Physical Determinants of Glioma Cell Migration and Disease Progression

A dissertation submitted to the faculty of

University of Minnesota by

REBECCA L. KLANK

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Written under the advisement of Dr. David J. Odde

September 2015

Acknowledgements

Contributors to this work: The Minnesota Supercomputing Institute, Stacy Decker Grunke, Benjamin Bangasser, Colleen Forster, Matthew Price, Thomas Odde, Karen SantaCruz, S. Joey McFarren, Barbera Tschida, Timothy Kuka, Khoa Cao, Steven Rosenfeld, Peter Canoll, Eva Turley, James McCarthy, John Ohlfest, and David Odde

Funding sources: NIH Imaging Training Fellowship, 2014 University of Minnesota Doctoral Dissertation Fellowship, and NIH R01-CA172986.

Dedication

This dissertation is dedicated to all of the people who helped me in my PhD pursuit. To my friends in the graduate program who offered their camaraderie during my time at the University of Minnesota, thank you for reminding me to laugh, arranging well-timed happy hours, and challenging me as an intellectual. To my colleagues in the Odde lab, particularly Benjamin Bangasser, Brian Castle, Kwaku Opoku, Dominique Seetapun, Brannon McCullough, Emily Tubman, Louis Prahl, and Ghaidan Shamsan, thank you for all of your technical assistance. To my collaborators, David Largaespada, Steven Rosenfeld, Peter Canoll, Jim McCarthy, and Tim Ebner, thank you for your insights, critiques, and contributions to this work. To the late John Ohlfest, thank you for sharing your unwavering optimism towards developing treatments for these devastating cancers and teaching me the value that we as researchers bring to the field. To my wonderful advisor, Dave Odde, thank you for all of the support you have given me over the last few years. Your vast knowledge, enthusiasm, and thoughtful guidance have molded me into the scientist I am today. To my loving family, thank you for always believing in me and encouraging me to reach my full potential. And finally, to my incredible husband, Craig, thank you for everything. Through every late night, every failed experiment, every instance of imposter's syndrome, you were always there to help me get through it.

Abstract

Glioblastoma (GBM) is a highly aggressive brain cancer (generally, "glioma") with poor patient prognosis, even with current standard treatments. In order to rationally develop novel treatments that can significantly extend patient survival, we must first understand at a basic scientific level how the disease progresses. GBM is thought to be fatal due to highly invasive cells that migrate beyond the visible bulk tumor and lead to tumor recurrence after therapeutic intervention. Therefore, we sought to investigate what makes GBM cells invasive at the single-cell level (Chapter 1). Using a genetically induced mouse glioma model and confocal imaging of intact tumor-containing brain slices, we found that, consistent with previous biophysical models, glioma cell migration is biphasic with respect to the concentration of the transmembrane cell adhesion molecule CD44. By contrast, cell proliferation is independent of CD44 level. Additionally, mouse model and human patient survival are also biphasic with respect to CD44 level, with poorest prognosis occurring at intermediate CD44 levels. Thus, migration and survival are both biphasic and are anti-correlated to each other, suggesting that CD44-dependent migration directly affects survival outcomes. We next investigated how these single-cell behaviors impact overall tumor growth and progression (Chapter 2). Noticing that previous models of GBM migration use parameter values for migration rate (defined by a diffusion coefficient, also known as a random motility coefficeint) that are much higher than our measurements of single-cell migration behavior in Chapter 1, a Brownian dynamics (BD) approach was used to simulate single-cell growth, proliferation, and

migration, and compare model assumptions. These studies showed that employing the physically-based assumption that tumor cells occupy volume, an assumption not captured in current reaction-diffusion (RD) simulations, resulted in increased tumor spreading behavior with the same input parameters. Specifically, non-overlapping cells can enter a jammed regime where interior cells are subdiffusive, and peripheral cells become biased outward and superdiffusive in a quasi-ballistic expansion. Thus, we show that, when we account for volume conservation, the relatively low values of diffusion coefficient, such as what was measured in Chapter 1, can generate fast progressing tumors that are similar to RD simulations which use diffusion coefficients much greater than what is observed experimentally for single migrating cells. Therefore, we suggest that cellular jamming behavior contributes to the fast spreading of GBM tumors, and that subsequent simulations of GBM growth should incorporate this assumption so that models are physically grounded and achieve consistency between single-cell behavior and bulk tumor progression. Overall, these studies demonstrate the potential importance of fundamental physical effects in driving tumor progression generally, and glioblastoma specifically.

Table of Contents

Acknowledgements i

Dedication ii

Abstract iii

List of Tables viii

List of Figures ix

List of Supplemental Figures x

List of Supplemental Movie Captions x

Introduction 1

Glioblastoma 2

Cell migration 4

GBM Migration 8

Synopsis 9

Chapter1: Biphasic dependence of glioma survival and cell migration on CD44 expression level 11

Summary 12

Introduction 12

Results 15

Conclusions 31

Materials and Methods 31

Animals and tumor models 31

Primary cell lines 33
Western Blotting 33
Analysis of published human data 34
Bioluminescence 35
Immunohistochemistry 35
Stochastic whole-cell model 37
Single cell migration dynamics in brain slice preparations 43
Supplementary Discussion 44
Possible mechanisms of CD44-mediated effects on migration in the mouse brain 44
Relationship to previous analyses of CD44 in glioma 47
Implications for bioinformatics, biomarker discovery, genetic analysis, and therapeutic strategy 51
Final Remarks 52
Supplemental Figures 54
Supplemental Movie Captions 60
Acknowledgments 60
Chapter2: Brownian dynamics simulation of tumor progression and application to glioblastoma 62
Summary 62
Introduction 64
Results 73

Discussion 85
Conclusions 89
Methods 91
One-Dimensional Brownian Dynamics Simulation 91
Data Analysis 93
Supplemental Materials 94
Derivation of diffusion within a concentrated solution 94
Supplemental Figures 97
Supplemental Movie Captions 102
Conclusions 103
References 110
Appendix A. Biographical Sketch 121

List of Tables

Table 1. Simulation parameters of 2D motor –clutch simulation of cell migration 27

Table 2. Bulk tumor-based theoretical estimates of diffusivity in the literature using and RD model 70

Table 3. Single cell experimental measurements of diffusivity in the literature 71

List of Figures

Figure 1. Magnetic resonance images (MRI) of a patient with GBM. 3

Figure 2. Diagram of the canonical view of cell migration 5

Figure 3. Control of CD44 level in the Sleeping Beauty de novo mouse glioma model and recapitulation of human glioma. 15

Figure 4. Mouse and human glioma survival are biphasic with respect to CD44 expression level. 18

Figure 5. Net proliferation rate in the mouse gliomas is independent of CD44 expression level 24

Figure 6. Single-cell motility in biophysical simulation and in ex vivo brain slices is biphasic with respect to CD44 expression level 29

Figure 7. Diagram of overlapping + carrying capacity and non-overlapping models 76

Figure 8. Spatio-temporal dynamics of the three BD models 79

Figure 9. Analysis of overall model behavior and parameter sensitivity 80

Figure 10. Comparison of velocity values between the overlapping+cc and non-overlapping models 81

Figure 11. Intratumoral spatial variation of single cell motion behavior 84

List of Supplemental Figures

Supplemental Figure 1. TGCA survival analysis 54

Supplemental Figure 2. Bioluminescence imaging 56

Supplemental Figure 3. Whole-cell model schematic 58

Supplemental Figure 4. Western blots of non-muscle myosin heavy chain isoforms 59

Supplemental Figure 5. Progeny trees of simple overlapping model 97

Supplemental Figure 6. Progeny trees of the overlapping + carrying capacity model 98

Supplemental Figure 7. Progeny trees of the non-overlapping model..... 99

Supplemental Figure 8. Long term behavior of simulations..... 100

Supplemental Figure 9. Effect of cell time interval on cell velocity estimates 101

List of Supplemental Movie Captions

Supplemental Movie 1. Whole-cell simulations of changing clutch number..... 60

Supplemental Movie 2. Tracked cells in a WT brain slice experiment..... 60

Supplemental Movie 3. Simulation of the simple overlapping uninhibited model 102

Supplemental Movie 4. Simulation of the overlapping + carrying capacity model 102

Supplemental Movie 5. Simulation of the non-overlapping model..... 102

Introduction

In the year 2015, 1.6 million Americans are expected to be newly diagnosed with some form of cancer (*1*). In the same year, 589,000 are expected to perish from cancerous diseases (*1*). While there are cancers particular to every organ in the body, each with unique characteristics and prognoses, collectively what makes cancer devastating for so many people is its ability to usurp ordinary physiological processes for its own uncontrolled expansion, ultimately destroying normal tissue and organ function. Therefore, in addition to drug discovery and design of early detection methods, current cancer research is dedicated to understanding the process of tumor progression in order to develop novel therapeutic strategies that could retard or perhaps reverse the course of these diseases.

Tumor progression, in general, by the simplest definition can be described as aberrant cell proliferation and migration into new tissue regions. Both of these processes – proliferation and migration – are key to the development of a lethal tumor. Aberrant proliferation, often caused by mutation to cell growth checkpoints such as phosphatidyl-inositol-3-OH kinase (PI3K), p53, or MDC1(*2*), creates an excess of cells which can damage surrounding normal tissue and result in loss of tissue function. However, tumors which are characterized only by excessive growth are often treatable by surgical removal as the tumor is localized to one region. It is the tumors with invasive characteristics that

spread beyond the initial tumor location, i.e. malignant tumors, which typically progress to lethality.

In malignant tumors, for example invasive carcinomas, cells are able to invade into surrounding healthy tissue by breaking through a region of connective tissue called the basement membrane (*3*). The invasion of tumor cells in malignant cancers spreads the disease to new areas of the tissue and creates a blurred tumor margin that makes surgical intervention more difficult (*4*). In some tumors, spreading is also achieved via metastasis, where individual cells are able to intravasate into the bloodstream and seed a new tumor in a secondary tissue via extravasation from the bloodsteam (*5*). Particular to this dissertation is the question of what determines tumor spreading in the cancer glioblastoma (GBM), which very rarely metastasizes (*6*), but spreads through the invasion of individual cells throughout the entire organ.

Glioblastoma

GBM is a grade IV astrocytoma (glioma) that is infamous for its fatality and is the most common type of brain cancer, accounting for approximately 17% of all central nervous system tumors (*7*). The incidence of GBM tumors in the United States is approximately 3.19 cases/100,000 person-years (*7*), which translates to roughly 10,000 cases per year. Unlike other central nervous system tumors with better prognoses, GBMs offer patients a dismal 5-year survival outcome of merely 4.75% (*7*). Moreover, the median survival time for GBM patients is only 15 months post-diagnosis (*8*, *4*). Despite

implementation of current standard treatment with surgery, alkylating agent-based chemotherapy, and focused radiotherapy (*4*, *9*), patients with GBM frequently experience tumor recurrence and ultimately succumb to the disease (Figure 1). Consequently, the current treatment methods are generally regarded as palliative rather than curative. To improve on these dismal outcomes, it is imperative to understand the fundamental mechanisms of GBM progression.

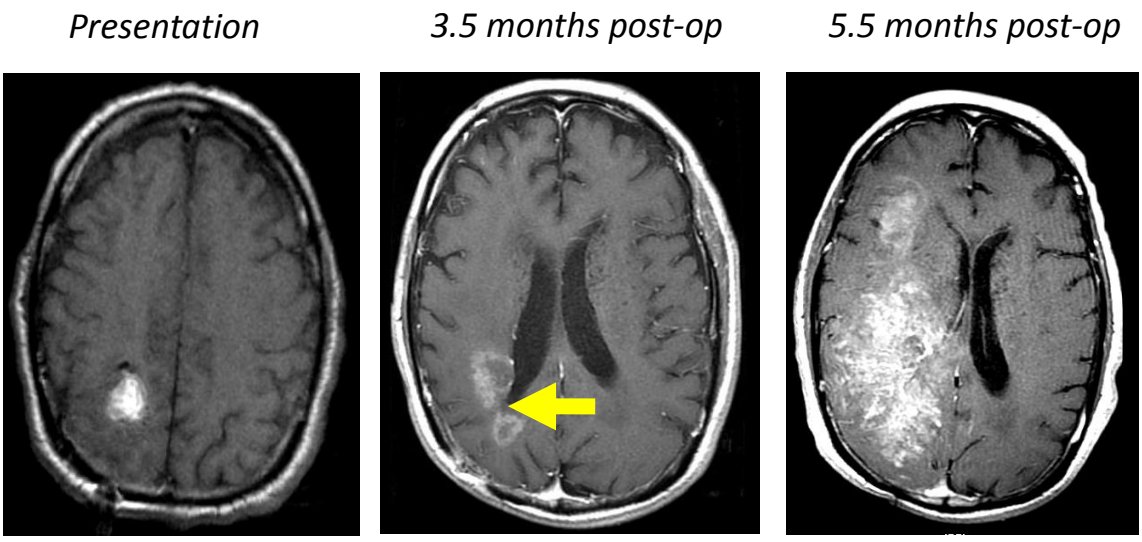


Figure 1. Magnetic resonance images (MRI) of a patient with GBM.
A gadolinium-enhanced MR image of a typical glioblastoma patient shows a bright lesion in the left hemisphere when the patient presented in the clinic (left panel). After surgical removal of the lesion, the patient MRI showed evidence of tumor recurrence near the initial site (yellow arrow, middle panel). The MR image taken 2 weeks later (right panel) shows that the lesion had spread to nearly the entire left hemisphere. This particular patient passed away only 2 weeks after the last MRI scan. Images were provided by Dr. Steven S. Rosenfeld, M.D. Ph.D.

