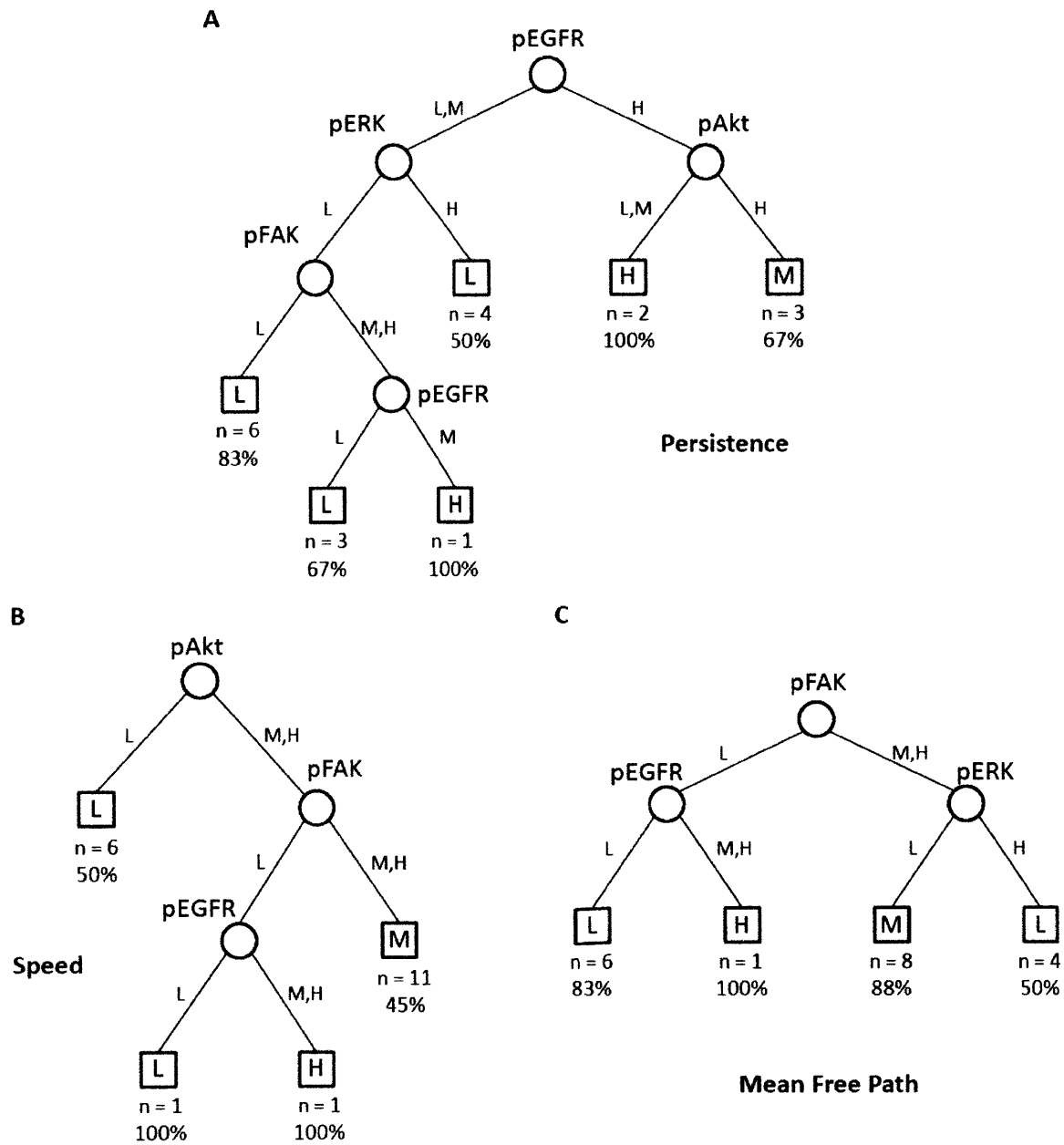
Mean Free Path

Figure 5.3. Decision trees using discretized signals as predictors of responses of (A) Persistence Time, (B) Speed, and (C) Mean Free Path. Trees generated using *Gini* split criterion.

131
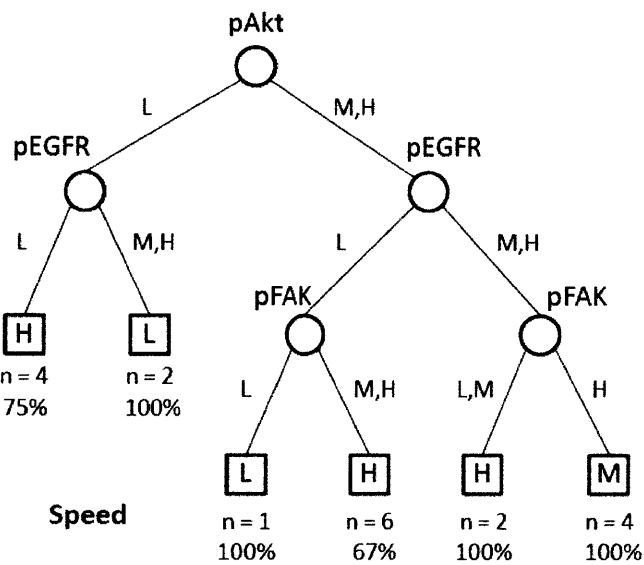
Figure 5.4. Signaling-predicted decision tree for MSC speed using *Twoing* criterion for tree splitting.

5.3.2. Hypothesis generation from DT models

Decision trees relate the cell signaling predictors to the ultimate cell migration responses in a graphically intuitive manner, providing a "function" $f(X) = y$ where X is the matrix of signals, and y is a vector of responses. Following a canonical cue-signal-response treatment of cell behavior, we assume this decision tree function to hold true and stable. In other words, this function describes how cell signaling maps to cell response, and a particular signaling profile will always map to a specific response profile. With this assumption, we can generate hypotheses from our decision tree models that alter the signaling matrix X, shifting the signaling landscape, resulting in $f(X)$ "calculating" to a different vector y of cell response.

We use the *Gini*-generated trees for Persistence Time and MFP (Figure 5.3A, C) and the *Twoing* Speed tree (Figure 5.4) to generate hypotheses. Several hypotheses can be formed from the graphical signal-response trees:

1. Decreasing ERK signaling from high to low increases cell migration MFP from low to medium (red arrow, Figure 5.5A).

2. Decreasing EGFR signaling to low lowers cell MFP from high to low (grey arrow, Figure 5.5A).

3. Decreasing Akt signaling increases cell Persistence Time from medium to high (grey arrow, Figure 5.5B).

4. Decreasing FAK in the presence of medium or high EGFR increases speed from medium to high (blue arrow, Figure 5.5C).

5. Decreasing FAK while EGFR is low decreases cell speed from high to low (black arrow, Figure 5.5C).

We next examined these hypotheses to evaluate experimental testability.

5.3.2.1. *Hypothesis 1: Decrease ERK signaling to increase MFP*

Decreasing ERK signaling from high to low would lower cell migration MFP from high to medium (red arrow, Figure 5.5A). The same decrease of ERK in the Persistence Time tree could increase cell PT if a) FAK is low, or b) FAK is not low and EGFR signaling is medium (red arrow, Figure 5.5B).

The right half of the MFP tree showed that when FAK is not low, high ERK leads to low MFP and low ERK leads to medium MFP. This tree predicted that ERK and MFP are inversely related and generated the hypothesis that decreasing ERK would "push" cells from the right-most branch (low MFP) to one with medium MFP (red arrow). The 4 conditions at the right-most leaf are all four of the Cn-adsorbed conditions and are not well classified at 50% accuracy. The discretized data for these conditions show that despite all four having high FAK and high ERK, two of the four (0.3 and 3 ug/mL) have low MFP whereas the other two (1 and 10 ug/mL) have medium MFP, so the classification is indeed split half and half (Table 1). To gain more insight, we generated a binary

decision tree for cell persistence (Figure 5.5B). All four of the Cn-adsorbed conditions classify

together in the persistence tree as well with 50% accuracy (low or medium EGFR, high ERK).

Applying the hypothesis from the MFP tree (lowering ERK would increase MFP), the persistence

tree predicts that lowering ERK would increase persistence under two conditions: 1) if FAK is low,

or 2) if FAK is not low, then if EGFR is medium. All of the Cn conditions have high discretized FAK

and thus would not meet condition one (Table 5.1). However, on 3 ug/mL Cn, EGFR is indeed

medium. Thus at 3 ug/mL Cn with tEGF, our MFP and persistence tree modeling predicts that

lowering ERK would increase MFP through an increase of persistence, a testable hypothesis.

	pEGFR	pAkt	pFAK	pERK	MFP	Pers
C0.3t	L	M	H	H	L	L
C1t	L	M	H	H	M	M
C3t	M	M	H	H	L	L
C10t	L	M	H	H	M	M

Table 5.1. Discretized signal and response data of the four collagen-adsorbed conditions on tEGF polymer surfaces.

5.3.2.2. *Hypothesis 2: Decrease EGFR signaling to lower MFP*

In the MFP tree of Figure 5.5A, the left branch from the FAK root predictor shows that by

decreasing EGFR signaling, we can lower the cell's MFP from high to low. Again because MFP is a

comprehensive parameter encompassing both Speed and Persistence Time, we looked at the other

two decision trees as well to see how EGFR signaling affects Speed and PT both to result in a lower

MFP. EGFR is the root predictor in the Persistence Time tree (Figure 5.5B), and both branches lead

to another predictor to further categorize cell PT. Thus, the effects of decreasing EGFR from high to

low on cell PT is unclear because the result would depend on the specific levels of all three of the other phosphoproteins. This convoluted effect of decreasing EGFR is not surprising since it is upstream of all three of our other measured phosphoproteins.

As for testing this hypothesis, we have actually in fact already "tested" it. In Chapter 4, we showed that decreasing EGFR with AG1478 decreased both cell Speed and Persistence Time, which naturally would lead to a decreased MFP (Figure 4.12, Figure 4.13). EGFR was inhibited on tEGF and control surfaces adsorbed with 3 ug/mL of fibronectin, two of the 19 conditions used to build our decision tree. So this successful "testing" of our hypothesis is actually a recapitulation of two training conditions, ruling it out as an actual testable hypothesis.

5.3.2.3. *Hypothesis 3: Decrease Akt signaling to increase Persistence*

Our Persistence Time decision tree shows that in the presence of high EGFR signaling, high levels of Akt lead to medium PT, whereas low/medium levels of Akt lead to high PT. This produces the hypothesis that when EGFR is high, decreasing Akt would increase PT (grey arrow, Figure 5.5B).

How this increase in PT would affect overall cell MFP is hard to determine. Our Mean Free Path tree actually omits Akt as a predictor node, so no direct MFP hypothesis can be generated from our existing model based on decreasing Akt. In the Speed tree, interestingly, Akt is the root predictor. The starting condition for this hypothesis (high Akt, high EGFR) predicts medium or high levels of cell speed, depending on the levels of FAK (right-most branch, Figure 5.5C). The Speed tree further predicts that decreasing Akt to medium would not change cell speed, but decreasing it to low would lower cell speed from medium/high to low. Since MFP is the product of speed and persistence time, the resulting MFP from a simultaneous increase in persistence and a decrease in cell speed is difficult to predict. Thus, while this hypothesis is testable with commercially-available Akt inhibitors, its ultimate applications in promoting comprehensive cell migration in tissue engineering scaffolds are not clear.