GBMs are notoriously invasive tumors, and the presence of disperse invasive cells which are able to evade focal treatments is thought to be the cause of GBM recurrence and ultimately death (*10*). In fact, in the early days of treatment, surgeons would take drastic measures to prevent tumor recurrence by removing patients' entire affected hemisphere, yet even at the time of disease presentation, invasive GBM cells had reached the opposite hemisphere and recurrence was unavoidable (*4*, *10*). Therefore, in the case of GBM, it is important to understand cell migration within the brain microenvironment for the development of effective treatments to this disease.

Cell migration

Cell migration is a fundamental process utilized in normal physiological functions such as embryogenesis and wound healing (*11*); however, in cancers this normal cellular behavior leads to the spreading and dissemination of tumors. The process of cell migration involves complex interactions between cytoskeletal components, motor proteins, receptor complexes, and extracellular ligands.

The canonical model for cell motility is illustrated in Figure 2. In this model, cells create a protrusion, establish focal contacts to generate traction, and contract while simultaneously releasing adhesions in the rear, causing the cell to propel forward (*12*).

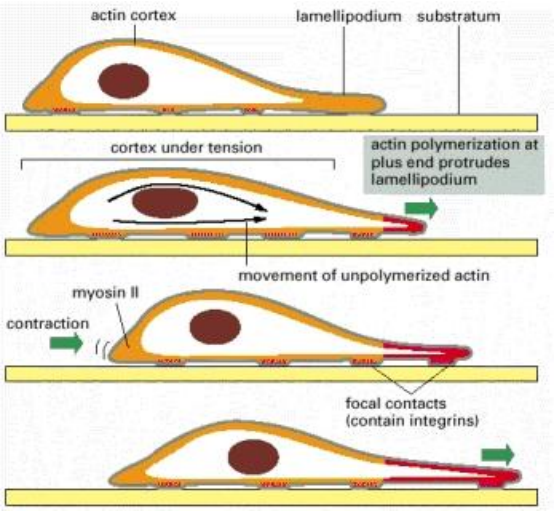


Figure 2. Diagram of the canonical view of cell migration

A migrating animal cell extends its lamellipodium, creates new adhesions, and releases its adhesions in the rear while contracting its actin filaments via myosin II motor forces, causing forward movement. Adapted from Molecular Biology of the Cell, Figure 16-86 (*12*).

The cytoskeletal protein most active in migration is actin. Actin filaments are composed of globular actin monomers which self-assemble to form double-stranded helical structures (*12*). A family of proteins known as myosin II motor proteins "walk" along filamentous actin (F-actin), with two functional sets of motor domains that cross-link adjacent actin filaments, causing contraction of the cytoskeleton. This interaction of actin and myosin within the cell is analogous to actin-myosin contraction in muscle (*12*). However, contraction of actin fibers is not sufficient to move cells, just as walking in a frictionless environment does not permit net translocation. Cells generate necessary traction by interacting with their environment through transmembrane cell adhesion molecules such as integrins.

Integrins are glycoprotein heterodimers which, on the extracellular side, bind to specific amino acid sequences in the extracellular matrix (ECM), and in the cytosol, connect to actin filaments through a series of linker proteins such as talin, vinculin, tensin, paxillin, and α-actinin (*12*, *10*, *13*). Thus the linkage between the force generating myosin molecules and the ECM creates a mechanical continuum between the internal cellular structures and the extracellular environment.

The ECM is composed of a variety of proteins and glycosaminoglycans (GAGs) that fill the extracellular space within a tissue. ECM proteins such as collagen, fibronectin, and laminin assemble to form 3D lattices which act as structural supports for cells (*12*). These molecules are often deposited, remodeled, and degraded by local cells by the release of matrix metalloproteases (MMPs) (*14*).

In addition to the basic linkages that contribute to cell migration, cells have also been found to migrate in response to certain cues; particularly cells have been shown to respond to chemical and mechanical cues. The most basic migratory behavior in response to chemical cues is the well-documented phenomenon of chemotaxis. Chemotaxis is the directional migration of cells either toward or away from a soluble gradient of chemical compounds. One example is the chemotaxis of glioma cells to secreted platelet derived growth factor (PDGF) which was computationally modeled by Massey *et al.* (*15*).

Migration has also been theorized to depend on the concentration of chemical ligands in the ECM and the cell's ability to bind to these molecules (*16*). As

demonstrated in simulations, as well as in *in vitro* studies (*17, 18*), cells display optimal migration behaviors when substrate attachments are neither weak nor strong, but intermediate. This is due to the fact that weak attachments do not generate sufficient adhesion to the environment to produce traction, and strong attachments do not release to allow the cell to move. This model showed not only a dependence of migration on ligand concentration and adhesion strength, but also a dependence on mechanics.

Other groups have also demonstrated a mechanical dependence on migration. For instance, Pelham and Wang showed that cells can be mechanically sensitive by culturing both NIH 3T3 fibroblast cells and normal rat kidney epithelial cells on mechanically tunable polyacrylamide gels (PAGs) conjugated with collagen ECM molecules. In their study, they demonstrated that cells had altered morphological and motile behavior as substrate stiffness was changed (*19*). Additionally, Lo *et al.* used PAG substrates to show that cells would migrate preferentially from soft to stiff regions, a phenomenon they termed "durotaxis" ("duro" meaning "hard") (*20*).

An important computational model for describing mechanosensing (and cell motility) is the motor-clutch model first described by Mitchison and Kirshner in 1988 (*21*). In this model, F-actin within a cell is physically linked to a substrate by transiently engaged molecular "clutches" which allow the cell to generate traction. In 2008 Chan and Odde expanded upon this model to include the compliance of the substrate and showed that leading edge progression and actin retrograde flow both depended on the rigidity of

the substrate (*22*). More recent adaptations of the model have shown that the maximal traction force occurs at the intermediate substrate stiffness, where migration is hypothesized to be fastest (*23*). These physical models of cell migration behavior inform our understanding of migration in general, and point to potential mechanisms which may contribute to GBM motility.

GBM Migration

Migration of GBM cells *in vivo* could be due to several factors including chemical mechanisms, such as hypoxic gradients (*24*, *25*) or paracrine signaling (*8*, *26*). GBMs have also been shown to migrate preferentially along certain anatomical features of the brain including vascular basement membranes (*8*, *27*, *28*), subependyma (*8*), and myelinated axons (*8*) (white matter tracts(*27–29*)). In fact, a retrospective statistical analysis of 121 GBM patients found that outcomes were more favorable in patients with tumors that originated in the deep grey matter (basal ganglia) than patients whose tumors originated in the lobar, white matter regions (*9*). This could be a response to the presence of specific ligand molecules like myelin (*30*), or a mechanical response to topographical features in white matter regions.

GBM cells have also been shown to be mechanosensitive in a study by Ulrich *et al.* using immortal GBM cell lines cultured on *in vitro* PAG substrates of different mechanical stiffness (*29*). Additionally, it was found by Beadle *et al*. that animal model glioma cells migrate differently in response to geometric confinement. Specifically, on

2D surfaces, cells migrated in a manner that was independent of myosin II activity, but in confined 3D environments migration was dependent on myosin II (*31*).

Additionally, the ECM molecules of the brain may be contributing to GBM migration. Brain ECM is unique in that many of the ECM proteins common to other tissues are present only at low concentrations. Instead, a large portion of the brain ECM is composed of hydrophilic GAG molecule such as chondroitin sulfate and hyaluronic acid (HA) (*10*, *27*, *32*). The primary receptor for HA is the transmembrane molecule CD44. CD44 is commonly overexpressed in cancers, including GBM (*33*), Moreover, similar to integrins, CD44 is linked to the underlying actin cytoskeleton through linker proteins (ERM/Ankyrn (*34*)), and therefore it is hypothesized that CD44 may be significantly contributing to characteristic GBM invasiveness (*35*). More details about the contributions of CD44 to GBM migration are discussed in Chapter 1.

Synopsis

Overall, this dissertation describes a biophysical approach to understanding brain cancer progression. In the first chapter, I discuss how the use of quantitative biophysical methods led to the discovery of a biphasic relationship between the expression level of the transmembrane cell adhesion molecule CD44 and survival outcomes in both an induced mouse model of high-grade glioma and in human high-grade glioma patients. Furthermore, utilization of existing biophysics theory on the biphasic relationship between cell adhesion and migration led us to investigate the role of CD44 on cell

migration within our mouse models, where we observed the predicted biphasic response. Interestingly, this response was anti-correlated to survival outcomes, thus offering a potential mechanism for the observed biphasic survival. These findings are clinically relevant in that they suggest that anti-CD44 treatments could have deleterious effects for a subset of patients, yet be beneficial to others.

In the second chapter, I present a new model of GBM tumor progression using a one-dimensional Brownian dynamics (BD) approach. Noting a major quantitative discrepancy between cell migration speed estimates used in current progression simulations and experimental measurements of single-cell migration rates within an animal model system, I re-evaluated the assumptions of existing models in the literature. BD simulation outputs demonstrated a limitation of current GBM models that assume cell volume is negligible and showed how cell proliferation and migration act in concert to promote a faster-progressing tumor when volume was considered. Moreover, this model shows how using physically-grounded parameter values can produce *in silico* tumors with the appropriate rate of progression. Therefore, in order to develop more predictive models that can be helpful in the clinic using patients' own cell parameters, current models must be adjusted to account for the effects of cell volumes.

Chapter1: Biphasic dependence of glioma survival and cell migration on CD44 expression level

Authors: Rebecca L. Klank[1,†], Stacy A. Decker Grunke[2,†], Benjamin L. Bangasser[1], Colleen L. Forster[3], Matthew A. Price[3], Thomas J. Odde[1], Karen S. SantaCruz[3], Steven S. Rosenfeld[4], Peter D. Canoll[5], Eva A. Turley[6], James B. McCarthy[3,#], John R. Ohlfest[2,#,‡], David J. Odde[1,#,*]

Affiliations

[1]Department of Biomedical Engineering, University of Minnesota, Minneapolis, Minnesota, USA

[2]Department of Pediatrics, University of Minnesota, Minneapolis, Minnesota, USA

[3]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, USA

[4]Brain Tumor and Neuro-Oncology Center, Cleveland Clinic, Cleveland, Ohio, USA

[5]Department of Pathology and Cell Biology, Columbia University Medical Center, New York, New York, USA

[6]Department of Oncology, London Health Science Center, Schulich School of Medicine, Western University, London, Ontario, Canada

†Co-first Authors

#Co-senior Authors

‡Deceased Contributor

Summary

While several studies link the cell surface marker CD44 to cancer progression, conflicting results show both positive and negative correlations with increased CD44 levels. Here we demonstrate that the survival outcomes of genetically-induced glioma-bearing mice and of high-grade human glioma patients are *biphasically* correlated with CD44 level, with poorest outcomes occurring at intermediate levels. Furthermore, three previously defined human glioma subtypes exhibited distinct linear correlations between CD44 and survival based on their relative location along the biphasic curve; most notably, the high-CD44-expressing mesenchymal subtype exhibited a *positive* trend of survival with increased CD44 level. Mouse cell migration rates in *ex vivo* brain slice cultures were also biphasically associated with CD44 level, with maximal migration corresponding to minimal survival. Cell simulations suggest that cell-substrate adhesiveness is sufficient to explain this biphasic migration. More generally, these results highlight the potential importance of non-monotonic relationships between survival and biomarkers associated with cancer progression.

Introduction

At its most basic level, cancer progression is driven by aberrant cell proliferation and migration into previously healthy tissue. In the case of glioblastoma (grade IV glial brain cancer), even with standard treatment therapies, median patient survival is approximately 15 months (*36*) due to the well-established invasiveness of the tumors (*8*,

37). The brain is relatively rich in hyaluronic acid (HA), rather than collagen and fibronectin (*27*), suggesting that CD44, a major transmembrane cell surface receptor for HA (*38*), could be an important mediator of glioma cell migration into the brain parenchyma (*39*, *40*). More generally, CD44 is reported to contribute to cancer invasiveness and has been implicated in epithelial to mesenchymal transition of many cell types (*33*). However, even though CD44 has been linked to cancers of the breast, lung, prostate, and brain, the literature is mixed on the importance of CD44 (*33*, *41*). In the case of glioma, some studies identified CD44 as a negative prognostic indicator of survival (*42–45*), while others have found no correlation (*46*). Still others reported that, while CD44 concentration is positively correlated with glioma grade, within the highest malignancy grade, patients with expression above the median for that group experienced longer survival than those below the median expression (*47*). At a fundamentally mechanistic level, such apparently contradictory results are not necessarily surprising due to the well-known biphasic dependence of cell migration on the strength of adhesion to the surrounding extracellular matrix (*16–18*). CD44 is a cell adhesion molecule that extracellularly binds to HA and intracellulary is mechanically linked to the actin cytoskeleton through ezrin/radixin/moesin proteins (*34*); therefore, from a biophysics perspective, cells would be expected to exhibit biphasic, not monotonic, migration behavior in response to increasing CD44 content. Since migration contributes to overall tumor progression, we further hypothesized a biphasic relationship between CD44

expression level and survival. However, such biphasic relationships have not been established *in vivo*, and furthermore their relevance to cancer progression is unclear.

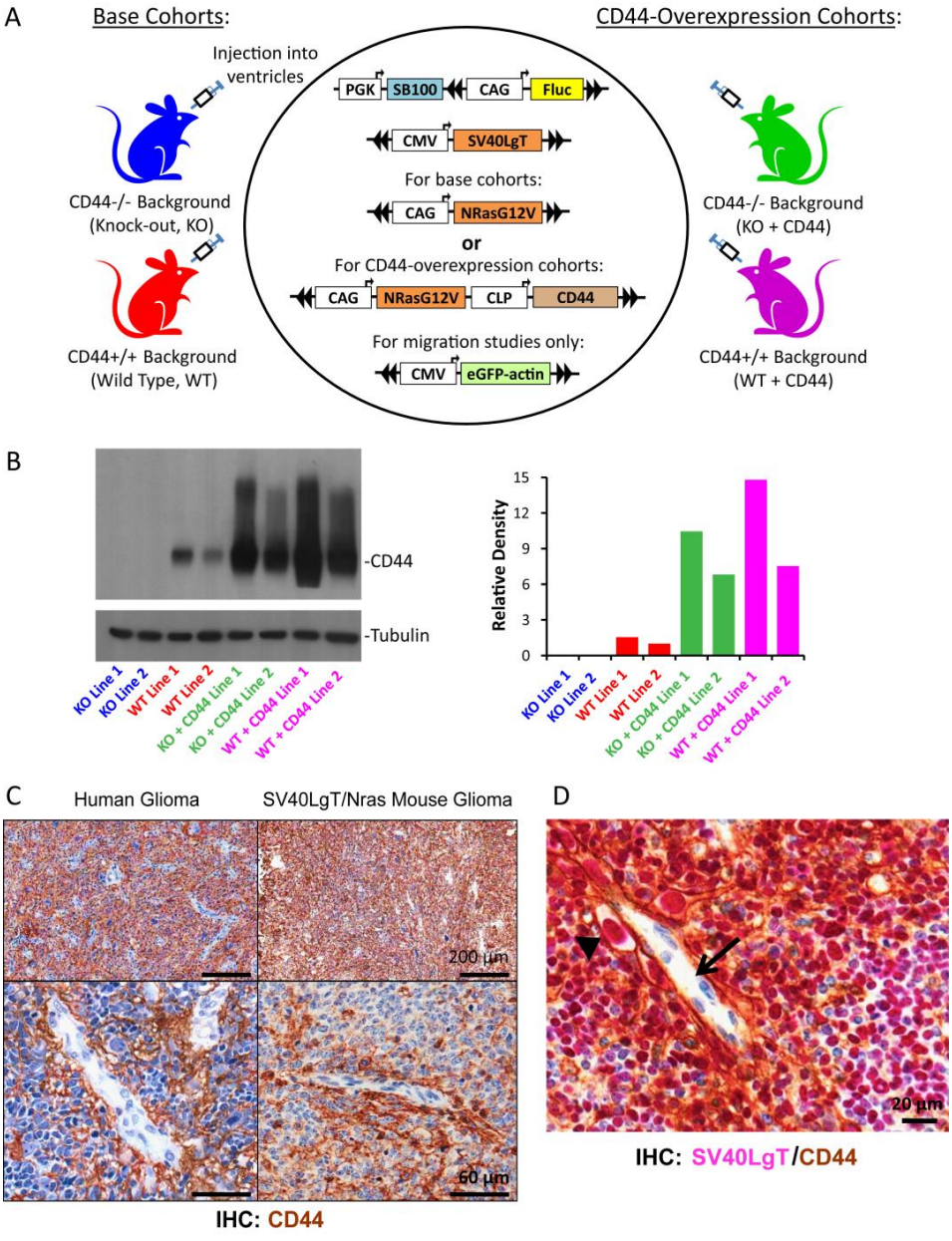


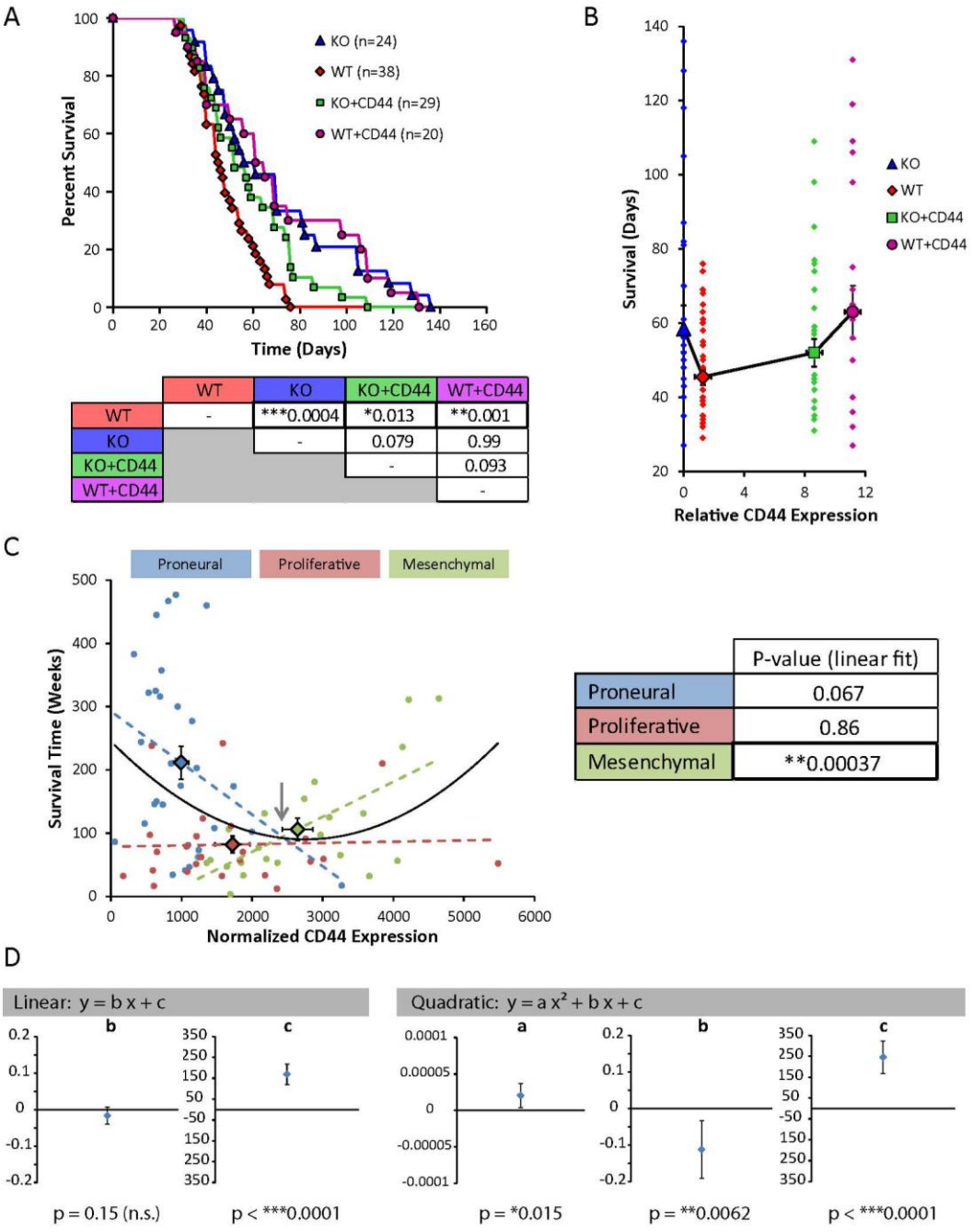
Figure 3. Control of CD44 level in the Sleeping Beauty de novo mouse glioma model and recapitulation of human glioma.

A. Schematic of the mouse glioma models. Plasmids encoding for Sleeping Beauty transposase (SB100, light blue), oncogenes (orange), and Firefly Luciferase (yellow), were injected into CD44-competent (red) or matched CD44-deficient (blue) mice. Exogenous CD44 (pink) was included in KO+CD44 (green) and WT+CD44 (magenta) overexpression cohorts. For brain slice migration studies, a fourth plasmid encoding eGFP-actin was included in each of the four conditions. The double arrows on the plasmids represent the SB100 binding sites. See materials and methods section for more details. B. Left: Western blot of CD44 expression for glioma cell lines derived from each of the four conditions (2 per condition). Right: Densitometry quantification of CD44 expression relative to WT line 2. C. Histological comparison of SV40LgT/Nras constructs in a WT mouse model to a human high-grade glioma shows similarities in CD44 staining. Above: Low magnification shows homogeneous staining patterns in human and mouse tissues. Below: Higher magnification shows cell-to-cell variability and strong CD44 staining in the perivascular region. D. Two-color IHC for CD44 (brown, membrane) and SV40LgT (pink, nuclear) in WT mouse models confirms that CD44 expression is from SV40LgT-expressing glioma cells in the perivascular region (arrowhead). Arrow shows endothelial cells forming the vascular lumen.

Results

To investigate the effects of CD44 on glioma progression, we used a *de novo* mouse glioma model which uses the Sleeping Beauty (SB) transposase system that has previously been shown to stably integrate oncogenic plasmid DNA into the host genome (*48*). Moreover, this model system has been shown to drive astrocytic tumors, consistent with World Health Organization grade III and grade IV gliomas. Our SB-based model uses transforming plasmids encoding for constitutively active Nras (NrasG12V) and simian virus 40 Large T antigen (SV40LgT) to mimic common human mutations frequently disrupted in human gliomas (e.g. RAS/PI3K, p53, RB). Altogether, four

conditions were investigated (Figure 3A): CD44-positive wildtype animals (+/+, WT), genetically-matched CD44 knockout animals (-/-, KO), KO animals with exogenous CD44 plasmid (KO+CD44), and WT animals with exogenous CD44 plasmid (WT+CD44) as models of CD44 overexpression. Western blotting of glioma cell lines derived from the *de novo* mouse model (2 per condition) was used to confirm that KO lines lacked CD44 and to determine the relative CD44 concentrations for each condition (Figure 3B). Quantification shows that both conditions injected with CD44 plasmid (KO+CD44 and WT+CD44) contained higher CD44 levels than WT, indicating that the addition of the CD44 plasmid generates an exogenous overexpression model, regardless of the mouse genotype. Therefore, the primary difference between the KO+CD44 and WT+CD44 conditions is the presence of endogenous levels of CD44 in stromal cells in the latter case. Immunohistochemistry (IHC) analysis demonstrated that our mouse model shows similar characteristics when compared to human glioma (Figure 3C) and that CD44 was expressed by tumorigenic SV40LgT+ cells in our models (Figure 3D).