135

5.3.2.4.   Hypothesis 4: Decrease FAK signaling to increase Speed

Depending on the levels of EGFR, decreasing FAK signaling has opposing effects on cell Speed. This is the first of two hypotheses involving the decrease of FAK. When EGFR levels are medium/high, our Speed tree predicts that decreasing FAK increases speed from medium to high (blue arrow, Figure 5.5C). Our Persistence Time tree predicts that in the presence of medium EGFR, decreasing FAK to low levels would not affect cell Persistence Time - the latter would stay high (blue arrow, Figure 5.5B). This high PT, combined with an increase in Speed, would lead to overall increase in cell MFP. This is the exact prediction that we see in our MFP tree (Figure 5.5A). Decreasing FAK would shift cell MFP from the right-hand branch cluster to the left-hand branch cluster, which when EGFR is medium/high predicts high MFP, showing consistency across our three generated decision tree models.

This consistency lends this hypothesis to be a good candidate for testing, though the hypothesis itself is not entirely novel. FAK phosphorylation tends to correlate with increasing surface adhesiveness due to there being more focal adhesions on highly adhesive surfaces, and the focal adhesions being larger in area. Because EGF-stimulated cell speed versus surface adhesiveness is biphasic, cell speeds are actually low on these highly adhesive conditions, hypothesized to be due to the cell's inability to efficiently detach in the rear [67]. Decreasing FAK effectively shifts cell speed toward the peak of the biphasic curve. With an unchanging Persistence Time, this increase in cell MFP from decreasing FAK essentially recapitulates the biphasic relationship between cell Speed and 2D surface adhesiveness.

136

5.3.2.5. *Hypothesis 5: Decrease FAK signaling to decrease Speed*

This hypothesis is the corollary to Hypothesis 4, predicting the effects of decreasing FAK when EGFR is low. The hypothesis originates from the Speed tree, which predicts that decreasing FAK while EGFR is low will decrease cell speed from high to low (black arrow, Figure 5.5C). This same decrease in FAK while EGFR is low will increase cell PT from low to high (black arrow, Figure 5.5B). As with Akt, we have a change in signaling profile that results in PT and Speed changes that oppositely affect cell MFP. In this case however, FAK is the root predictor in the MFP tree. Decreasing FAK from medium/high to low shifts cell MFP from the right-hand branch to the left-hand branch, and given that EGFR is low, the MFP tree predicts that MFP would also be low (black arrow, Figure 5.5A). Depending on the cell's levels of ERK signaling, this is either a "no change" if ERK is high, or a decrease from medium MFP if ERK is low. This hypothesis is testable, and FAK inhibitors are commercially available.
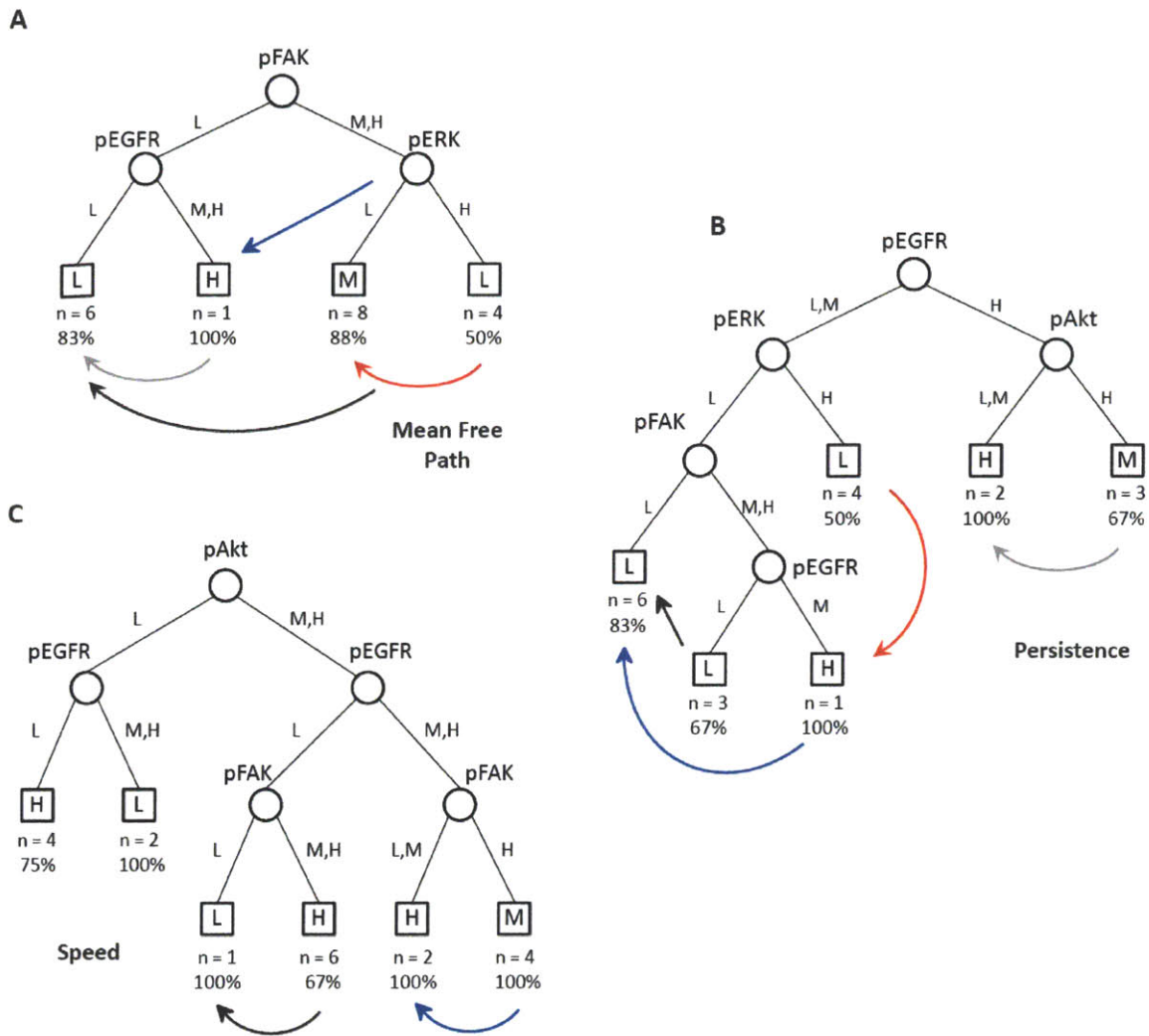
Figure 5.5. Decision trees generate testable hypotheses for predictor changes that can change cell response. Trees shown are for cell (A) Mean Free Path, (B) Persistence Time, and (C) Speed. Arrows graphically show hypotheses generated, and arrows of the same color show the same hypotheses across all three trees.

5.3.3. ERK inhibition increases directional persistence

Hypothesis 1, 3, and 5 were all interesting testable hypotheses generated from our decision tree modeling. Hypothesis 1 presents a prediction that would increase MSCs' ability to traverse and "fill out" a tissue engineering scaffold, though the particular hypothesis itself is very specific. Hypothesis 3 and 5 have questionable tissue engineering applications, though. Hypothesis 3 predicts changes in MSC migration Speed and PT that would cause opposing effects in cell MFP, and since Akt is not a predictor node in our MFP decision tree, its importance in cell MFP is unclear. Hypothesis 5, on the other hand, is a prediction that decreases cell MFP which would decrease MSCs' ability to traverse a scaffold, not useful for our *ex vivo* expansion of MSC application. Given all of this, we decided to test the first hypothesis: ERK inhibition on Cn 3 surfaces with tEGF would increase PT and MFP.

5.3.3.1.   ERK inhibitor dose-response

We screened two MEK-1 inhibitors, CI-1040 (PD 184352) and U0126, to determine usage doses and efficacy. Figure 5.6 shows levels of phosphorylated ERK 30 minutes following MSC stimulation with soluble EGF and treatment with increasing concentrations of CI-1040 and U0126 on tissue culture plastic. We ran two biological replicates for each condition, and the error bars show +/- the standard error of mean (SEM). The two left-most bars are cells without any inhibitor treatment and thus are plotted in dark gray as conditions common to both CI-1040 and U0126. 10 nM of soluble EGF significantly stimulated phospho-ERK levels, even in full serum media, to 4.3-fold of pERK with no EGF stimulation. The two drugs inhibited this ERK stimulation in a comparable manner, though the 0.01 nM treatment with CI-1040 was not tested. 0.01 nM of U0126 had negligible effects on the EGF-stimulated ERK levels, while 0.1 nM of both inhibitors showed small decreases in ERK, and 1 nM of both inhibitors brought ERK signaling back down to basal levels without EGF stimulation. 10 nM of either inhibitor did not further decrease this basal ERK signal.

Since the two drugs showed comparable efficacies, we arbitrarily chose U0126 to inhibit ERK and test our decision-tree generated hypothesis.
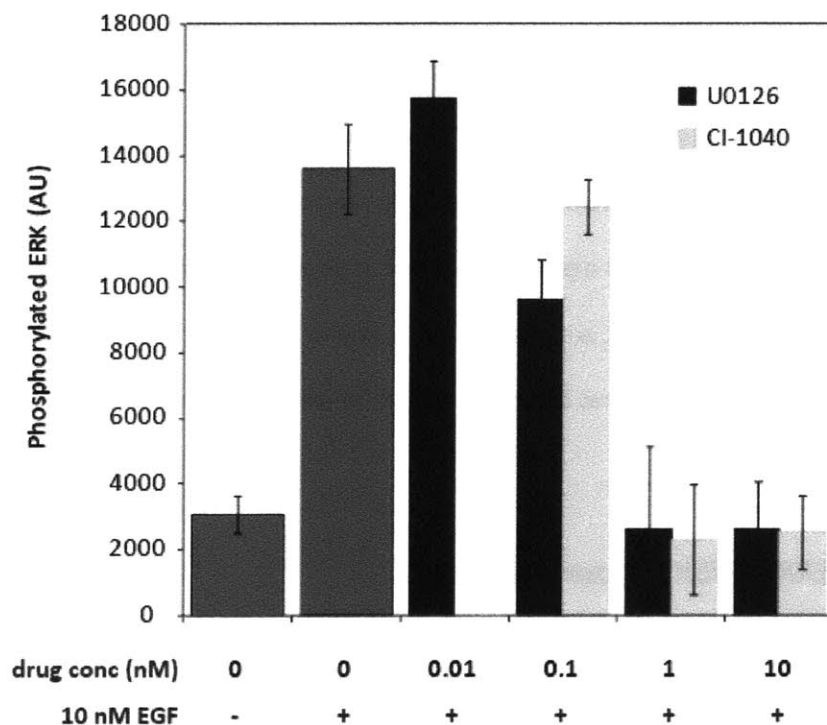
Figure 5.6. MEK inhibitor dose-response. Phosphorylated ERK levels were measured via Luminex. Error bars show +/- SEM from two biological replicates.

5.3.3.2. *MSC migration response with ERK inhibition*

Given the U1026 dose-response, we chose to track migration of MSCs on 3 ug/mL collagen-adsorbed tEGF surfaces treated with 0.3 nM and 1 nM of U1026 to test the effects of partial and full ERK inhibition respectively.

Figure 5.7 shows the effects of ERK inhibition on the migration responses of MFP, Persistence Time, and Speed. Our decision tree models predicted that on tEGF surfaces adsorbed

140

Figure 5.7. Effects of ERK inhibition on average migration responses of (A) Mean Free Path, (B) Persistence Time, and (C) Speed. All experiments were conducted on tEGF copolymer surfaces adsorbed with 3 ug/mL collagen. Error bars show +/- SEM, and statistical significance between bracketed conditions is denoted by * (p < 0.01) and # (p < 0.05).

with 3 ug/mL of collagen, inhibiting ERK signaling would increase cell MFP by increasing cell

Persistence. This is exactly what we see in Figure 5.7. Inhibitng ERK using 0.3 nM of U0126 led to

an increase in MSC MFP from 15.9 um to 23.9 um (50.3% increase), with a p-value of 0.06, just

outside of the acceptable range of statistical significance (Figure 5.7A). Interestingly, inhibiting

ERK even more using a higher drug concentration of 1 nM reversed the trend from no ERK

inhibition. Full ERK inhibition decreased cell MFP from partial inhibition, bringing it down to 18.6

um (28.4% decrease).

This differential effect of partial versus full ERK inhibition on migration was not predicted

by our decision tree model, however. Because time-integrated ERK was so high on collagen

surfaces compared to the others, discretizing of ERK into three data range bins in actuality resulted

in only two bins. Thus, the data used to construct our decision trees only had two levels of ERK:

low versus high. Given that the predicted response from ERK inhibition matched the experimental

results of partial inhibition, this suggests that the "low" ERK in our decision tree may in actuality be

141

"intermediate" ERK, and we did not capture truly low levels of ERK. Another possibility is simply that we did not measure enough conditions and/or signals to capture the differential effects of high versus partial ERK inhibition.

Change in cell Persistence Time with ERK inhibition was also as we predicted (Figure 5.7B). 0.3 nM of drug significantly decreased PT from 16.9 minutes with no inhibition to 22.3 minutes (31.9%). Similar to cell speed, further ERK inhibition with 1 nM of drug reversed the trend and decreased cell PT to 18.3 minutes (21.9% decrease), though this decrease from 0.3 nM of drug was not statistically significant. Our decision tree model of cell speed did not include ERK as a predictor node, and thus could not provide a prediction for ERK inhibition on cell Spped. Figure 5.7C shows that ERK inhibition with intermediate drug concentration decreases cell Speed significantly, from 67.1 um/hour to 62.3 um/hour (7.2%). Full ERK inhibition, just like PT and MFP, reversed this trend recovering cell speed somewhat to 63.5 um/hour. Though only a 2% increase, this change was also statistically significant with a p-value of 0.037.
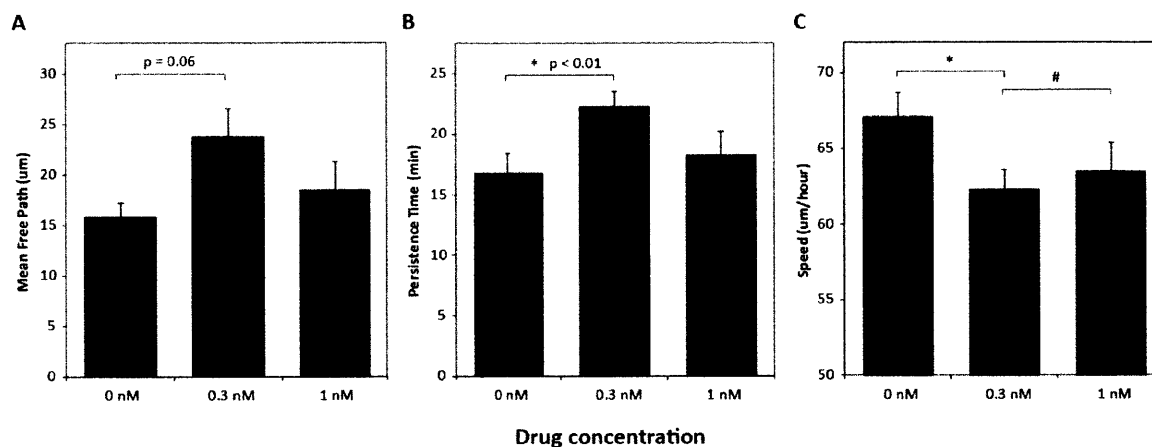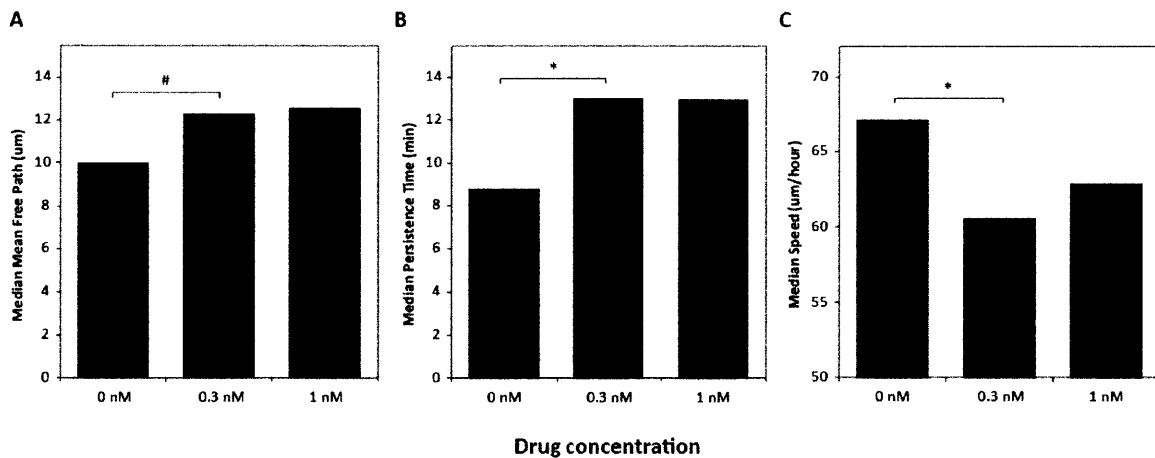

Figure 5.8. Effects of ERK inhibition on median migration responses of (A) Mean Free Path, (B) Persistence Time, and (C) Speed. All experiments were conducted on tEGF copolymer surfaces adsorbed with 3 ug/mL collagen. Statistical significance between bracketed conditions is denoted by * ($p < 0.01$) and # ($p<0.05$).

It is not surprising that inhibiting ERK reduces cell speeds as ERK is linked to contractile

force generation as well as rear detachment [62,139], but what is surprising is that even at full ERK

inhibition, MSC migrate speeds remain over 60 um/hour. This may be partially due to the way that

the Visible analysis determines cell location. Membrane ruffling affects Visible's determination of

cell coordinates, so a stationary cell with highly active membrane activitity may actually have a

non-zero migration speed because the "center of motion" crosshairs are constantly changing.

Under ERK-inhibition, membrane activity may be increased, leading to Visible analyses that

artificially inflate migration speeds. The error bars on the PT and MFP plots are also smaller than

the migration responses reported in Chapter 4. Even then, because PT lends itself to having outliers

and is not a normally distributed parameter within a population of cells, we also plotted the

medians for cell MFP, Persistence Time, and Speed as affected by ERK inhibition (Figure 5.8). The

medians generally trended similarly to the averages. MFP medians increased with increasing drug

concentration, though the increase from no drug to 0.3 nM U0126 was much greater and also

statistically significant ($p = 0.02$). ERK inhibition increased cell Persistence Time and decreased cell

Speed, the same observations from looking the averages of these parameters.

5.3.3.3.   *Effects of ERK inhibition on other signals*

Another assumption of the decision tree generated hypotheses is that the particular signals

involved in the hypothesis are the only ones that are changing. For the particular hypothesis we

tested, we assumed that all other phosphoproteins would stay at the same or similar levels in the

presence of ERK inhibition. Because ERK is such a ubiquitous protein involved in multiple

pathways including those of Akt and FAK signaling, this assumption may not be valid. Thus, we

measured phosphorylated Akt and FAK in MSCs treated with ERK inhibitor U0126 on 3 ug/mL

collagen surfaces with and without tEGF.

Using drug concentrations of 1, 3, and 10 nM, we generally did not observe changes in the levels of Akt and FAK 30 minutes after treatment (Figure 5.9). 3 nM of U0126 increased both Akt and FAK signaling, but this level of inhibitor was ten times the concentration that we used to test our hypothesis. 1 nM of the ERK inhibitor increased Akt about 20% and had no effect on FAK signaling, and even 10 nM of the inhibitor did not affect Akt and FAK signaling. This suggests that for the four phosphoproteins that we measured, inhibiting ERK only changed the signaling matrix **X** by its ERK-related values.

A

B
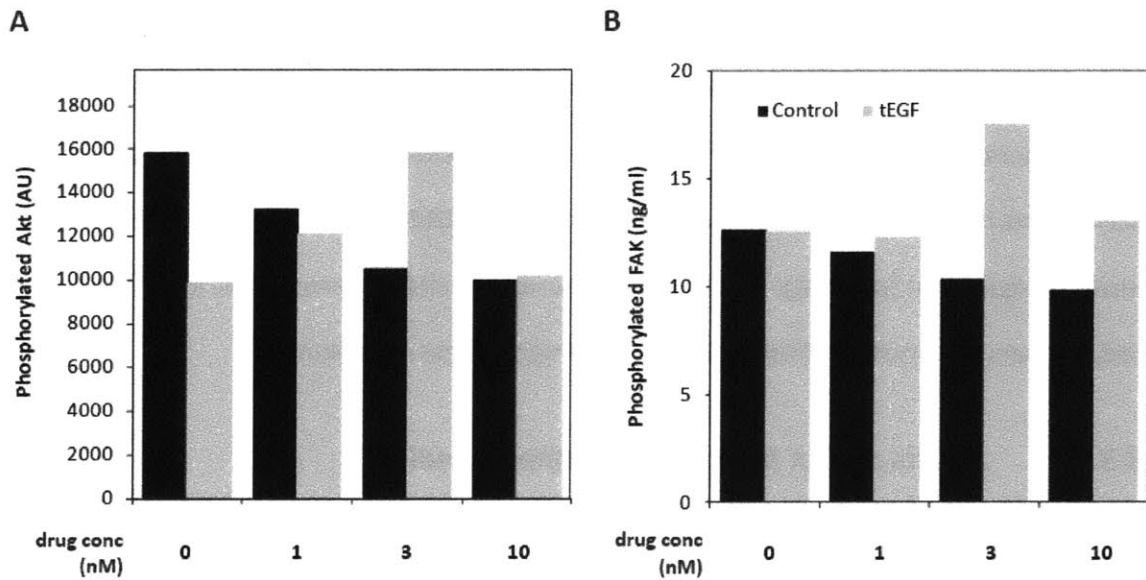
Figure 5.9. Effects of ERK inhibition on MSC (A) Akt and (B) FAK signaling 30 minutes after tEGF treatment. All experiments were conducted on tEGF surfaces adsorbed with 3 ug/mL collagen.

5.4. Summary and conclusions

This chapter discusses the key findings of this thesis:

1) Multivariate combinations of signaling pathways affect MSC migration. Single pathways were uninformative of cell response.

2) EGFR, ERK, Akt, and FAK form a subset of governing proteins that can accurately predict MSC migration in response to tEGF and ECM cues.

3) A combination therapy approach with both biomaterials and small molecules can more effectively and precisely control MSC migration.

While we explicitly explored MSC migration responses in the course of this thesis, these idea and findings are principles that can apply to all studies of cell response. The underlying paradigm of external cues affecting cellular behavior through signaling pathways is true for any behavior, be it migration, proliferation, or survival.

6. Thesis conclusions and future directions of study

Bone marrow derived multipotent stromal cells hold great potential for tissue engineering applications because of their ability to home to injury sites and to differentiate along the multipotent lineages of osteogenic, chondrogenic, ana adipogenic lineages to aid in tissue repair and regeneration. However, their full therapeutic potential is as yet limited due to cell scarcity. Designing biomimetic scaffolds for the *ex vivo* expansion of isolated MSCs before re-integration into the target tissue is one approach to combat this challenge. The effective design of these scaffolds hinges on a comprehensive understanding of MSC biology so that the scaffolds can be engineered to drive MSCs survival, proliferation, migration, and differentiation.