	WT	KO	KO+CD44	WT+CD44
WT	-	***0.0004	*0.013	**0.001
KO		-	0.079	0.99
KO+CD44			-	0.093
WT+CD44				-

	P-value (linear fit)
Proneural	0.067
Proliferative	0.86
Mesenchymal	**0.00037

Figure 4. Mouse and human glioma survival are biphasic with respect to CD44 expression level. A. Kaplan-Meier plot of animal survival for each of the four model conditions. Table below lists the p-values for each pairwise comparison within an overall multiple comparisons statistical framework (*p<0.05, ** p<0.01, *** p<0.001). B. Biphasic dependence of survival on CD44 protein expression level. Large colored markers indicate the median survival. Dots corresponding to each color indicate the survival times of individual animals. Horizontal error bars show the median ± coefficient of variation/√2. Vertical error bars correspond to ± standard error of the mean for survival. C. Biphasic dependence of human glioma patient survival on CD44 mRNA expression (Rab14-normalized) from data published in Phillips et al. 2006 (*49*). Individual data points are color-coded by subtype, as determined by Phillips et al. Quadratic fitting of all the data (black curve) reveals a statistically significant non-monotonic/biphasic fit (p=0.015) with upward concavity. A linear correlation analysis was conducted on each subtype separately to determine subtype-specific linear trends (right table). Notably, the mesenchymal subtype exhibited a strong positive linear correlation with CD44 level. Gray arrow indicates the point where the three linear fits approximately converge, closely coinciding with the minimum in the biphasic fit. D. Left: Values of each of the coefficients in the linear fit of all the human survival data are plotted with error bars indicating the 95% confidence intervals. While the data overall exhibits a slight downward linear trend, the slope is not statistically different from zero (p=0.15). Right: Values of each of the coefficients in the quadratic (biphasic) fit are plotted with error bars indicating the 95% confidence intervals. All coefficients in the quadratic fit are statistically significant, confirming a biphasic dependence of survival on CD44 level, with intermediate levels exhibiting the worst outcomes.

Survival outcomes were tested for each of the four conditions depicted in Figure 4A. Comparisons of the Kaplan-Meier survival curves reveal that both KO mice and KO+CD44 mice exhibited a moderate, yet significant, survival increase compared to WT animals. Inclusion of the CD44 plasmid in WT animals (WT+CD44), which, if the monotonic assumption were correct, would hypothetically produce the worst survival outcome, exhibited the longest median survival compared to all other conditions (Figure 4A). By plotting the median survival of each condition against the relative CD44

concentration, our mouse model shows that survival is biphasic with respect to protein level expression of CD44 (Figure 4B), where intermediate, WT levels exhibit the worst survival outcomes.

We next investigated whether this biphasic relationship was present in the human disease. A recent study by Bhat *et al.* implicated CD44 as a negative prognosticator of survival in human glioblastoma patients (*45*). In their study, the authors compared the survival of newly diagnosed glioblastoma patients with high CD44 expression to that of low CD44 expression, based on qualitative classification of stained histological sections. Using this high/low binning approach, they found a significant difference in the two survival curves, where higher CD44 correlated with worse survival outcomes. In our analysis, we utilized a slightly larger published dataset of human glioma survival from Phillips *et al.* 2006 (*49*) which included quantifiable patient mRNA expression levels. Using categorical methods similar to Bhat *et al.*, we pooled the data in two groups, one above the median CD44 expression level and one below, and found that survival of high expressing group was statistically lower than that of the low expressing group ($p=0.02$, data not shown). Alternatively, instead of pooling the data into two groups, we also took a continuum approach to data analysis, which does not require an arbitrary threshold to assign groups, and plotted survival times as a function of CD44 expression level (Figure 4C). The implication of the categorical analysis is that more CD44 is worse for patient outcome; however, when we fit a line to the continuum data, we found that, while there is

a slight negative slope (slope= -0.02, not shown), the 95% confidence interval of this parameter includes the value zero, meaning that the linear trend is not statistically significant (Figure 4D – left). Given the results from the mouse experiments, we then tested to see if a quadratic (biphasic) fit would be more appropriate. The quadratic fitting of the data proved to be statistically significant (p=0.015 for the quadratic term), indicating that patient outcomes actually *improve* once CD44 levels are sufficiently high (Figure 4C – black curve). The 95% confidence intervals for all coefficients in the quadratic fit do not include zero, meaning that all of the coefficients are statistically significant (Figure 4D – right). Moreover, it should be noted that we did not constrain our quadratic fit to a positive quadratic term, yet both animal and human data showed biphasic relationships that are concave-up, demonstrating that there is a minimum in survival in both cases at relatively intermediate CD44 levels. Independent of the statistical analysis, it is evident from direct examination of the data itself that there are very few long-term survivors at intermediate CD44 level, and that all of the long-term survivors (>200 weeks) either had low CD44 (<1500) or high CD44 (>3500). Similar analysis was completed using data available through The Cancer Genome Atlas (TCGA). Although the numerical values of CD44 expression differ between the two datasets, the results from TCGA show a similar concave-up biphasic trend (p=0.018, Supplemental Figure 1A), while the linear trend was again not statistically significant (p=0.694). Thus, we find that survival is biphasically dependent on CD44 level for both human glioma

patients and genetically-induced glioma-bearing mice, with poorest prognosis in both cases at intermediate CD44 level.

The original genome-wide transcriptomic analyses by Phillips *et al.* in 2006 were seminal to the field in terms of establishing at least 3 glioma subtypes. Although more recent subtype classifications have emerged (*50*), subsequent classification analyses have confirmed the robustness and consistency of the proneural and mesenchymal subtypes in particular (*45*). Therefore, we used the original Phillips classification scheme to determine if our continuum approach to expression-level analyses would elucidate any subtype-specific trends. As shown in Figure 4C, we found that proneural subtypes tended to be lower in CD44 expression, proliferative subtypes were intermediate, and mesenchymal subtypes tended to have the highest CD44 expression. This is perhaps not surprising as CD44 was used as a histological marker for the mesenchymal subtype (*49*) and has been shown to positively correlate with a transition to a mesenchymal signature in glioma (*51*).

Consistent with a biphasic model, linear correlation analysis for survival as a function of CD44 within these three subtypes showed distinct trends based on Rab-14 normalized CD44 expression: 1) the proneural subtype demonstrated a modest negative linear correlation between survival and CD44 level ($p=0.067$), 2) the proliferative subtype demonstrated no correlation ($p=0.86$), and 3) the mesenchymal subtype demonstrated a strong positive correlation ($p=0.00037$). The latter result is particularly

interesting because it implies that, while the mesenchymal subtype uses CD44 as a histological marker and exhibits overall poor survival outcome (*49*), CD44 is actually a *positive* prognostic indicator within the mesenchymal subtype, with better survival observed at higher CD44 levels. In addition, it is worth noting that the three best-fit linear correlations for the three subtypes (proneural, proliferative, and mesenchymal) all intersect at a point that nearly coincides with the quadratic fit to the overall data. This point, at normalized CD44 level=2500, represents the worst-case scenario for high grade glioma outcome, where prognosis is poorest, a feature of the data that is evident even without statistical analysis. Overall, our analysis of the high grade glioma subtypes is consistent with our analysis of the pooled continuum data, where low and high CD44 expressors have, on average, better survival outcomes, and intermediate expressors have the worst average survival outcomes.

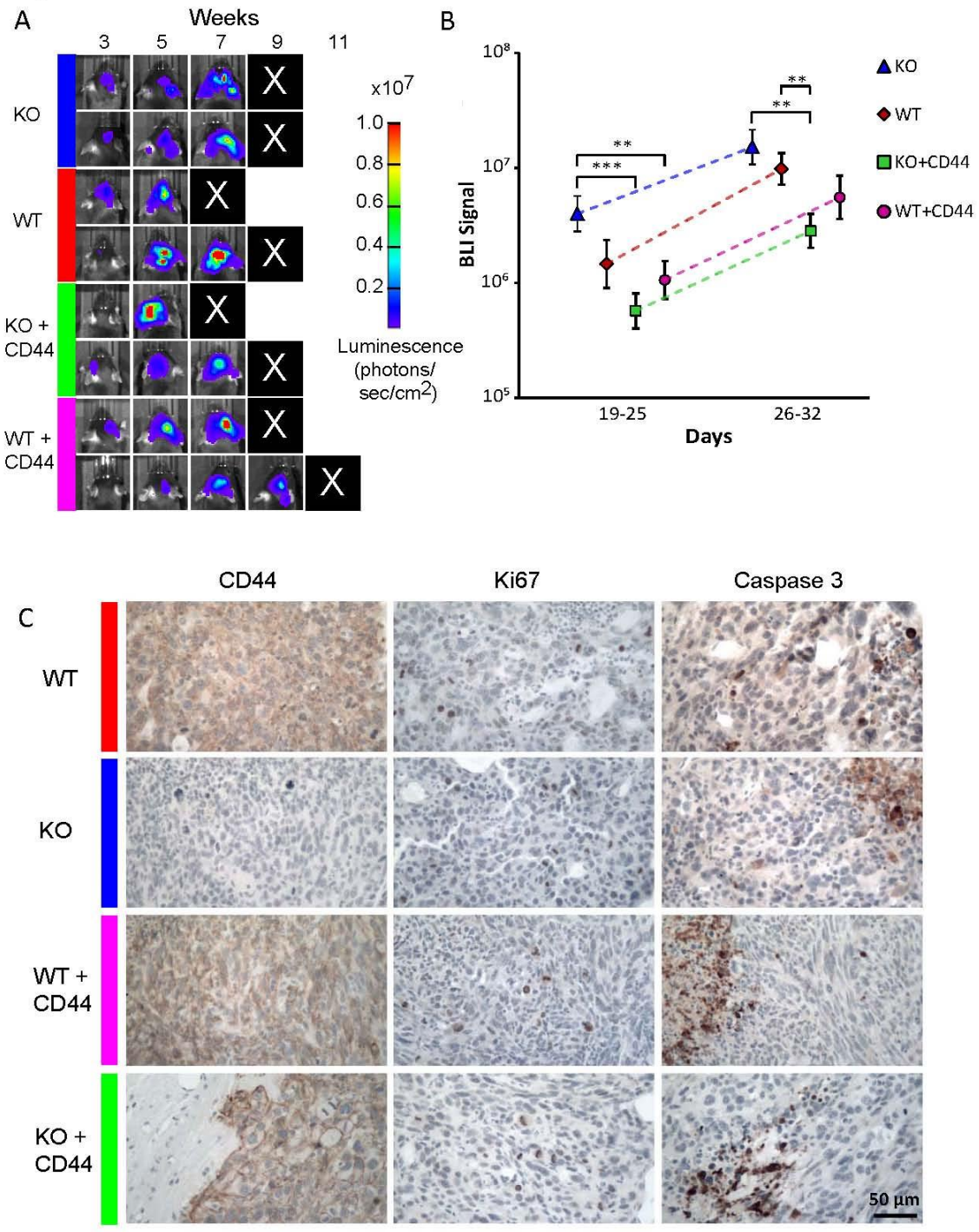

A
Weeks
3
5
7
9
11
KO
WT
KO +
CD44
WT +
CD44
X
x10^7
1.0
0.8
0.6
0.4
0.2
Luminescence
(photons/
sec/cm2)
B
BLI Signal
10^8
10^7
10^6
10^5
**
***
**
**
19-25
26-32
Days
KO
WT
KO+CD44
WT+CD44
C
CD44
Ki67
Caspase 3
WT
KO
WT +
CD44
KO +
CD44
50 µm

Figure 5. Net proliferation rate in the mouse gliomas is independent of CD44 expression level
A. Bioluminescence imaging of the four CD44 expression conditions showing glioma progression over time. B. Quantification of BLI integrated intensity. Mean integrated intensities at weeks 3 and 4 (days 19-25 and 26-32) are plotted ± SEM. At early time points, KO values are highest between all groups (**p<0.01, ***p<0.001). Slopes between weeks 3 and 4 (dashed lines) are not statistically different between all conditions (p=0.88), indicating similar net growth rates. C. Immunohistology confirmed CD44 status and showed no discernable differences in Ki67 or Caspase 3 staining across the four conditions. Left: CD44 positive cells are labeled in brown. Middle: Brown indicates Ki67-positive (proliferating) cells for each of the conditions. Right: Brown indicates Caspase 3-positive (apoptotic) cells for each of the conditions. All slides are counterstained with hematoxylin.