This thesis presents a quantitative and comprehensive study of the effects of immobilized epidermal growth factor and ECM on MSC migration for bone tissue engineering applications. Furthermore, we start to explore the underlying signaling that govern MSC migration. Using a systems-level decision tree analysis to relate MSC signal to migration response, we discovered a non-intuitive therapeutic target in ERK signaling as a potential effector of directional persistence of MSC migration. Together with the ECM and tEGF biomaterials surfaces, we demonstrated a proof of concept for a two-layer combination approach to wound healing therapy: a "coarse" biomaterials control using collagen tEGF surfaces to stimulate reliable MSC motility speeds followed by small

molecules "fine-tuning" of that speed to increase persistence to ultimately result in a more precise control of MSC migration and increase in MSC MFP.

While the work presented here is good start toward understanding MSC migration, cell migration *in vivo* occurs in 3D environments as opposed to the 2D substrates of our current study. Increasing evidence shows that additional biophysical processes are involved with 3D migration versus 2D. For example, cells may need to clear a path through its 3D ECM environment before motility is possible, an irrelevant concern on 2D substrates [140,141]. Cell migration *in vivo* may also progress along "tracks" such as fibers such as collagen [72], another consideration not taken into account in 2D. Subsequently, the signaling pathways discovered to govern cell responses in 2D do not always translate to 3D and/or additional pathways are necessarily to fully encapsulate the signal-response landscape of 3D migration. As we identified independent cues for tuning MSC migration speed and response in this thesis, optimizing for both to be high in the applications of *ex vivo* expansion of MSCs, this may not always be the most desirable outcome particularly in 3D.

Pioneering work theorizes that mechanical and biophysical properties of 3D environments affect cell migration behavior, which was subsequently shown experimentally [142,143]. Specific to MSCs, matrix and scaffold pore size, stiffness, and adhesiveness have all been shown to affect 3D migration speed, but the signaling mechanisms are yet to be explored and migration persistence quantified [144]. Measuring cell signaling in 3D environments is a non-trivial challenge, but nevertheless critical to our full understanding of MSC migration *in vivo* and in 3D scaffolds. This represents a major direction for future studies.

Another limitation to the work discussed here is the usage of an immortalized cell line. While convenient and even ideal for preliminary studies of MSCs, we were unable to fully explore the multitude of signaling pathways that affect migration because many phosphoproteins were undetectable or unchanging in the hTERT-MSCs. While this thesis shows a proof of concept for therapeutic target discovery for controlling migration to benefit bone injury repair, it remains to be

seen if ERK remains a prominent potential target in an expanded landscape of additional measured signals. Using the immortalized cell line allowed us to avoid issues such as senescence at higher passages and population heterogeneities, but the reality is that these are significant challenges in MSC-based therapeutics. Comprehensive understanding of MSC behavior is impossible without studying and addressing these intrinsic characteristics of these cells. Thus, another major direction for future studies is to employ the principles demonstrated in this thesis for the study of primary MSCs.

7. References

[1] M.F. Pittenger, a M. Mackay, S.C. Beck, R.K. Jaiswal, R. Douglas, J.D. Mosca, M. a Moorman, D.W. Simonetti, S. Craig, and D.R. Marshak, "Multilineage potential of adult human mesenchymal stem cells." *Science (New York, N.Y.)*, vol. 284, Apr. 1999, pp. 143-7.

[2] N. Zech, "Adult Stem Cell Manipulation and Possible Clinical Perspectives," *Journal fur Reproductionsmedizin und Endokrinologie*, vol. 1, 2004, pp. 91-99.

[3] A.J. Friedenstein, K.V. Petrakova, A.I. Kurolesova, and G.P. Frolova, "Heterotopic of bone marrow. Analysis of precursor cells for osteogenic and hematopoietic tissues." *Transplantation*, vol. 6, Mar. 1968, pp. 230-47.

[4] J. Tolar, K. Le Blanc, A. Keating, and B.R. Blazar, "Hitting the Right Spot with Mesenchymal Stromal Cells (MSCs)." *Stem cells (Dayton, Ohio)*, Jul. 2010, pp. 1446-1455.

[5] J.T. Williams, S.S. Southerland, J. Souza, A.F. Calcutt, and R.G. Cartledge, "Cells isolated from adult human skeletal muscle capable of differentiating into multiple mesodermal phenotypes." *The American surgeon*, vol. 65, Jan. 1999, pp. 22-6.

[6] A. Erices, P. Conget, and J.J. Minguell, "Mesenchymal progenitor cells in human umbilical cord blood." *British journal of haematology*, vol. 109, Apr. 2000, pp. 235-42.

[7] P.A. Zuk, M. Zhu, P. Ashjian, D.A. De Ugarte, J.I. Huang, H. Mizuno, Z.C. Alfonso, J.K. Fraser, P. Benhaim, and M.H. Hedrick, "Human adipose tissue is a source of multipotent stem cells." *Molecular biology of the cell*, vol. 13, Dec. 2002, pp. 4279-95.

[8] S.A. Kuznetsov, M.H. Mankani, S. Gronthos, K. Satomura, P. Bianco, and P.G. Robey, "Circulating skeletal stem cells." *The Journal of cell biology*, vol. 153, May. 2001, pp. 1133-40.

[9] D.G. Phinney, G. Kopen, W. Righter, S. Webster, N. Tremain, and D.J. Prockop, "Donor variation in the growth properties and osteogenic potential of human marrow stromal cells." *Journal of cellular biochemistry*, vol. 75, Dec. 1999, pp. 424-36.

[10] M. Dominici, K. Le Blanc, I. Mueller, I. Slaper-Cortenbach, F. Marini, D. Krause, R. Deans, A. Keating, D. Prockop, and E. Horwitz, "Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement." *Cytotherapy*, vol. 8, Jan. 2006, pp. 315-7.

[11] R. Meisel, A. Zibert, M. Laryea, U. Göbel, W. Däubener, and D. Dilloo, "Human bone marrow stromal cells inhibit allogeneic T-cell responses by indoleamine 2,3-dioxygenase-mediated tryptophan degradation." *Blood*, vol. 103, Jun. 2004, pp. 4619-21.

[12]   D.J. Prockop, "Repair of tissues by adult stem/progenitor cells (MSCs): controversies, myths, and changing paradigms.," *Molecular therapy : the journal of the American Society of Gene Therapy*, vol. 17, Jun. 2009, pp. 939-46.

[13]   L.G. Menon, K. Kelly, H.W. Yang, S.-K. Kim, P.M. Black, and R.S. Carroll, "Human bone marrow-derived mesenchymal stromal cells expressing S-TRAIL as a cellular delivery vehicle for human glioma therapy.," *Stem cells (Dayton, Ohio)*, vol. 27, Sep. 2009, pp. 2320-30.

[14]   A.J. Nauta and W.E. Fibbe, "Immunomodulatory properties of mesenchymal stromal cells.," *Blood*, vol. 110, Nov. 2007, pp. 3499-506.

[15]   A. Nakamura, M. Akahane, H. Shigematsu, M. Tadokoro, Y. Morita, H. Ohgushi, Y. Dohi, T. Imamura, and Y. Tanaka, "Cell sheet transplantation of cultured mesenchymal stem cells enhances bone formation in a rat nonunion model.," *Bone*, vol. 46, Feb. 2010, pp. 418-24.

[16]   M. Akahane, H. Shigematsu, M. Tadokoro, T. Ueha, T. Matsumoto, Y. Tohma, A. Kido, T. Imamura, and Y. Tanaka, "Scaffold-free cell sheet injection results in bone formation.," *Journal of tissue engineering and regenerative medicine*, vol. 4, Jul. 2010, pp. 404-11.

[17]   H. Dashtdar, H.A. Rothan, T. Tay, R.E. Ahmad, R. Ali, L.X. Tay, P.P. Chong, and T. Kamarul, "A preliminary study comparing the use of allogenic chondrogenic pre-differentiated and undifferentiated mesenchymal stem cells for the repair of full thickness articular cartilage defects in rabbits.," *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, Mar. 2011.

[18]   B. Marquass, R. Schulz, P. Hepp, M. Zscharnack, T. Aigner, S. Schmidt, F. Stein, R. Richter, G. Osterhoff, G. Aust, C. Josten, and A. Bader, "Matrix-Associated Implantation of Predifferentiated Mesenchymal Stem Cells Versus Articular Chondrocytes: In Vivo Results of Cartilage Repair After 1 Year.," *The American journal of sports medicine*, Apr. 2011.

[19]   G.F. Muschler, C. Nakamoto, and L.G. Griffith, "Engineering principles of clinical cell-based tissue engineering.," *The Journal of bone and joint surgery. American volume*, vol. 86-A, Jul. 2004, pp. 1541-58.

[20]   R.H. Lee, A.A. Pulin, M.J. Seo, D.J. Kota, J. Ylostalo, B.L. Larson, L. Semprun-Prieto, P. Delafontaine, and D.J. Prockop, "Intravenous hMSCs improve myocardial infarction in mice because cells embolized in lung are activated to secrete the anti-inflammatory protein TSG-6.," *Cell stem cell*, vol. 5, Jul. 2009, pp. 54-63.

[21]   G.J. Block, S. Ohkouchi, F. Fung, J. Frenkel, C. Gregory, R. Pochampally, G. Dimattia, D.E. Sullivan, and D.J. Prockop, "Multipotent Stromal Cells (MSCs) are Activated to Reduce Apoptosis in Part by Upregulation and Secretion of Stanniocalcin-1 (STC-1).," *Stem cells (Dayton, Ohio)*, Dec. 2008.

[22]   A. Shekaran and A.J. García, "Extracellular matrix-mimetic adhesive biomaterials for bone repair.," *Journal of biomedical materials research. Part A*, vol. 96, Jan. 2011, pp. 261-72.

[23]   K.E. Healy and R.E. Guldberg, "Bone tissue engineering.," *Journal of musculoskeletal & neuronal interactions*, vol. 7, 2007, pp. 328-30.

[24] S. Baba, T. Inoue, Y. Hashimoto, D. Kimura, M. Ueda, K. Sakai, N. Matsumoto, C. Hiwa, T. Adachi, and M. Hojo, "Effectiveness of scaffolds with pre-seeded mesenchymal stem cells in bone regeneration -Assessment of osteogenic ability of scaffolds implanted under the periosteum of the cranial bone of rats-.," *Dental materials journal*, vol. 29, Nov. 2010, pp. 673-81.

[25] B. Dozza, C. Di Bella, E. Lucarelli, G. Giavaresi, M. Fini, P.L. Tazzari, S. Giannini, and D. Donati, "Mesenchymal stem cells and platelet lysate in fibrin or collagen scaffold promote non-cemented hip prosthesis integration.," *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, vol. 29, Jan. 2011, pp. 961-8.

[26] C. Szpalski, J. Barr, M. Wetterau, P.B. Saadeh, and S.M. Warren, "Cranial bone defects: current and future strategies.," *Neurosurgical focus*, vol. 29, Dec. 2010, p. E8.

[27] C. Xu, P. Su, X. Chen, Y. Meng, W. Yu, A.P. Xiang, and Y. Wang, "Biocompatibility and osteogenesis of biomimetic Bioglass-Collagen-Phosphatidylserine composite scaffolds for bone tissue engineering.," *Biomaterials*, vol. 32, Feb. 2011, pp. 1051-8.

[28] A. Chatterjea, A.J.S. Renard, C. Jolink, C.A. van Blitterswijk, and J. de Boer, "Streamlining the generation of an osteogenic graft by 3D culture of unprocessed bone marrow on ceramic scaffolds.," *Journal of tissue engineering and regenerative medicine*, Mar. 2011.

[29] A. Gloria, R. De Santis, and L. Ambrosio, "Polymer-based composite scaffolds for tissue engineering.," *Journal of applied biomaterials & biomechanics : JABB*, vol. 8, pp. 57-67.

[30] A.J.W. Johnson and B.A. Herschler, "A Review of the mechanical behavior of CaP and CaP/polymer composites for applications in bone replacement and repair.," *Acta biomaterialia*, vol. 7, Jul. 2010, pp. 16-30.

[31] J.R. Porter, A. Henson, and K.C. Popat, "Biodegradable poly(epsilon-caprolactone) nanowires for bone tissue engineering applications.," *Biomaterials*, vol. 30, Mar. 2009, pp. 780-8.

[32] F. Wang, Z. Li, K. Tamama, C.K. Sen, and J. Guan, "Fabrication and characterization of prosurvival growth factor releasing, anisotropic scaffolds for enhanced mesenchymal stem cell survival/growth and orientation.," *Biomacromolecules*, vol. 10, Sep. 2009, pp. 2609-18.

[33] Q.P. Pham, F.K. Kasper, L. Scott Baggett, R.M. Raphael, J.A. Jansen, and A.G. Mikos, "The influence of an in vitro generated bone-like extracellular matrix on osteoblastic gene expression of marrow stromal cells.," *Biomaterials*, vol. 29, Jun. 2008, pp. 2729-39.

[34] M. Rodrigues, L.G. Griffith, and A. Wells, "Growth factor regulation of proliferation and survival of multipotential stromal cells.," *Stem cell research & therapy*, vol. 1, Jan. 2010, p. 32.

[35] A. Malgieri, E. Kantzari, M.P. Patrizi, and S. Gambardella, "Bone marrow and umbilical cord blood human mesenchymal stem cells: state of the art.," *International journal of clinical and experimental medicine*, vol. 3, Jan. 2010, pp. 248-69.

[36] J.M. Curran, F. Pu, R. Chen, and J.A. Hunt, "The use of dynamic surface chemistries to control msc isolation and function.," *Biomaterials*, Apr. 2011.

[37] C.M. Kemmis, A. Vahdati, H.E. Weiss, and D.R. Wagner, "Bone morphogenetic protein 6 drives both osteogenesis and chondrogenesis in murine adipose-derived mesenchymal cells depending on culture conditions.," *Biochemical and biophysical research communications*, vol. 401, Oct. 2010, pp. 20-5.

[38] S. Weiss, T. Hennig, R. Bock, E. Steck, and W. Richter, "Impact of growth factors and PTHrP on early and late chondrogenic differentiation of human mesenchymal stem cells.," *Journal of cellular physiology*, vol. 223, Apr. 2010, pp. 84-93.

[39] S. Tsutsumi, A. Shimazu, K. Miyazaki, H. Pan, C. Koike, E. Yoshida, K. Takagishi, and Y. Kato, "Retention of multilineage differentiation potential of mesenchymal cells during proliferation in response to FGF.," *Biochemical and biophysical research communications*, vol. 288, Oct. 2001, pp. 413-9.

[40] J. Farré, S. Roura, C. Prat-Vidal, C. Soler-Botija, A. Llach, C.E. Molina, L. Hove-Madsen, J.J. Cairó, F. Gòdia, R. Bragós, J. Cinca, and A. Bayes-Genis, "FGF-4 increases in vitro expansion rate of human adult bone marrow-derived mesenchymal stem cells.," *Growth factors (Chur, Switzerland)*, vol. 25, Apr. 2007, pp. 71-6.

[41] A. Schmidt, D. Ladage, T. Schinkothe, U. Klausmann, C. Ulrichs, F.-J. Klinz, K. Brixius, S. Arnhold, B. Desai, U. Mehlhorn, R.H.G. Schwinger, P. Staib, K. Addicks, and W. Bloch, "Basic fibroblast growth factor controls migration in human mesenchymal stem cells," *Stem Cells*, vol. 24, 2006, pp. 1750-1758.

[42] J. Fiedler, N. Etzel, and R.E. Brenner, "To go or not to go: Migration of human mesenchymal progenitor cells stimulated by isoforms of PDGF.," *Journal of cellular biochemistry*, vol. 93, Nov. 2004, pp. 990-8.

[43] B. Krausgrill, M. Vantler, V. Burst, M. Raths, M. Halbach, K. Frank, S. Schynkowski, K. Schenk, J. Hescheler, S. Rosenkranz, and J. Müller-Ehmsen, "Influence of cell treatment with PDGF-BB and reperfusion on cardiac persistence of mononuclear and mesenchymal bone marrow cells after transplantation into acute myocardial infarction in rats.," *Cell transplantation*, vol. 18, Jan. 2009, pp. 847-53.

[44] G. Forte, M. Minieri, P. Cossa, D. Antenucci, M. Sala, V. Gnocchi, R. Fiaccavento, F. Carotenuto, P. De Vito, P.M. Baldini, M. Prat, and P. Di Nardo, "Hepatocyte growth factor effects on mesenchymal stem cells: proliferation, migration, and differentiation.," *Stem cells (Dayton, Ohio)*, vol. 24, Jan. 2006, pp. 23-33.

[45] K. Tamama, V.H. Fan, L.G. Griffith, H.C. Blair, and A. Wells, "Epidermal growth factor as a candidate for ex vivo expansion of bone marrow-derived mesenchymal stem cells.," *Stem cells (Dayton, Ohio)*, vol. 24, Mar. 2006, pp. 686-95.

[46] M. Krampera, A. Pasini, A. Rigo, M.T. Scupoli, C. Tecchio, G. Malpeli, A. Scarpa, F. Dazzi, G. Pizzolo, and F. Vinante, "HB-EGF / HER-1 signaling in bone marrow mesenchymal stem cells : inducing cell expansion and reversibly preventing multilineage differentiation," *Cell*, vol. 106, 2005, pp. 59-66.

[47] I. Martin, A. Muraglia, G. Campanile, R. Cancedda, and R. Quarto, "Fibroblast growth factor-2 supports ex vivo expansion and maintenance of osteogenic precursors from human bone marrow.," *Endocrinology*, vol. 138, Oct. 1997, pp. 4456-62.

[48] J. Fiedler, C. Brill, W.F. Blum, and R.E. Brenner, "IGF-I and IGF-II stimulate directed cell migration of bone-marrow-derived human mesenchymal progenitor cells.," *Biochemical and biophysical research communications*, vol. 345, Jul. 2006, pp. 1177-83.

[49] J. Fiedler, F. Leucht, J. Waltenberger, C. Dehio, and R.E. Brenner, "VEGF-A and PlGF-1 stimulate chemotactic migration of human mesenchymal progenitor cells.," *Biochemical and biophysical research communications*, vol. 334, Aug. 2005, pp. 561-8.

[50] M.E. Lukashev and Z. Werb, "ECM signalling: orchestrating cell behaviour and misbehaviour.," *Trends in cell biology*, vol. 8, Nov. 1998, pp. 437-41.

[51] D.J. Behonick and Z. Werb, "A bit of give and take: the relationship between the extracellular matrix and the developing chondrocyte.," *Mechanisms of development*, vol. 120, Nov. 2003, pp. 1327-36.

[52] S.-H. Kim, J.E. Turnbull, and S.E. Guimond, "Extracellular matrix and cell signalling - the dynamic cooperation of integrin, proteoglycan and growth factor receptor.," *The Journal of endocrinology*, Feb. 2011.

[53] D.S. Harburger and D. a Calderwood, "Integrin signalling at a glance.," *Journal of cell science*, vol. 122, Jan. 2009, pp. 159-63.

[54] A.K. Kundu and A.J. Putnam, "Vitronectin and collagen I differentially regulate osteogenesis in mesenchymal stem cells.," *Biochemical and biophysical research communications*, vol. 347, Aug. 2006, pp. 347-57.

[55] R.M. Salasznyk, R.F. Klees, W. a Williams, A. Boskey, and G.E. Plopper, "Focal adhesion kinase signaling pathways regulate the osteogenic differentiation of human mesenchymal stem cells.," *Experimental cell research*, vol. 313, Jan. 2007, pp. 22-37.

[56] A. Ode, G.N. Duda, J.D. Glaeser, G. Matziolis, S. Frauenschuh, C. Perka, C.J. Wilson, and G. Kasper, "Toward biomimetic materials in bone regeneration: Functional behavior of mesenchymal stem cells on a broad spectrum of extracellular matrix components.," *Journal of biomedical materials research. Part A*, vol. 95, Sep. 2010, pp. 1114-24.

[57] M. Thibault, C. Hoemann, and M. Buschmann, "Fibronectin, vitronectin, and collagen I induce chemotaxis and haptotaxis of human and rabbit mesenchymal stem cells in a standardized transmembrane assay.," *Stem Cells and DevelopmentDevelopment*, vol. 16, 2007, pp. 489-502.

[58] E. Spaeth, a Klopp, J. Dembinski, M. Andreeff, and F. Marini, "Inflammation and tumor microenvironments: defining the migratory itinerary of mesenchymal stem cells.," *Gene therapy*, vol. 15, May. 2008, pp. 730-8.

155

[59] H. Yagi, A. Soto-Gutierrez, B. Parekkadan, Y. Kitagawa, R.G. Tompkins, N. Kobayashi, and M.L. Yarmush, "Mesenchymal Stem Cells: Mechanisms of Immunomodulation and Homing.," *Cell transplantation*, vol. 19, Jun. 2010, pp. 667-679.

[60] D.A. Lauffenburger and A.F. Horwitz, "Cell migration: a physically integrated molecular process.," *Cell*, vol. 84, Feb. 1996, pp. 359-69.

[61] A. Wells, J. Kassis, J. Solava, T. Turner, and D.A. Lauffenburger, "Growth factor-induced cell motility in tumor invasion.," *Acta oncologica (Stockholm, Sweden)*, vol. 41, Jan. 2002, pp. 124-30.

[62] A.J. Ridley, M.A. Schwartz, K. Burridge, R.A. Firtel, M.H. Ginsberg, G. Borisy, J.T. Parsons, and A.R. Horwitz, "Cell migration: integrating signals from front to back.," *Science (New York, N.Y.)*, vol. 302, Dec. 2003, pp. 1704-9.

[63] M. Raftopoulou and A. Hall, "Cell migration: Rho GTPases lead the way," *Developmental Biology*, vol. 265, Jan. 2004, pp. 23-32.

[64] M.J. Kim, E. Kim, S.H. Ryu, and P.G. Suh, "The mechanism of phospholipase C-gamma1 regulation.," *Experimental & molecular medicine*, vol. 32, Sep. 2000, pp. 101-9.

[65] M.F. Fournier, R. Sauser, D. Ambrosi, J.-J. Meister, and A.B. Verkhovsky, "Force transmission in migrating cells.," *The Journal of cell biology*, vol. 188, Jan. 2010, pp. 287-97.

[66] P. a DiMilla, K. Barbee, and D. a Lauffenburger, "Mathematical model for the effects of adhesion and mechanics on cell migration speed.," *Biophysical journal*, vol. 60, Jul. 1991, pp. 15-37.