Since CD44 has been linked to uncontrolled cell growth (*33*, *41*), we next examined the effects of CD44 on net tumor cell proliferation. The inclusion of *FLuc* in our animal models (Figure 3A) allowed serial measurements of tumor burden using bioluminescence imaging (BLI). BLI of the four conditions shows the relative progression of the tumors under different CD44 backgrounds (Figure 5A). All four backgrounds were able to establish tumors, indicating that tumor initiation is independent of CD44 content. Integrated BLI signal at weeks 3 and 4 showed that initial tumor burden was greatest for KO animals, on average (Figure 5B). More importantly, changes in integrated BLI signal between weeks 3 and 4 were not different between all cases (p=0.88), meaning that tumor net proliferation rates were indistinguishable across all four conditions (Figure 5B – dashed lines). Growth rate calculations were not conducted after week 4 to avoid potential bias caused by substantial animal death (Figure 5A, Supplemental Figure 2).

Since growth rates were the same for all conditions, it is likely that the initial differences in tumor burden between groups, which did not correlate with outcome, were due to small differences in the number of cells transformed at the time of injection. To further assess net tumor cell proliferation, we also performed IHC staining of all four CD44 conditions for Ki67 and Caspase 3 as indicators of proliferation and apoptosis, respectively. Staining for Ki67 and Caspase3 showed considerable variability in all four conditions and did not seem to correlate with survival or CD44 level (Figure 5C). Therefore, we concluded that inherent differences in net proliferation do not directly explain biphasic survival as a function of CD44 level. As others have found that antibody-mediated CD44 inhibition does not alter proliferation *in vitro*, but does affect migration (*52*), we instead focused on cell migration as an important driver of these cancers.

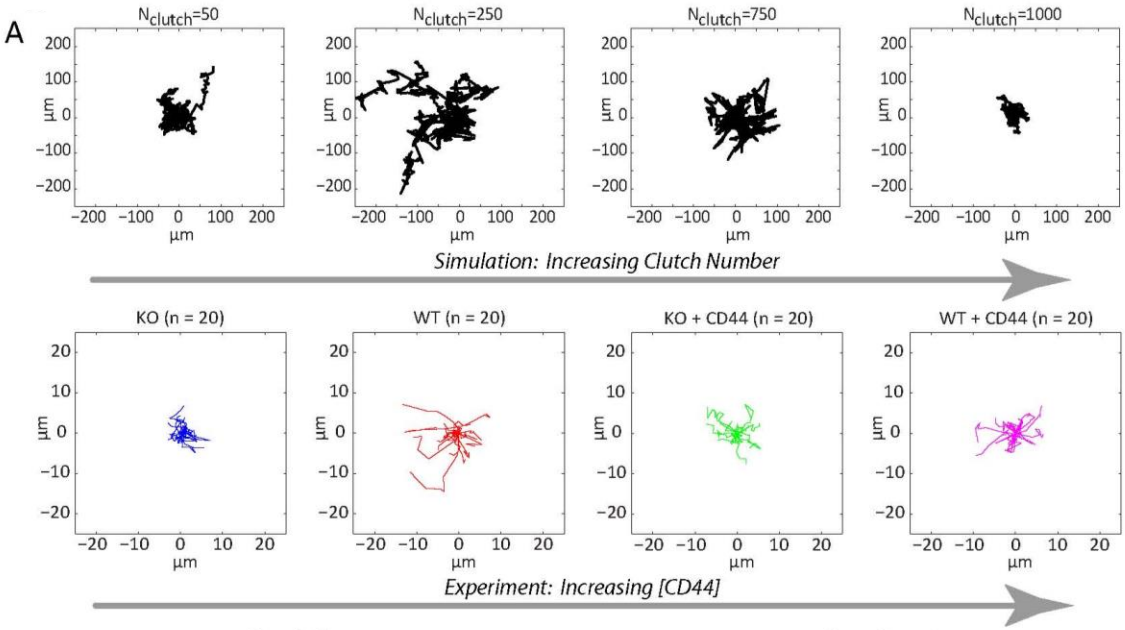
Cell migration is a dynamic physical process involving cell attachment to, and subsequent detachment from, the cell substratum (*53*). As attachment and detachment are opposing processes, it is perhaps unsurprising that maximal cell migration requires an intermediate level of adhesion to the substrate, as demonstrated through seminal theoretical analysis by DiMilla and colleagues in the 1990s (*16*). While the DiMilla *et al.* model established an important principle that was subsequently confirmed experimentally *in vitro* (*17*, *18*), it required *a priori* determination of symmetry breaking to establish cell directionality and assumed an infinitely stiff environment and a constant traction force

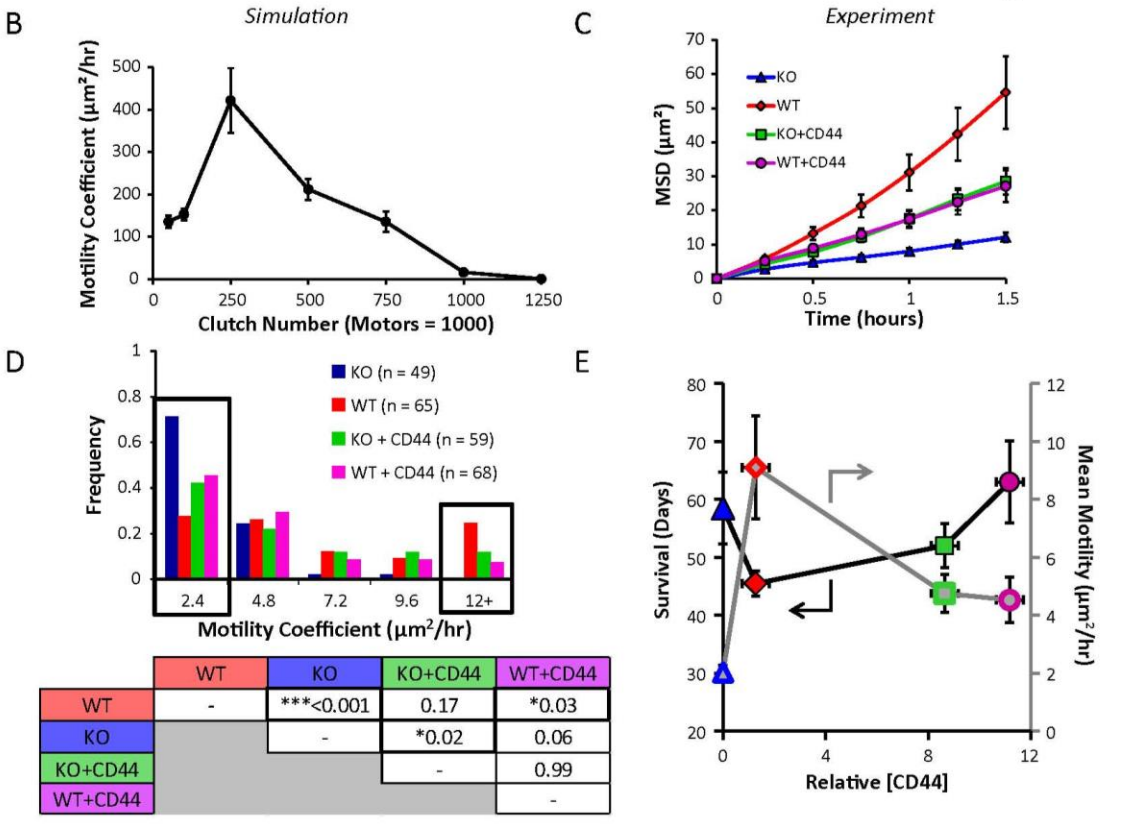
capacity, whereas more recent cell motility models allow for stochastic polarization of the cell, account for compliant microenvironments and molecular motor-mediated force-velocity relationships, wherein increased load on the motors reduces their speed and traction forces rise and fall dynamically. To implement these more realistic assumptions (*22*, *54*), we developed a 2-dimensional whole-cell model based on our previously published "motor-clutch" stochastic Monte Carlo model for cell protrusion and traction dynamics (*22*, *23*) (Supp. Methods, Table 1). In the 2D motor-clutch model, single modules, each representing a cell protrusion, are comprised of myosin motors, stiff actin filaments, and transiently bound "clutches" which are modeled as Hookean springs. Multiple motor-clutch modules are attached to a common cell body where traction forces are integrated to sum to zero in accordance with Newton's Second Law (ignoring inertial forces; Supplemental Figure 3).

Table 1. Simulation parameters of 2D motor –clutch simulation of cell migration

List of parameter symbols, their meanings, and values for the whole-cell simulations used in this paper. Parameters with asterisks (*) represent maxima or minima for the parameters. (Citations in table: (*22*, *55*–*59*))

Symbol	Parameter	Value	Source
N_m	Total number of motors	1000	Adjustable (estimated ~10^2-10^5 based on Molloy *et al.* 1995)
N_c	Total number of clutches	Variable (50-1250)	Adjustable (N_c~N_m based on Bangasser *et al.* 2013 (Biophys.))
A_{tot}	Total possible actin protrusion length	100 µm	Typical cell length observed
v_p*	Maximum actin polymerization velocity	200 nm/s	Chan and Odde 2008
k_{mod}*	Maximum module birth rate	1 s^{-1}	Based on k_{mod} and eq. S8
k_{mod}	Typical module birth rate	0.001 s^{-1}	Must be similar to k_{cap}
k_{cap}	Module capping rate	0.001 s^{-1}	Estimated based on Schafer *et al.* 1996 and experimental observation
l_{init}	Initial module length	5 µm	Adjustable
l_{min}	Minimum module length	0.1 µm	Adjustable
κ_{cell}	Cell spring constant	1000 pN/nm	Adjustable (>κ_{sub})
$n_{c,cell}$	Number of cell body clutches	10	Adjustable (<N_c)
n_m*	Maximum number of module motors	100	Estimated based on Plotnikov *et al.* 2012
F_m	Motor stall force	2 pN	Molloy *et al.* 1995
v_m*	Motor unloaded velocity	120 nm/s	Chan and Odde 2008
n_c*	Maximum number of module clutches	0.1*N_c	Adjustable
k_{on}	Clutch on-rate	0.3 s^{-1}	Chan and Odde 2008
k_{off}*	Clutch unloaded off-rate	0.1 s^{-1}	Lele *et al.* 2008
κ_c	Clutch spring constant	0.8 pN/nm	Bangasser *et al.* 2013 (Biophys.)
F_b	Clutch bond rupture force	2 pN	Jiang *et al.* 2003
κ_{sub}	Substrate spring constant	10 pN/nm	Chan and Odde 2008





	WT	KO	KO+CD44	WT+CD44
WT	-	***<0.001	0.17	*0.03
KO		-	*0.02	0.06
KO+CD44			-	0.99
WT+CD44				-

Figure 6. Single-cell motility in biophysical simulation and in ex vivo brain slices is biphasic with respect to CD44 expression level

A. Upper: Wind rose plots of a representative subset of simulated motor-clutch model single-cell trajectories with increasing clutch number (representing CD44 level) over 10 hours of simulated time. Motor number is held constant at 1000 per cell. Lower: Wind rose plots of 20 randomly selected experimental cell trajectories for each animal condition over 1.5 hours. B. Simulation results demonstrating a biphasic, concave-down dependence of random motility coefficient with respect to numbers of clutches (mean ± SEM). C. A Mean squared displacement (MSD) as a function of time for each of the animal conditions. D. Histogram of random motility coefficients for each tracked cell calculated from single-cell MSDs over 1.5 hours. The table below shows the pairwise comparisons of distributions for each condition using the non-parametric, rank-based Kruskal-Wallis test (*$p<0.05$, ** $p<0.01$, *** $p<0.001$, all with Bonferroni adjustment). Wild-type migration is statistically faster than both KO and WT+CD44, but not KO+CD44. E. Both animal survival (black, left axis) and cell motility (gray, right axis) are biphasic with respect to CD44 expression and anti-correlated with each other such that survival is poorest where migration is fastest.

To test whether the motor-clutch model gives rise to biphasic migration rates as a function of adhesiveness in a compliant environment, we altered the number of clutches (adhesion molecules) while maintaining a constant number of motors. Quantitative analyses of cell displacements (Figure 6A – upper) allowed us to calculate the relationship between mean squared displacement (MSD) and time. From this we calculated the random motility coefficient for each cell, a quantitative measure of dispersive capacity that is directly analogous to a diffusion coefficient (See Methods) (Figure 6B). Consistent with experimental cell observations (Supplemental Figure 4), motors were held constant at 1000 per cell. We found that maximal cell migration occurred when the clutch number was 250. However, deviation from this number of

clutches, while holding number of motors constant, resulted in sub-optimal migration (Supp. Movie 1), leading to a biphasic relationship between clutch number and motility. From these results, we predicted that changing CD44 concentration would effectively change the clutch number, and therefore predicted a biphasic relationship would exist between CD44 content and cell migration rates in our animal models.