[67] G. Maheshwari, a Wells, L.G. Griffith, and D. a Lauffenburger, "Biophysical integration of effects of epidermal growth factor and fibronectin on fibroblast migration.," *Biophysical journal*, vol. 76, May. 1999, pp. 2814-23.

[68] A.A. Brahmbhatt and R.L. Klemke, "ERK and RhoA differentially regulate pseudopodia growth and retraction during chemotaxis.," *The Journal of biological chemistry*, vol. 278, Apr. 2003, pp. 13016-25.

[69] S.J. Franco and A. Huttenlocher, "Regulating cell migration: calpains make the cut.," *Journal of cell science*, vol. 118, Sep. 2005, pp. 3829-38.

[70] C. Huang, K. Jacobson, and M.D. Schaller, "MAP kinases and cell migration.," *Journal of cell science*, vol. 117, Sep. 2004, pp. 4619-28.

[71] J. Kassis, D.A. Lauffenburger, T. Turner, and A. Wells, "Tumor invasion as dysregulated cell motility.," *Seminars in cancer biology*, vol. 11, Apr. 2001, pp. 105-17.

[72] H. Yamaguchi, J. Wyckoff, and J. Condeelis, "Cell migration in tumors.," *Current opinion in cell biology*, vol. 17, Oct. 2005, pp. 559-64.

[73] E.G. Levin, "Cancer therapy through control of cell migration.," *Current cancer drug targets*, vol. 5, Nov. 2005, pp. 505-18.

[74] Y.Y. Zaytseva, N.K. Wallis, R.C. Southard, and M.W. Kilgore, "The PPAR{gamma} Antagonist T0070907 Suppresses Breast Cancer Cell Proliferation and Motility via Both PPAR{gamma}-dependent and -independent Mechanisms.," *Anticancer research*, vol. 31, Mar. 2011, pp. 813-23.

[75] M.K. Shanmugam, K.A. Manu, T.H. Ong, L. Ramachandran, R. Surana, P. Bist, L.H.K. Lim, A.P. Kumar, K.M. Hui, and G. Sethi, "Inhibition of CXCR4/CXCL12 signaling axis by ursolic acid leads to suppression of metastasis in transgenic adenocarcinoma of mouse prostate model.," *International journal of cancer. Journal international du cancer*, Apr. 2011.

[76] A. Liu, H. Chen, W. Wei, S. Ye, W. Liao, J. Gong, Z. Jiang, L. Wang, and S. Lin, "Antiproliferative and antimetastatic effects of emodin on human pancreatic cancer.," *Oncology reports*, Apr. 2011.

[77] T. Ideker, "Building with a scaffold: emerging strategies for high- to low-level cellular modeling," *Trends in Biotechnology*, vol. 21, Jun. 2003, pp. 255-262.

[78] S.J. Vayttaden, S.M. Ajay, and U.S. Bhalla, "A spectrum of models of signaling pathways.," *Chembiochem : a European journal of chemical biology*, vol. 5, Oct. 2004, pp. 1365-74.

[79] K. a Janes and D. a Lauffenburger, "A biological approach to computational models of proteomic networks.," *Current opinion in chemical biology*, vol. 10, Feb. 2006, pp. 73-80.

[80] B. Chakravarti, B. Mallik, and D.N. Chakravarti, "Proteomics and systems biology: application in drug discovery and development.," *Methods in molecular biology (Clifton, N.J.)*, vol. 662, Jan. 2010, pp. 3-28.

[81] S.L. Porter, G.H. Wadhams, and J.P. Armitage, "Signal processing in complex chemotaxis pathways," *Nature Reviews Microbiology*, vol. 9, Feb. 2011, pp. 153-65.

[82] S. Hautaniemi, S. Kharait, A. Iwabu, A. Wells, and D.A. Lauffenburger, "Modeling of signal-response cascades using decision tree analysis.," *Bioinformatics (Oxford, England)*, vol. 21, May. 2005, pp. 2027-35.

[83] A. Iwabu, K. Smith, F.D. Allen, D.A. Lauffenburger, and A. Wells, "Epidermal growth factor induces fibroblast contractility and motility via a protein kinase C delta-dependent pathway.," *The Journal of biological chemistry*, vol. 279, Apr. 2004, pp. 14551-60.

[84] D.N. Jackson and D.A. Foster, "The enigmatic protein kinase Cdelta: complex roles in cell proliferation and survival.," *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, vol. 18, Apr. 2004, pp. 627-36.

[85] S. Kharait, S. Hautaniemi, S. Wu, A. Iwabu, D.A. Lauffenburger, and A. Wells, "Decision tree modeling predicts effects of inhibiting contractility signaling on cell motility.," *BMC systems biology*, vol. 1, Jan. 2007, p. 9.

[86]   E. Jones and X. Yang, "Mesenchymal stem cells and bone regeneration: Current status.," *Injury*, Apr. 2011.

[87]   F. Anjos-Afonso and D. Bonnet, "Prospective identification and isolation of murine bone marrow derived multipotent mesenchymal progenitor cells.," *Best practice & research. Clinical haematology*, vol. 24, Mar. 2011, pp. 13-24.

[88]   G.J. Rho, B.M. Kumar, and S.S. Balasubramanian, "Porcine mesenchymal stem cells--current technological status and future perspective.," *Frontiers in bioscience : a journal and virtual library*, vol. 14, Jan. 2009, pp. 3942-61.

[89]   N. Maheshri and E.K. O Shea, "Living with noisy genes: how cells function reliably with inherent variability in gene expression.," *Annual review of biophysics and biomolecular structure*, vol. 36, Jan. 2007, pp. 413-34.

[90]   C. Beauséjour, "Bone marrow-derived cells: the influence of aging and cellular senescence.," *Handbook of experimental pharmacology*, Jan. 2007, pp. 67-88.

[91]   V. Vacanti, E. Kong, G. Suzuki, K. Sato, J.M. Canty, and T. Lee, "Phenotypic changes of adult porcine mesenchymal stem cells induced by prolonged passaging in culture.," *Journal of cellular physiology*, vol. 205, Nov. 2005, pp. 194-201.

[92]   W. Wagner, A.D. Ho, and M. Zenke, "Different facets of aging in human mesenchymal stem cells.," *Tissue engineering. Part B, Reviews*, vol. 16, Aug. 2010, pp. 445-53.

[93]   T. Okamoto, T. Aoyama, T. Nakayama, T. Nakamata, T. Hosaka, K. Nishijo, T. Nakamura, T. Kiyono, and J. Toguchida, "Clonal heterogeneity in differentiation potential of immortalized human mesenchymal stem cells.," *Biochemical and biophysical research communications*, vol. 295, Jul. 2002, pp. 354-61.

[94]   M. Kobune, Y. Kawano, Y. Ito, H. Chiba, K. Nakamura, H. Tsuda, K. Sasaki, H. Dehari, H. Uchida, O. Honmou, S. Takahashi, A. Bizen, R. Takimoto, T. Matsunaga, J. Kato, K. Kato, K. Houkin, Y. Niitsu, and H. Hamada, "Telomerized human multipotent mesenchymal cells can differentiate into hematopoietic and cobblestone area-supporting cells.," *Experimental hematology*, vol. 31, Aug. 2003, pp. 715-22.

[95]   J.L. Simonsen, C. Rosada, N. Serakinci, J. Justesen, K. Stenderup, S.I.S. Rattan, T.G. Jensen, and M. Kassem, "Telomerase expression extends the proliferative life-span and maintains the osteogenic potential of human bone marrow stromal cells.," *Nature biotechnology*, vol. 20, Jun. 2002, pp. 592-6.

[96]   B.M. Abdallah, M. Haack-Sørensen, J.S. Burns, B. Elsnab, F. Jakob, P. Hokland, and M. Kassem, "Maintenance of differentiation potential of human bone marrow mesenchymal stem cells immortalized by human telomerase reverse transcriptase gene despite [corrected] extensive proliferation.," *Biochemical and biophysical research communications*, vol. 326, Jan. 2005, pp. 527-38.

[97] D.J. Irvine, a V. Ruzette, a M. Mayes, and L.G. Griffith, "Nanoscale clustering of RGD peptides at surfaces using comb polymers. 2. Surface segregation of comb polymers in polylactide.," *Biomacromolecules*, vol. 2, Jan. 2001, pp. 545-56.

[98] D.J. Irvine, a M. Mayes, and L.G. Griffith, "Nanoscale clustering of RGD peptides at surfaces using Comb polymers. 1. Synthesis and characterization of Comb thin films.," *Biomacromolecules*, vol. 2, Jan. 2001, pp. 85-94.

[99] D. Leckband, S. Sheth, and A. Halperin, "Grafted poly(ethylene oxide) brushes as nonfouling surface coatings." *Journal of biomaterials science. Polymer edition*, vol. 10, Jan. 1999, pp. 1125-47.

[100] A. Au, C. a Boehm, A.M. Mayes, G.F. Muschler, and L.G. Griffith, "Formation of osteogenic colonies on well-defined adhesion peptides by freshly isolated human marrow cells.," *Biomaterials*, vol. 28, Apr. 2007, pp. 1847-61.

[101] A. Au, "Molecular Substrate Design for the Selective Adhesion, Proliferation and Differentiation of Marrow Connective Tissue Progenitors," Massachusetts Institute of Technology, 2005.

[102] R.B. Dickinson and R.T. Tranquillo, "Optimal estimation of cell movement indices from the statistical analysis of cell tracking data," *AIChE Journal*, vol. 39, Dec. 1993, pp. 1995-2010.

[103] G.A. Dunn, "Characterizing a kinesis response: Time averaged measures of cell speed and directional persistence," *Agents and actions supplements*, vol. 12, 1983, pp. 14-33.

[104] D.A. Lauffenburger and J.J. Linderman, *Receptors*, New York: Oxford University Press, 1993.

[105] D.S. Rhoads and J.-L. Guan, "Analysis of directional cell migration on defined FN gradients: role of intracellular signaling molecules." *Experimental cell research*, vol. 313, Nov. 2007, pp. 3859-67.

[106] M.F. Ware, A. Wells, and D.A. Lauffenburger, "Epidermal growth factor alters fibroblast migration speed and directional persistence reciprocally and in a matrix-dependent manner," *Journal of Cell Science*, vol. 2432, 1998, pp. 2423-2432.

[107] N. Kumar, M.H. Zaman, H.-D. Kim, and D.A. Lauffenburger, "A high-throughput migration assay reveals HER2-mediated cell migration arising from increased directional persistence.," *Biophysical journal*, vol. 91, Aug. 2006, pp. L32-4.

[108] H.-D. Kim, T.W. Guo, A.P. Wu, A. Wells, F.B. Gertler, and D.A. Lauffenburger, "Epidermal growth factor-induced enhancement of glioblastoma cell migration in 3D arises from an intrinsic increase in speed but an extrinsic matrix- and proteolysis-dependent increase in persistence.," *Molecular biology of the cell*, vol. 19, Oct. 2008, pp. 4249-59.

[109] X.-D. Chen, "Extracellular matrix provides an optimal niche for the maintenance and propagation of mesenchymal stem cells.," *Birth defects research. Part C, Embryo today : reviews*, vol. 90, Mar. 2010, pp. 45-54.

[110] J.D. Humphries, A. Byron, and M.J. Humphries, "Integrin ligands at a glance.," *Journal of cell science*, vol. 119, Oct. 2006, pp. 3901-3.

[111] A. Sorkin and L.K. Goh, "Endocytosis and intracellular trafficking of ErbBs.," *Experimental cell research*, vol. 315, Mar. 2009, pp. 683-96.

[112] I.H. Madshus and E. Stang, "Internalization and intracellular sorting of the EGF receptor: a model for understanding the mechanisms of receptor trafficking.," *Journal of cell science*, vol. 122, Oct. 2009, pp. 3433-9.

[113] D.A. Puleo, R.A. Kissling, and M.S. Sheu, "A technique to immobilize bioactive proteins, including bone morphogenetic protein-4 (BMP-4), on titanium alloy.," *Biomaterials*, vol. 23, May. 2002, pp. 2079-87.

[114] H. Tsujigiwa, H. Nagatsuka, M. Gunduz, A. Rodriguez, R.S. Rivera, R.Z. Legeros, M. Inoue, and N. Nagai, "Effects of immobilized recombinant human bone morphogenetic protein-2/succinylated type I atelocollagen on cellular activity of ST2 cells.," *Journal of biomedical materials research. Part A*, vol. 75, Oct. 2005, pp. 210-5.

[115] H. Schliephake, A. Aref, D. Scharnweber, S. Bierbaum, S. Roessler, and A. Sewing, "Effect of immobilized bone morphogenic protein 2 coating of titanium implants on peri-implant bone formation.," *Clinical oral implants research*, vol. 16, Oct. 2005, pp. 563-9.

[116] U. Fischer, U. Hempel, D. Becker, S. Bierbaum, D. Scharnweber, H. Worch, and K.-W. Wenzel, "Transforming growth factor beta1 immobilized adsorptively on Ti6Al4V and collagen type I coated Ti6Al4V maintains its biological activity.," *Biomaterials*, vol. 24, Jul. 2003, pp. 2631-41.

[117] B.K. Mann, R.H. Schmedlen, and J.L. West, "Tethered-TGF-beta increases extracellular matrix production of vascular smooth muscle cells.," *Biomaterials*, vol. 22, Mar. 2001, pp. 439-44.

[118] V.H. Fan, K. Tamama, A. Au, R. Littrell, L.B. Richardson, J.W. Wright, A. Wells, and L.G. Griffith, "Tethered epidermal growth factor provides a survival advantage to mesenchymal stem cells.," *Stem cells (Dayton, Ohio)*, vol. 25, May. 2007, pp. 1241-51.

[119] M.O. Platt, A.J. Roman, A. Wells, D.A. Lauffenburger, and L.G. Griffith, "Sustained epidermal growth factor receptor levels and activation by tethered ligand binding enhances osteogenic differentiation of multi-potent marrow stromal cells.," *Journal of cellular physiology*, vol. 221, Nov. 2009, pp. 306-17.

[120] Y. Ito, "Covalently immobilized biosignal molecule materials for tissue engineering," *Soft Matter*, vol. 4, 2008, p. 46.

[121] C.S. Swindle, K.T. Tran, T.D. Johnson, P. Banerjee, A.M. Mayes, L. Griffith, and A. Wells, "Epidermal growth factor (EGF)-like repeats of human tenascin-C as ligands for EGF receptor.," *The Journal of cell biology*, vol. 154, Jul. 2001, pp. 459-68.

[122] "The NCBI handbook [Internet]," Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information, , p. Chapter 18.

[123] L.B. Richardson, "EGF Receptor-mediated Fibroblast Signaling and Motility: Role Of Nanoscale Spatial Ligand Organization," Massachusetts Institute of Technology, 2005.

[124] E.J. Joslin, L.K. Opresko, A. Wells, H.S. Wiley, and D.A. Lauffenburger, "EGF-receptor-mediated mammary epithelial cell migration is driven by sustained ERK signaling from autocrine stimulation," *Journal of cell science*, vol. 120, Oct. 2007, pp. 3688-99.

[125] A.P. Popova, P.D. Bozyk, A.M. Goldsmith, M.J. Linn, J. Lei, J.K. Bentley, and M.B. Hershenson, "Autocrine production of TGF-beta1 promotes myofibroblastic differentiation of neonatal lung mesenchymal stem cells," *American journal of physiology. Lung cellular and molecular physiology*, vol. 298, Jun. 2010, pp. L735-43.

[126] Z. Hamidouche, O. Fromigué, U. Nuber, P. Vaudin, J.-C. Pages, R. Ebert, F. Jakob, H. Miraoui, and P.J. Marie, "Autocrine fibroblast growth factor 18 mediates dexamethasone-induced osteogenic differentiation of murine mesenchymal stem cells," *Journal of cellular physiology*, vol. 224, Aug. 2010, pp. 509-15.

[127] R.J. Petrie, A.D. Doyle, and K.M. Yamada, "Random versus directionally persistent cell migration," *Nature reviews. Molecular cell biology*, vol. 10, Aug. 2009, pp. 538-49.

[128] A.B. Fielding and S. Dedhar, "The mitotic functions of integrin-linked kinase," *Cancer metastasis reviews*, vol. 28, Jul. 2009, pp. 99-111.

[129] L.C. Kim, L. Song, and E.B. Haura, "Src kinases as therapeutic targets for cancer," *Nature reviews. Clinical oncology*, vol. 6, Oct. 2009, pp. 587-95.

[130] M. Vicente-Manzanares, X. Ma, R.S. Adelstein, and A.R. Horwitz, "Non-muscle myosin II takes centre stage in cell adhesion and migration," *Nature reviews. Molecular cell biology*, vol. 10, Nov. 2009, pp. 778-90.

[131] D.D. Schlaepfer, C.R. Hauck, and D.J. Sieg, "Signaling through focal adhesion kinase," *Progress in biophysics and molecular biology*, vol. 71, Jan. 1999, pp. 435-78.

[132] S. Narumiya, M. Tanji, and T. Ishizaki, "Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion," *Cancer metastasis reviews*, vol. 28, Jun. 2009, pp. 65-76.

[133] K.M. Yamada and S. Even-ram, "Integrin regulation of growth factor signalling and adhesion," *Nature Cell Biology*, vol. 4, 2002.

[134] P. Rajasethupathy, S.J. Vayttaden, and U.S. Bhalla, "Systems modeling: a pathway to drug discovery," *Current opinion in chemical biology*, vol. 9, Aug. 2005, pp. 400-6.

[135] A.S. Azmi, Z. Wang, P.A. Philip, R.M. Mohammad, and F.H. Sarkar, "Proof of Concept: A review on how network and systems biology approaches aid in the discovery of potent anticancer drug combinations," *Molecular cancer therapeutics*, vol. 9, Nov. 2010, pp. 3137-44.

[136] P. Geurts, A. Irrthum, and L. Wehenkel, "Supervised learning with decision tree-based methods in computational and systems biology," *Molecular bioSystems*, vol. 5, Dec. 2009, pp. 1593-605.

[137] C. Kingsford and S.L. Salzberg, "What are decision trees?," *Nature biotechnology*, vol. 26, Sep. 2008, pp. 1011-3.

[138] L. Breiman, J. Friedman, C. Olshen, and C. Stone, *Classification and Regression Trees*, Boca Raton, FL: CRC Press, 1984.

[139] A.J. Ridley, "Pulling back to move forward.," *Cell*, vol. 116, Feb. 2004, pp. 357-8.

[140] P. Friedl and K. Wolf, "Tumour-cell invasion and migration: diversity and escape mechanisms.," *Nature reviews. Cancer*, vol. 3, May. 2003, pp. 362-74.

[141] M. Ehrbar, A. Sala, P. Lienemann, A. Ranga, K. Mosiewicz, A. Bittermann, S.C. Rizzi, F.E. Weber, and M.P. Lutolf, "Elucidating the Role of Matrix Stiffness in 3D Cell Migration and Remodeling.," *Biophysical journal*, vol. 100, Jan. 2011, pp. 284-93.

[142] M.H. Zaman, R.D. Kamm, P. Matsudaira, and D.A. Lauffenburger, "Computational model for cell migration in three-dimensional matrices.," *Biophysical journal*, vol. 89, Aug. 2005, pp. 1389-97.

[143] M.H. Zaman, L.M. Trapani, A.L. Sieminski, A. Siemeski, D. Mackellar, H. Gong, R.D. Kamm, A. Wells, D.A. Lauffenburger, and P. Matsudaira, "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*, vol. 103, Jul. 2006, pp. 10889-94.

[144] S.R. Peyton, Z.I. Kalcioglu, J.C. Cohen, A.P. Runkle, K.J. Van Vliet, D.A. Lauffenburger, and L.G. Griffith, "Marrow-Derived stem cell motility in 3D synthetic scaffold is governed by geometry along with adhesivity and stiffness.," *Biotechnology and bioengineering*, vol. 108, May. 2011, pp. 1181-93.

8.   Appendix – PRW Matlab Code

The following sections provide the Matlab code written to implement the Persistent

Random Walk model. Algorithm flow is as follows: `PersFit()` is the main function (PersFit.m)

which prompts the user for an Excel spreadsheet with cell coordinates data. Two columns of

coordinate data exist for each cell in the spreadsheet, with the first column corresponding to the

time-adjacent changes in x-coordinates, and the second column the changes in y-coordinates.

`PersFit` in turn calls `msdnonoverlap` and `msd_overlap` to calculate the Mean Square

Displacement of each cell using both non-overlapping and overlapping intervals (see Chapter 4). It

then fits for Persistence Time values using linear least squares, calling on `prwfuncalc_SW`, the

PRW equation that describes the relationship between all of the parameters of the PRW model.