To investigate the role of CD44 expression level on glioma cell migration in the mouse brains, we measured the motility of cells in *ex vivo* brain slice cultures (*31*). As shown in Figures 6C and 6D, WT cells exhibited the fastest migration overall and were found to exhibit statistically different migration than KO and WT+CD44 ($p<0.001$ and $p=0.03$, respectively). Although both KO+CD44 and WT+CD44 models exhibited similar motility, KO+CD44 was not found to be different than WT ($p=0.17$), likely due to the presence of a higher percentage of faster moving KO+CD44 cells compared to WT+CD44 (Figure 6C – black boxes). However, when both overexpression cases are pooled together, they are statistically different from WT ($p<0.02$, data not shown). KO animals displayed the least motility of all cases, but were not found to be statistically different from the high CD44 content WT+CD44 condition. Together, these *ex vivo* single cell analyses demonstrate a biphasic relationship between CD44 protein expression and glioma cell motility, where intermediate CD44 content (i.e. WT expression levels) leads to the greatest overall motility (Figure 6E – gray curve). These biphasic migration results are qualitatively consistent with simulation predictions (Figure 6B). Notably,

biphasic migration outcomes are anti-correlated with biphasic survival outcomes, where maximal migration corresponds with minimal survival (Figure 6E – black curve), thus suggesting that CD44 may ultimately control survival by controlling cell migration and invasion in a non-monotonic manner.

Conclusions

In summary, we find that a simple high/low (monotonic) description of the role of CD44 in glioma progression, commonly used for CD44 and cancer gene expression generally, masks important aspects of CD44 gene function. Instead, we find that a low/medium/high, or non-monotonic, description reveals a previously unappreciated biphasic dependence that naturally follows from CD44's known role in cell adhesion and the biophysics of cell migration. More generally, the results motivate the reframing of biomarker discovery, bioinformatics, and cancer gene analysis into a non-monotonic framework, where genes are allowed to exhibit optimality, which is common in both engineered and natural systems.

Materials and Methods

Animals and tumor models

C57BL/6J (CD44+/+, WT) and B6.129(Cg)-CD44tm1Hbg/J (CD44-/-, KO) mice were purchased from Jackson labs (Bar Harbor, ME). Animals were housed in pathogen free conditions under controlled 12-hour light dark cycle with free access to food and water according to Institutional Animal Care and Use Committee approved protocol

0904A64582. Malignant gliomas were induced in neonatal mice by DNA injection into the right lateral ventricle as described previously (*48*). Hypothermia anesthesia was achieved for postnatal day 1-2 mice by placement in ice for 3-5 minutes. Neonatal mice were secured in a cooled neonatal stereotactic frame. Hamilton syringes were fitted with a 30 gauge beveled needle. Neonatal mice were injected with 350 ng of plasmid DNA mixed with polyethyleneimine (jetPEI, Polyplus, Berkeley, CA) in a total volume of 0.7 µL at a rate of 0.3 µL/min. The following plasmids were used: pT2/C-Luc/PGK-SB100, pT2/Cag-NrasV12, pT/CMV-LgTAg, pT2/Cag-NrasV12/CLP-CD44, pKT2/CMVGFP-actin. The CD44 rescue construct was cloned using *in vitro* recombination (InFusion, Clontech, Mountain View, CA) and verified by sequencing and restriction digests. All plasmids were prepped in 0.1x TE buffer (1mM Tris mixed with 1mM EDTA) for *in vivo* transfection using HiPure Maxiprep kit (Invitrogen, Grand Island, NY). Animals were monitored daily for morbidity. Moribund animals were terminally anesthetized by intraperitoneal injection of ketamine/xylazine cocktail followed by transcardial perfusion with 0.1M phosphate buffer solution and 4% paraformaldehyde. Brains were collected and post fixed in 10% neutral buffered formalin and transferred to 80% ethanol prior to paraffin embedding. Log-rank statistical analyses were generated using Graphpad Prism software.

Primary cell lines

Moribund animals were terminally anesthetized with a ketamine/xylazine cocktail and transcardially perfused with isotonic saline. Tumors were collected, minced with a scalpel, and expanded in serum free media as tumorspheres. The media consisted of DMEM/F12 (Gibco, Grand Island, NY) with 0.5x B27, 1x N2, 1% penicillin/streptomycin, and normocin. Twenty ng/ml EGF and bFGF (R&D Systems, Minneapolis, MN) were added to the cultures every 3-4 days and spheres were dissociated every 5-7 days with a mix of non-enzymatic dissociation solution (Sigma-Aldrich, St. Louis, MO) and TryPLE (Invitrogen, Grand Island, NY).

Western Blotting

Cells were plated in T25 flasks at $1x10^6$ cells/flask and allowed to grow for two days in complete growth medium. Cell culture supernatants were removed from the flask and suspended cells collected by centrifugation. Adherent cells were released with non-enzymatic cell dissociation buffer (Life Technologies Inc., Grand Island, NY) and then pooled with non-adherent cells. Total cell pellets were washed once with cold PBS and lysed in modified RIPA buffer (50 mM Tris HCL pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% lauryl sulfate, 100mM sodium vanadate, and Complete Mini protease inhibitors – Roche Applied Science, Indianapolis, IN) on ice. Lysates were sonicated for 30 seconds and centrifuged at 4^oC for 10 min at 14,000 rpm to remove debris. Lysates were quantitated by BCA (Thermo Fisher Scientific, Pittsburgh, PA) and

equal protein loaded on the gel. After membrane transfer, the protein bands were imaged using standard western blot techniques and ECL (Thermo Fisher Scientific, Pittsburgh, PA). Anti-mouse CD44 (IM7) antibody was purchased from BD Biosciences (San Diego, CA), and HRP-conjugated anti-sheep secondary antibody was purchased from R&D Systems (Minneapolis, MN); anti-α tubulin monoclonal antibody was purchased from EMD Millipore (St. Louis, MO); HRP-conjugated anti-mouse and rabbit secondary antibodies were purchased from Jackson Immunoresearch (West Grove, PA); antibody against Nonmuscle Myosin Heavy Chain (NMHC) isoforms II-A, -B, and -C were purchased from Covance (Denver, PA).

Analysis of published human data

CD44 expression data (Phillips *et. al* 2006 (*49*)) was accessed from the Gene Expression Omnibus accession number GSE4271. From this database, gene NM_000610, corresponding to CD44, was selected for analysis. To be consistent with the analysis by Phillips *et al*., CD44 expression was normalized to the mean of three Rab14 genes (IDs: 200927_s_at, 200928_s_at, 211503_s_at). Using this method, each patient had a unique normalization factor. Survival data was available for 77 patients. Data for patients without endpoint survival (n=23) was not included in the analysis. Regression analyses for linear and quadratic fits were completed using R statistical software.

Bioluminescence

Animals were monitored for tumor progression using noninvasive bioluminescence imaging. Animals were injected intraperitoneally with 100 µl of 28.5 mg/ml luciferin (GoldBio, St. Louis, MO). Mice were anesthetized using 3% isofluorane and imaged on an IVIS 50 instrument (Xenogen, Alameda, CA). Images were acquired ten minutes after injection with a five minute exposure (Xenogen LivingImage, Alameda, CA). Bioluminescent signals were quantified using the LivingImage software by selecting a circular Region of Interest of 2.7 cm in diameter and analyzed using Graphpad Prism software. Integrated intensity signals across conditions were analyzed by one-way ANOVA. For statistics on tumor growth rates, signal changes between weeks 3 and 4 were calculated for each animal, and the distributions of delta values were compared by one-way ANOVA.

Immunohistochemistry

Unstained brain tumor sections (4 µm) were de-paraffinized and rehydrated using standard methods. For CD44 and SV40LgT antigen retrieval, slides were incubated in pH 9.0 buffer (BORG Decloaking reagent, Biocare Medical, Concord, CA) in a steamer for 30 minutes at 95-98°C, followed by a 20 minute cool down period. Slides were rinsed between steps in running tap water followed by immersion in 1x Tris-buffered saline/0.1% Tween-20 (TBST; pH 7.4). Endogenous peroxidase activity was quenched by slide immersion in 3% hydrogen peroxide solution (Peroxidazed, Biocare Medical,

Concord, CA) for 10 minutes. Rodent Block M (Biocare Medical, Concord, CA) was placed on all sections for 30 minutes. Blocking solution was removed and slides were incubated overnight at 4°C with the appropriate primary antibody diluted 1:400 in 10% blocking solution/90% TBST. The primary antibodies were bHABP (Calbiochem, Billerica, MA), rat anti-mouse CD44 IgG (IM7, BD Biosciences, San Diego, CA), SV40TAg (v-300, Santa Cruz Biotechnology), Ki67 (Abcam, Cambridge, MA), and cleaved-caspase 3 (Biocare Medical, Concord, CA). CD44 was detected with Rat on Mouse AP-Polymer (Biocare Medical, Concord, CA) and SV40TAg was detected with Rabbit on Rodent Polymer (Biocare Medical, Concord, CA) using the manufacturer's specifications. The chromagens used were diaminobenzidine (DAB) (Covance, Princeton, NJ) and Vulcan Fast Red (Biocare Medical, Concord, CA) following manufacturer's specifications. For double labeling of SV40LgT/CD44, the same protocol was followed as individually starting with CD44 (DAB). When the first stain was complete, denaturing solution (Biocare Medical, Concord, CA) was applied for 5 minutes at room temperature followed by Rodent Block M, which was reapplied for 15 minutes, and then the SV40LgT was detected as before. When staining was finished, all slides were rinsed in TBST and counterstained with CAT Hematoxylin diluted 1:3 (Biocare Medical, Concord, CA) for 5 minutes, rinsed in tap water, dehydrated and coverslipped. Slides were scanned and downloaded using the Aperio ePathology System or images were captured on a Leica DMI6000B inverted microscope with Retiga 2000R camera with QCapture software.

Stochastic whole-cell model

The stochastic whole-cell migration model was developed based on previously published motor-clutch models from our group (*22*, *23*, *60*). Similar to these models, the whole-cell model uses a Monte Carlo approach to simulate the interaction of molecular motors and compliant molecular adhesions (i.e. "clutches") that stochastically engage and disengage based on the kinetic rate constants. However, the previous model only simulated the mechanical interactions of individual cell protrusions, which we term "modules". Our whole-cell model simulates an entire cell by connecting several motor-clutch modules together at a central cell body node, and then balancing the forces at this node (Supplemental Figure 3).

Each cell contains j modules, each of which follows the principles previously described (*23*). Each module contains $n_{c,j}$ clutches which bind the F-actin cytoskeleton to a compliant sub-cellular substrate at a constant on-rate, k_{on}. Clutches unbind according to Bell's Law (*61*), where unbinding becomes more likely as the force on each clutch increases.

$$k_{off\,(i,j)} = k_{off}^{*} \exp\left(\frac{F_{clutch(i,j)}}{F_{bond}}\right) \qquad \text{(Eq. 1)}$$

Here, k_{off}* is the off rate in the absence of force, $F_{clutch(i.j)}$ is the force on the i^{th} clutch in the j^{th} module, and F_{bond} is the characteristic bond rupture force.

Clutches are modeled as Hookean springs, so that the force on each clutch is

$$F_{clutch(i,j)} = \kappa_c (x_{i,j} - x_{sub,j}) \quad \text{(Eq. 2)}$$

where κ_c is the clutch spring constant and $x_{i,j}$ and $x_{sub,j}$ are the positions of the i^{th} clutch and the underlying substrate, respectively for the j_{th} module.

Molecular motors within each module drive F-actin retrograde flow at a velocity described by

$$v_{m,j} = v_m^* \left(1 - \frac{F_j}{n_{m,j} F_m}\right) \quad \text{(Eq. 3)}$$

where v_m* is the unloaded velocity of retrograde flow, $n_{m,j}$ is the number of motors within the j^{th} module, F_m is the force required to stall one motor, and F_j is the force magnitude exerted on the module. Here, the force on the module is equal to the force on the substrate spring, which at equilibrium must balance the summed forces of the clutches and the force on the cell center, defined by

$$F_j = \kappa_{sub} x_{sub,j} = \kappa_c \sum_{i=1}^{n_{c,j}} x_{c,i} = \kappa_{cell} x_{cell} \quad \text{(Eq. 4)}$$

where κ_{sub} is the substrate stiffness, $x_{sub,j}$ is the substrate position underlying a particular j^{th} module, κ_c is the clutch spring stiffness, $x_{c,i}$ is the extension of the i^{th} clutch molecule within the module, κ_{cell} is the stiffness of the cell center, and $x_{cell,j}$ is the extension of the cell spring for the j^{th} module.