8.1. PersFit.m

```
% PersFit - calcluates speed, persistence according to PRW model
% Shan Wu
% Adapted from original code by Brian Harms, HD Kim

clear all;
close all;

global Speed i
global numPoints


% constants
PX2UM = 0.65789; % 1 pixel = 0.65789 um
xmin = -100;
xmax = 100;
ymin = -100;
ymax = 100;
tint = 10; % 10 minutes between time points
coordMin = 7; % min # of timepoints for each cell to include

% % read in the excel file
```

```matlab
% filename = input('\nEnter the excel file: \n','s');
% rawdata = xlsread(filename);

directory_name = uigetdir('d:\research\MatLab code\PRW\');

% read in the excel file
filename = input('\nEnter the excel file: \n','s');
% sheetname = input('\nEnter the worksheet name: \n','s');

rawdata = xlsread([directory_name,'\', filename,], 'Sheet1');
[PATHSTR,NAME,EXT,VERSN] = fileparts(filename);

[r, c] = size(rawdata);
totCells = c/2;                    % total number of cells
totCellInPlot = 0;                 % total # cells in wind-rose plot
numPoints = zeros(1,totCells);     % total # data points for each cell

timeIntLength = [1:r]';
timeIntLength = timeIntLength*tint;

figure('Name', 'wind-rose plot', 'NumberTitle','off');

% separate data into cells only if that cells has more than coordMin
% coordinates
j = 1;
for i = 1:totCells

    % truncate all NaN dxdy's for each cell
    % all cells start at origin (0,0)
    I = find(isnan(rawdata(:,i*2-1))==0);
    if (length(I) < coordMin)
        continue;
    end

    numPoints(j) = length(I);
    cells{j} = [rawdata(I,i*2-1:i*2)]; %0 0;
    j = j+1;

end

for i = 1:length(cells)

    i
    time = [];

    % Go to next iteration of non-zero points < threshold
    if (numPoints(i) < coordMin)
        continue;
    end

    % set distance traveled initially to 0
    pathlength{i} = 0;
```

```matlab
    % calculate total pixels/um moved up to each timepoint
    % calculate speed at each time point, assume 0 at first timepoint
    for j = 2:numPoints(i)
        cumPxData{i}(j,:) = sum(cells{i}(1:j,:));
        cumUmData{i}(j,:) = cumPxData{i}(j,:).*PX2UM;
        dx = cells{i}(j,1).*PX2UM;
        dy = cells{i}(j,2).*PX2UM;
        distance{i}(j-1,:) = (dx^2+dy^2)^0.5;
        pathlength{i} = pathlength{i} + distance{i}(j-1,:);
%         UmSpeed(j,i) = PxSpeed(j,i)*PX2UM;
    end

    % plot on wind-rose
    % plotting um's moved per timepoint in x and in y, from origin
    totCellInPlot = totCellInPlot + 1;
    plot(cumUmData{i}(:,1),cumUmData{i}(:,2)); hold on;
    axis([xmin xmax ymin ymax]);
    xlabel('x (um)');
    ylabel('y (um)');

    % calculate MSD using nonoverlapping intervals
    msdnonoverlapSW;

    % calculate MSD using overlapping intervals
    msd_overlap;

    time = timeIntLength(1:numPoints(i)-1);
    Speed(i,1) = pathlength{i}/(numPoints(i)-1)./tint;
    [Pcalc(i,1), Calcresnorm(i,1), Calcresiduals{i}] =
lsqcurvefit(@prwfuncalc_SW, tint, time, msd{i}(:,1), 0, r.*tint,
optimset('Display','off'));
    [Pcalc_o(i,1), Calcresnorm(i,1), Calcresiduals{i}] =
lsqcurvefit(@prwfuncalc_SW, tint, time, msd_o{i}(:,1), 0, r.*tint,
optimset('Display','off'));

    % calculate average, displacement, and chemotactic index
    displacement(i) =
(cumPxData{i}(numPoints(i),1)^2+cumPxData{i}(numPoints(i),2)^2)^0.5;
    DI(i, 1) = pathlength{i}./displacement(i);

    % calculate model values
    model_calc{i} = 2*Speed(i).^2.*Pcalc(i).*[time - Pcalc(i).*(1-
exp(-time./Pcalc(i)))];
    model_calc_o{i} = 2*Speed(i).^2.*Pcalc_o(i).*[time -
Pcalc_o(i).*(1-exp(-time./Pcalc_o(i)))];

    % calculate nonoverlap r^2 = 1 - SS_tot/SS_err
    SS_tot = sum((msd{i}(:,1)-msd_avg(i)).^2);
    SS_err = sum((msd{i}(:,1)-model_calc{i}).^2);
    r2(i,:) = 1-SS_err./SS_tot;
```

```matlab
    % calculate overlap r^2 = 1 - SS_tot/SS_err
    SS_tot = sum((msd_o{i}(:,1)-msd_o_avg(i)).^2);
    SS_err = sum((msd_o{i}(:,1)-model_calc_o{i}).^2);
    r2_o(i,:) = 1-SS_err./SS_tot;

end

% Calculate % lower than threshold
I = find(r2<0.6);
I2 = find(r2_o<0.6);
r2_thres = length(I)/length(r2);
r2_o_thres = length(I2)/length(r2_o);

% calculate cumulative index
% DI = DI(I);
DI_avg = sum(DI)./length(DI);
DI_sem = std(DI)/sqrt(length(DI));

figure('Name', 'DI vs. Pers', 'NumberTitle','off');
plot(DI, Pcalc, 'o');
xlabel('DI');
ylabel('Persistence Time');
figure('Name', 'DI per cell', 'NumberTitle','off');
plot(DI, 'o');
xlabel('Cell number');
ylabel('DI');

% plot dP vs. dr^2 b/t nonoverlap & overlap
figure('Name', 'd(pers) vs. d(r^2)', 'NumberTitle','off');
dP = Pcalc_o - Pcalc;
dr2 = r2_o - r2;
plot(dP, dr2, '.');
xlabel('Change in persistence time (min)');
ylabel('Change in r^2');

% plot overlapping vs. nonoverlapping P
figure('Name', 'P overlap vs. nonoverlap', 'NumberTitle','off');
plot(Pcalc, Pcalc_o, '.');
xlabel('P nonoverlapping');
ylabel('P overlapping');

% plot Speed vs. r^2
figure('Name', 'Speed vs. r^2', 'NumberTitle','off');
plot(r2, Speed.*60, 'b.'); hold on;
plot(r2_o, Speed.*60, 'r.');
title('Speed vs. r^2');
xlabel('r^2');
ylabel('Speed (um/hr)');
legend('nonoverlapping', 'overlapping');

% plot Persistence vs. r^2
figure('Name', 'Pers vs. r^2', 'NumberTitle','off');
```

```matlab
plot(r2, Pcalc, 'b.'); hold on;
plot(r2_o, Pcalc_o, 'r.');
title('Persistence vs. r^2');
xlabel('r^2');
ylabel('Persistence time (min)');
legend('nonoverlapping', 'overlapping');

% plot Persistence vs. speed
figure('Name', 'Pers vs. Speed', 'NumberTitle','off');
plot(Speed, Pcalc, 'b.'); hold on;
plot(Speed, Pcalc_o, 'r.');
title('Speed vs. persistence');
xlabel('Speed (um/hr)');
ylabel('Persistence (min)');
legend('nonoverlapping', 'overlapping');

% plot change in r^2 vs. r^2
figure('Name', 'd(r^2) vs. r^2', 'NumberTitle', 'off');
plot(r2, dr2, 'b.');
title('Change in r^2 vs. r^2');
xlabel('r^2');
ylabel('Change in r^2');

% One does not have to see all individual plots:
seeplot = input(['Would you like to look at ' num2str(length(cells)) '
individual cell plots? '],'s');
if (seeplot == 'y') || (seeplot == 'Y')
    plotthem = 1;
else
    plotthem = 0;
end

temp_counter = 0;
for i = 1:length(msd)
    time = [];
    time = timeIntLength(1:numPoints(i)-1);

    % Plot nonoverlap calculated Speed/fitted Persistence Curve
    figure('Position',[scrsz(3)/6 scrsz(4)/6 scrsz(3)/1.4
scrsz(4)/1.4],...
        'Name',['Model fitting to MSD data - Cell #'
num2str(i)],'NumberTitle','off')
    subplot(2,2,3);
    plot(time,msd{i}(:,1),'xk',time,model_calc{i},'-r');
    title(['Fit P = ' num2str(Pcalc(i)) ' min - r^2 = '
num2str(r2(i))]);
    xlabel('Nonoverlapping time interval [min]');
    ylabel('Mean Squared Displacement [um^2]');

    % Plot overlapping fitted speed/fitted persistence curve
    subplot(2,2,4);
    plot(time,msd_o{i}(:,1),'xk',time,model_calc_o{i},'-r');
```

```matlab
    title(['Fit P = ' num2str(Pcalc_o(i)) ' - r^2 = '
num2str(r2_o(i))]);
    xlabel('Overlapping time interval (hr)');
    ylabel('Mean Squared Displacement (um)');

    % Plot cell trajectory
    subplot(2,2,1);
    plot(cumUmData{i}(:,1),cumUmData{i}(:,2), 'k.-')
    title(['Cell trajectory - ' num2str(numPoints(i)) '
coordinates']);
    xlabel('x [um]');
    ylabel('y [um]');

    % Plot speed at each time interval curve
    subplot(2,2,2);
    Splot = distance{i}./tint.*60;
    plot(time,Splot,'xb');
    title(['Speed at each time - Average S = ' num2str(Speed(i)) '
(um/hr)']);
    xlabel('Time [h]');
    ylabel('Speed [um/h]');

    if plotthem == 0
        close(gcf)
    else
    % Wait after 10 plots.
        if round(i./10) == i./10
            fprintf(['Cells # ' num2str(temp_counter.*10 + 1) ' - '
num2str(i)]);
            temp_counter = temp_counter + 1;
            wait = input('\nPress Enter to continue','s');
        end
    end
end

Speed = Speed.*60;

% Wait for plots to be looked at
if plotthem == 1
    if round(i./10) == i./10
    else
        fprintf(['Cells # ' num2str(temp_counter.*10 + 1) ' - '
num2str(i)]);
        wait = input('\nPress Enter to continue','s');
    end
end
```

8.2. msdnonoverlapSW.m

```
% msdnSW.m - Shan Wu 7.13.2007

% Calculates the mean-square-displacements from cell path coordinates
% for time intervals using non-overlapping intervals (see Dickinson
% and Tranquillo, AIChE J, 1993).

% Briefly, for each cell, non-overlapping time intervals of length 1
% through maximum are averaged to obtain the msd for the particular
% time interval length

% global intSpeed

% maximum number of different intervals lengths
maxInt(i) = numPoints(i) - 1;

% number of non-overlapping intervals of each length
nonOverlapInt = floor(maxInt(i)./(1:maxInt(i))');

% for each # of time intervals, sum up displacements of that interval
% length across the entire cell path using non-overlapping intervals
for j = 1:maxInt(i)

    msd{i}(j,1) = 0;                    % set initial msd to 0 for int length

    for k = 1:nonOverlapInt(j)

        startPos = (k-1)*j+1;
        % start of interval is k-1 intervals away from start
        endPos = k*j+1;
        % end of interval is k intervals away from start

        xdist = cumUmData{i}(endPos,1)-cumUmData{i}(startPos,1);
        ydist = cumUmData{i}(endPos,2)-cumUmData{i}(startPos,2);
        totDist = xdist^2+ydist^2;

        msd{i}(j,1) = msd{i}(j,1) + totDist;

    end

    msd{i}(j,1) = msd{i}(j)/nonOverlapInt(j);
    % msd is average over the total number of intervals

    msd{i}(j,2) = nonOverlapInt(j);

    % # of intervals per msd
    msd_avg(i) = sum(msd{i}(:,1))./length(msd{i});

end
```

8.3. msc_overlap.m

```matlab
% msd_overlap.m by Shan Wu 4.4.2009

% Calculates the mean-square-displacements from cell path coordinates
% for time intervals using overlapping intervals (see Dickinson and
% Tranquillo, AIChE J, 1993).

% Briefly, for each cell, overlapping time intervals of length 1
% through maximum are averaged to obtain the msd for the particular
% time interval length.  Cell tracks are not statistically indepedent

% global intSpeed

% maximum number of different intervals lengths
maxInt(i) = numPoints(i) - 1;

% number of overlapping intervals of each length
overlapInt = [(numPoints(i)-1):-1:1]';

% for each # of time intervals, sum up displacements of that interval
% length across the entire cell path using non-overlapping intervals
for j = 1:maxInt(i)

    msd_o{i}(j,1) = 0;                          % set initial msd for this
interval length to zero

    for k = 1:overlapInt(j)

        startPos = k;
          % start of interval is k-1 intervals away from start
        endPos = k + j;
          % end of interval is k intervals away from start

        xdist = cumUmData{i}(endPos,1)-cumUmData{i}(startPos,1);
        ydist = cumUmData{i}(endPos,2)-cumUmData{i}(startPos,2);
        totDist = xdist^2+ydist^2;

        msd_o{i}(j,1) = msd_o{i}(j,1) + totDist;

    end

    msd_o{i}(j,1) = msd_o{i}(j)/overlapInt(j);
    % msd is average over the total number of intervals

    msd_o{i}(j,2) = overlapInt(j);
    % # of intervals per msd

    msd_o_avg(i) = sum(msd_o{i}(:,1))./length(msd_o{i});

end
```

8.4. prwfuncalc_SW.m

```
function f = prwfuncalc_SW(P, time)

global Speed i

% f = 2*(intSpeed(1:numPoints(i)-1,i)).^2.*P.*[time - P.*(1-exp(-
time./P))];

f = 2*Speed(i).^2.*P.*[time - P.*(1-exp(-time./P))];
```