The cell center also must have clutches ($n_{c,cell}$) to capture the drag due to the cell body. The force associated with the cell body clutches is given by the sum of the displacements of all the associated clutches multiplied by the clutch spring constant, which, again, must be balanced by the opposing force of the substrate spring at equilibrium. This relationship is defined by

$$F_{cell} = \kappa_c \sum_{i=1}^{n_{c,cell}} x_{c,i} = \kappa_{sub} x_{sub,cell} \qquad \text{(Eq. 5)}$$

Overall, all of the forces acting on the cell must sum to zero. Therefore,

$$F_{cell} + \sum_{j=1}^{n_{mod}} F_j = 0 \qquad \text{(Eq. 6)}$$

However, with these rules, unloaded actin would flow rearward without limit, so we imposed rules for actin polymerization and depolymerization, taking into account the finite actin pool. We assumed that F-actin polymerization speed at the leading edge of a module depends on the proportion of available G-actin, A_G, according to the equation

$$v_p = v_p^* \frac{A_G}{A_{tot}} \qquad \text{(Eq. 7)}$$

where v_p is the velocity of polymerization, v_p* is the maximal polymerization velocity, and A_{tot} is the total actin pool, in units of distance. Briefly, actin was represented in units of distance rather than monomers to accommodate the fact that a single protrusion has many actin filaments that are branched and/or bundled together. To simplify the

geometry of this problem, the length of actin was used to represent actin's contribution to the length of the module. Polymerization was halted by module capping at rate k_{cap}. F-actin automatically depolymerized once it was retrogradely transported beyond the location of the motors, at which point the corresponding 'distance' of actin was returned to the G-actin pool.

Likewise, the lifespan of a protrusion is not infinite within a cell, so we also included module birth and death criteria. Modules of initial length l_{init} were generated at rate k_{mod}, which is defined as the product of the maximal birth rate, k_{mod}*, and the proportion of available G-actin raised to the fourth power (*62*, *63*), as described by

$$k_{\mathrm{mod}} = k_{\mathrm{mod}}^{*}\left(\frac{A_G}{A_{tot}}\right)^4 \qquad \text{(Eq. 8)}$$

When a new module is generated, it recruits a fraction of the maximal number of motors and clutches (n_m* for motors, n_c* for clutches) based on the available pools, where N_m is the total number of motors per cell and N_c is the total number of clutches per cell

$$n_{m,j+1} = n_m^{*}\,\frac{N_m - \sum_{j=1}^{n_{\mathrm{mod}}} n_{m,j}}{N_m} \qquad \text{(Eq. 9)}$$

$$n_{c,j+1} = n_c^* \frac{N_c - \left(\sum_{j=1}^{n_{\mathrm{mod}}} n_{c,j} \right) - n_{c,cell}}{N_c} \quad \text{(Eq. 10)}$$

The orientation of the new module was assigned by randomly selecting a vector emanating radially from the cell center. Module death occurred when the length of the module was below l_{min}. Once modules were destroyed, all associated motors, clutches, and remaining actin were returned to their respective free pools.

To start the simulation, cells were initialized with 3 modules, a total actin pool of 100 μm, 1000 total motors, and variable number of total clutches. Simulations used a variable time-step Gillespie (*64*) Stochastic Simulation Algorithm (SSA), in which exactly one event happens per time-step. Despite numerous possible events (i.e., clutch binding, clutch unbinding, module birth, or module capping), only two random numbers must be generated per iteration to determine 1) the time increment and 2) the event that occurred, defined by

$$t_{event} = \frac{-\ln(rand_1)}{\sum_{i=1}^{n} k_i} \quad \text{(Eq. 11)}$$

$$\sum_{j=1}^{i-1} k_j < rand_2 \sum_{j=1}^{n} k_j < \sum_{j=1}^{i} k_j \quad \text{(Eq. 12)}$$

where n is the total number of possible events and k_i is the rate constant of the i^{th} possible event. For each iteration of the simulation, the random event was determined, as well as the elapsed time, t_{event}, for that event to occur. Next, F-actin velocities were calculated for each module and locations were updated based on the time elapsed by t_{event}. The displacement of the F-actin necessitated updating positions for clutches and substrate deformation. Additionally, the reference points for uncapped modules were updated to account for F-actin polymerization. Finally, the selected event was executed, after which modules shorter than l_{min} were terminated and the resulting force-balance relationship was used to determine the new cell body position and all spring extensions. This process was iterated until simulated time reached 10 hours.

Using the coordinates of the simulated cell center over time, we calculated the random motility coefficients according to the equation

$$\mu = \frac{\langle r^2 \rangle}{4t} \qquad \text{(Eq. 13)}$$

where $< r^2 >$ is the 2D mean squared displacement of the cell center over time interval t, and 4 describes the degrees of freedom based on the dimensionality of the problem (in this case 2-D). The value of μ for each simulated cell can be found from the slope of the mean squared displacement (MSD) vs. time plot, which is generated using the overlapping method (*65*). Slope was calculated based on the weighted uncertainty in each data point in MSD vs. time (data points for higher t have greater uncertainty due to finite

duration of the cell movement trajectories). The mean random motility coefficient μ for each condition was obtained by averaging the slopes for all simulated cells. For this study, we calculated μ for several different total clutch numbers (N_c) to determine the effects of clutch number on overall cell migration.

Single cell migration dynamics in brain slice preparations

Cohorts of WT, KO, KO+CD44 and WT+CD44 mice transformed with SV40LgT and NrasG12V and co-expressing eGFP-actin were used for migration analysis (Figure 3A). Mice bearing eGFP-actin tagged tumors were sacrificed when bioluminescence signals were in the range of $5x10^5$ to $8x10^6$ photons/second. Mice were terminally anesthetized with ketamine/xylaxine and perfused transcardially with isotonic saline. Brains were submerged in 2% agarose gel molds, allowed to solidify on ice for 30 minutes, and sectioned into 300 μm thick sagittal slices using a vibrating blade microtome (Leica, Buffalo Grove, IL). Slices were placed in chilled DMEM until the brain was completely sectioned. GFP visualization goggles were used to identify slices containing tumor cells expressing eGFP-actin. These slices were placed on 0.4 μm Millipore culture inserts inside glass bottom petri dishes (MatTek Corporation, Ashland, MA) containing DMEM as described previously (*31*). Slices were imaged on a Zeiss LSM 7 Live swept-field laser confocal microscope at 15 minute intervals for up to 20 hours in 5% CO_2 and humidified conditions (Carl Zeiss Microscopy, Jena, Germany). Images were collected with a 10x Plan-ApoChromat, 0.45 NA objective lens. Z stacks

composed of 10-20 planes were collected at 12 μm intervals and collapsed to a single image using maximum intensity projection in the z-direction. Images were registered by an affine transformation using ImageJ *StackReg* plug-in (École Polytechnique Fédérale De Lausanne) to account for stage drift and tissue relaxation. Tracking of individual cells was automated in Matlab for reduced subjectivity. Two-dimensional Fast Fourier Transform (FFT) filtering was conducted on images to remove low frequency background fluorescence variations. Cells were outlined and filled in Matlab using a Laplacian of Gaussian *edge* function and *imfill* function, respectively. The centroid coordinates of the resulting shape were tracked throughout the image series (Supp. Movie 2). Wind rose plots were generated by plotting 20 randomly selected centroid trajectories initialized at the origin for an elapsed time of 90 minutes. Cell centroid coordinates were used to calculate the mean squared displacement (MSD) of the cells over time by the overlapping method (*65*). To quantify the dispersion of cells, the MSD over time was used to calculate the random motility coefficient μ according to equation 13 (assuming 2-D geometry, evaluated at 90 minutes). Distributions of random motility coefficients for the different conditions were compared using a non-parametric Kruskal-Wallis test.

Supplementary Discussion

Possible mechanisms of CD44-mediated effects on migration in the mouse brain

The biphasic migration theory that we used to explain our results above was based on a previously published computational cell migration model by DiMilla *et al*. wherein

an optimal, intermediate adhesiveness was found to result in the fastest cell speeds. Adhesiveness, κ, was defined as the quotient of substrate ligand concentration and the dissociation constant of the receptor-ligand complex (K_d). By altering the concentration of adsorbed laminin proteins, DiMilla *et al*. confirmed the model-based prediction of migration with respect to substrate adhesiveness *in vitro* using myoblast cells (*16*). Similar biphasic relationships have since been demonstrated in human smooth muscle cells (*17*) and CHO cells (*18*). Our study is the first to our knowledge to demonstrate such biphasic migration dependence in an *ex vivo* disease model.

DiMilla *et al.*'s work and other subsequent studies specifically investigated the effects of changing substrate ligand concentration. Our study uniquely identifies a biphasic migration trend with respect to adhesiveness through changes in receptor number. Conceptually, the two mechanisms are related, since changing either receptor or ligand concentration changes the ability of the cell to adhere to its substrate. In either case, if adhesiveness is too great, the cell is unable to release rearward adhesions, and if adhesiveness is too small, the cell is unable to maintain adhesions to support the load of pulling motors. However, both models rely on the mechanical coupling of the myosin motor pulling forces to the underlying substrate, mediated by adhesive receptors. In this study, we implicate CD44 as a potential molecular "clutch" in mediating migration, consistent with these models and previous studies documenting its well-known role in cell-substrate adhesion (*40, 41, 66*).

However, it is important to also acknowledge CD44's many other potential functions related to motility. First of all, CD44 is a known signaling moiety that can activate Rho family proteins to affect focal adhesion assembly and cytoskeletal rearrangements (*33*, *41*, *67*, *68*), which have been shown to contribute to glioma cell invasion (*26*, *69*). While it is unclear here if CD44 is operating as a signaling entity or mechanical clutch, recent *in vitro* work by Kim and Kumar demonstrated that the addition of soluble HA to glioblastoma cells in *in vitro* transwell assays reduced cell migration similar to that of BSA blocked controls, suggesting that, although CD44 groups are interacting with HA, it is the mechanical coupling of CD44 to substrate-bound HA that leads to increased invasion, rather than through CD44-HA-mediated signaling (*66*). In future studies, it would be interesting to add soluble HA or HA antibodies directly to brain slices to measure the effects of blocking CD44-HA mechanical coupling.

A second potential mechanism of CD44-mediated motility is through modification of integrin binding. Glycoproteins such as CD44 have been suggested to both facilitate integrin clustering (*70*), and increase the probability of ligand binding (*71*). Through these mechanisms, CD44 could modulate the K_d for integrin binding to extracellular matrix, and as predicted by DiMilla *et al*. (*16*), result in a biphasic migration relationship. Similarly, CD44 has been shown to associate with MMP-9 (*72–74*), which could potentially be degrading matrix to allow for cell invasion. This would alter the local ligand concentration, which again would predict a biphasic relationship based on

the theory of DiMilla *et al*.. Therefore, CD44 may not be working directly as a molecular clutch, but indirectly affecting the adhesiveness of the cell to its substrate.

Additionally, there are a few other important considerations that could be contributing to CD44-dependent motility, but were not investigated directly. For instance, modification of CD44 through cleavage (*74*, *75*), splice variants (*76*) or glycosylation (*77–79*) could alter CD44's ability to bind HA and/or interact with the cytoskeleton. As we hypothesize mechanical coupling to be to the basis of CD44-mediated motility, these mechanisms could potentially facilitate or inhibit migration. However, these caveats do not change our results, but merely their molecular mechanistic underpinnings. It is worth noting that analysis of CD44 splicing in glioma found that unspliced, full-length standard CD44 (CD44s) was the only detectable species, and that CD44 variants (CD44v3, CD44v4/5, CD44v6) were undetectable (*46*). The exact molecular mechanism by which biphasic migration dependence on CD44 is achieved is currently unclear and will be investigated in future studies, but here we introduce a minimalist model that suggests CD44-dependent adhesiveness is sufficient to explain these results.

Relationship to previous analyses of CD44 in glioma

Our biphasic survival results are consistent with published results in the glioma literature, and can even possibly help resolve apparent controversies. Analysis by Bhat *et al*. showed proneural subtypes are low in CD44, whereas mesenchymal subtypes are high in CD44 (*45*), which is consistent with our analysis of the Phillips dataset (*49*). Pietras *et*

al. used a PDGF-driven proneural model to assess the survival of CD44-/-, CD44+/-, and CD44+/+ mice and found that animals homozygous for CD44 had the worst survival outcomes, followed by animals heterozygous for CD44, and then KO animals (*44*). Our study found a similar negative linear trend of survival as a function of CD44 level within the proneural subtype for human patients. Whereas we found a positive correlation in mesenchymal subtypes, Pietras *et al*. found no correlation in mesenchymal subtypes. However, they used categorical methods and had very few mesenchymal 'CD44 low' subjects to compare. Pietras *et al.* also found that human patients with lower CD44 had better overall survival than the rest of the glioblastomas analyzed based on data from The Cancer Genome Atlas (TCGA) (*44*), consistent with our categorical results. Correlating patient survival with histological specimens, Jijiwa *et al.* found that survival outcomes are worse for patients staining positively for CD44 compared to those who stained negatively (*43*). Similarly, analysis of the Rembrandt database by Anido *et al.* led them to conclude that upregulation of CD44 together with Id-1 corresponded with poor survival outcomes compared to all glioblastomas (*42*). However, it should be noted that the analysis by Anido *et al*. was confounded by the overexpression of Id-1, which has been shown to positively correlate with glioma cell invasiveness (*80*), and could therefore be contributing to their findings.

All of the studies in the paragraph above noted a negative correlation between survival and CD44 level. One explanation for these data is that a true negative correlation

exists, and as suggested by Wei *et al.* who demonstrated that survival outcomes *improved* with higher CD44 expression within glioblastoma patients, the increased survival of highly expressing patients could be confounded by the fact that cells lower in CD44 are more resistant to chemotherapeutic alkylating agents (*47*). In our analysis of human survival we did find a slight negative slope in the data, although it was not statistically significant ($p=0.15$), a finding that is also supported by results from Ranuncolo *et al.*(*46*). Another explanation is that there are relatively few very high CD44 expressing patients to show the upward trend that we saw in the Phillips *et al.* and TCGA datasets and in our mouse studies, so other studies may have only captured the initial downward phase of the curve. This could hypothetically be because the highly expressing cases are less common, or because endpoint reporting in ongoing studies is biased towards the faster-perishing subjects, and an incomplete study may not capture the long-lived low CD44/high CD44 patients. In fact, we saw evidence of potential survival bias in our analysis of the TCGA dataset. While the biphasic trend was significant within the deceased subjects (and thus the bias was not strong enough to affect our results in this case), we note that the variance in expression level between deceased and surviving subjects was statistically different, with surviving patients having a broader distribution (Supplemental Figure 1). This suggests that the surviving group has a larger contribution to the high/low extremes, and thus not allowing sufficient time to collect data for the long-lived patients could bias the analysis. Therefore, we expect that as the TCGA study comes to completion, there will be more long-lived patients at the low and high CD44 levels, and less at intermediate level,

which will then further reinforce the biphasic effect. Additionally, if one compares our mouse model data in the context of the human data from Phillips *et al*., the overexpressing mouse conditions had ~10x higher CD44 concentrations compared to WT. That would mean, with our experiments, the overexpressing conditions would give rise to values of 10,000 or more (10x average expression of ~1000) for human data (Figure 4C); however, all patients in the Phillips *et al.* study had expression much lower than this, suggesting that it is possible that many of the other studies are not capturing the upward trend of the highly expressing cases because these cases are less prevalent. In the limit, one might also expect such very high expressing patients could be diagnosed with lower grade gliomas, or even be asymptomatic.

However, the larger problem is that all of the above studies are based on categories that require arbitrary thresholds to distinguish "high" and "low" groups. Others have alluded to this problem, specifically in IHC, that research groups using different intensity cut-offs could reach inconsistent conclusions, even within the same disease (*41*, *46*). In addition to the thresholding issue, by categorizing/binning data, valuable information is lost that can potentially be informative of the mechanism. When placed on a continuum, it is easier to discern: a) the error in the trend, b) the sensitivity of the trend (e.g. slope), and c) the shape of the trend (linear, exponential, quadratic, etc.). For example, from the above studies one could only state that outcomes are worse on average for patients with CD44 expression above an arbitrary threshold compared to those below

the threshold, but this does not inform clinicians about how common it is for a high CD44 expressing patient to have a long survival (or the opposite), how a modest change in CD44 would be expected to affect a patient, or how extremely high CD44 levels could actually be somewhat beneficial.

Implications for bioinformatics, biomarker discovery, genetic analysis, and therapeutic strategy

For our human data analysis, we used the published data from Phillips *et al.* (*49*) because it provided readily available, genome-wide, quantitative mRNA expression data and subtype information. We attempted to conduct similar analyses using the NCI-hosted REMBRANDT database, but encountered obstacles. Although the database gives users Kaplan-Meier survival data categorized by high, intermediate, and low expression, the current design does not allow users to access and display all of the data on a patient-by-patient basis, for example in a scatterplot as shown in Figure 4C (note: the REMBRANDT NCI portal has recently been retired and transferred to the Georgetown Database of Cancer).

If CD44 inhibition is to be used as a therapeutic strategy for cancer, as has been suggested by many (*40*, *52*, *81*, *82*), and recently tested in clinical trials (*83*), it should be noted that partial inhibition of CD44 in patients that are highly overexpressing could potentially drive tumors into a hyper-migratory state, thus promoting progression. In the clinic, a biphasic model predicts that intermediate CD44 level patients could benefit from

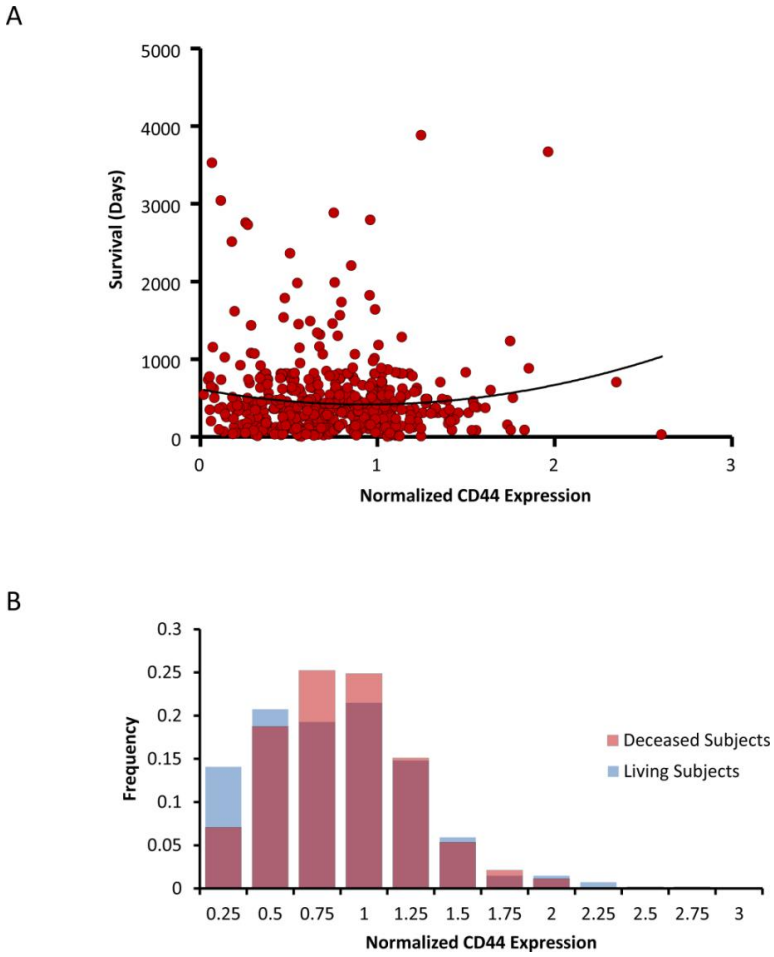
CD44 inhibition, while low CD44 level patients would not benefit nearly as much, and high CD44 level patients could actually worsen if they are reduced to intermediate levels. If patients from all three regimes are averaged together in a clinical trial, the trial would likely fail; however, by stratifying patients to select those most likely to benefit, i.e. the intermediate CD44 level patients (normalized CD44 level between ~1500 and ~3500, Figure 4C), the clinical trial would be more likely to succeed and provide some measure of relief for the patients with the most dismal prognoses. Alternatively, subtyping could be used to stratify patients, as our analysis revealed a negative correlation in proneural types, no correlation in proliferative types, and a positive correlation in mesenchymal types. Therefore, targeted CD44 inhibition of higher expressing proneural patients would most likely provide a benefit, whereas mesenchymal patients would likely experience a detrimental effect.

Final Remarks

Previous literature has documented a somewhat confusing and apparently contradictory dependence of survival on CD44 in glioma, if one assumes a monotonic trend. Here we demonstrated a simple biphasic relationship between CD44 level and survival outcome in both animal models and human glioma patients – a result that would be, and has been, obscured by standard monotonic analysis methods. Similarly, there is debate over the effects of CD44 in many other cancers (reviewed extensively by Naor (*41*)). We speculate that some of these cancers could also exhibit a biphasic trend with

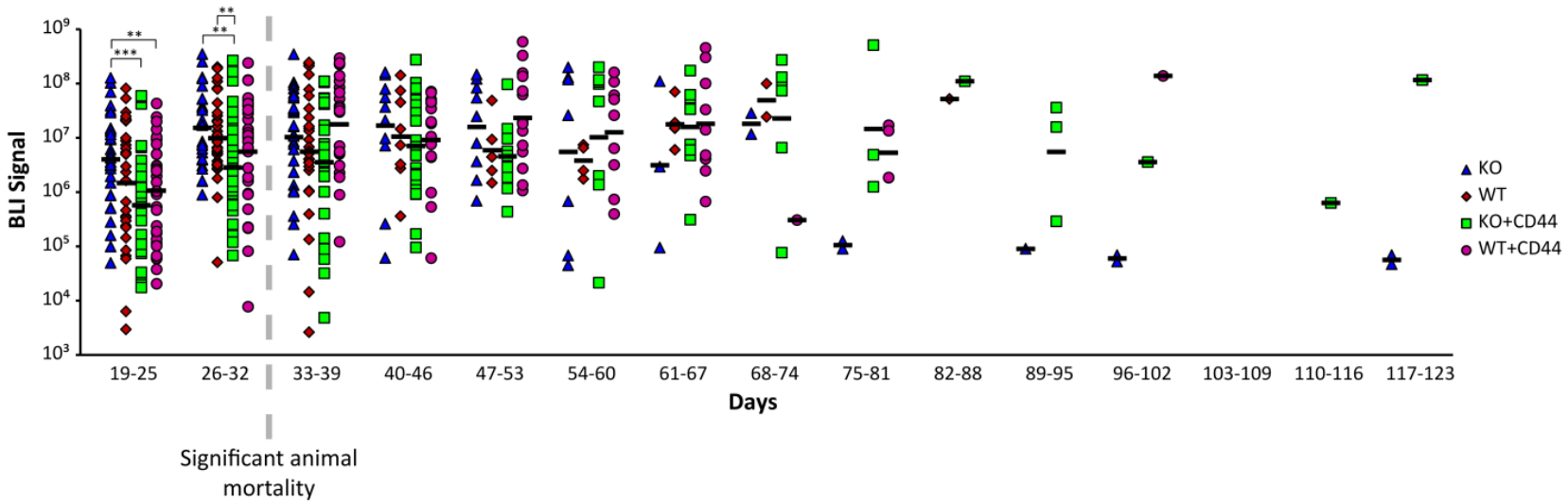
CD44 expression similar to glioma that has been obscured for the same reasons outlined above. Moreover, there are likely other molecules that exhibit a biphasic trend, such as periostin, which was recently found to induce biphasic migration in pancreatic cancer cells *in vitro* (*84*). Generally, it may be that there are many other biomarkers, such as CD44 in glioma or periostin in pancreatic cancer, that lack a linear (or monotonic) prognostic relationship, but yet have biphasic (or non-monotonic) prognostic value. Therefore, in the era of "big data" and the search for novel biomarkers, it will be important to connect genomic/proteomic analyses to biophysical models and, where appropriate, consider the possibility of higher order relationships and optimality.

Supplemental Figures



Supplemental Figure 1. TGCA survival analysis

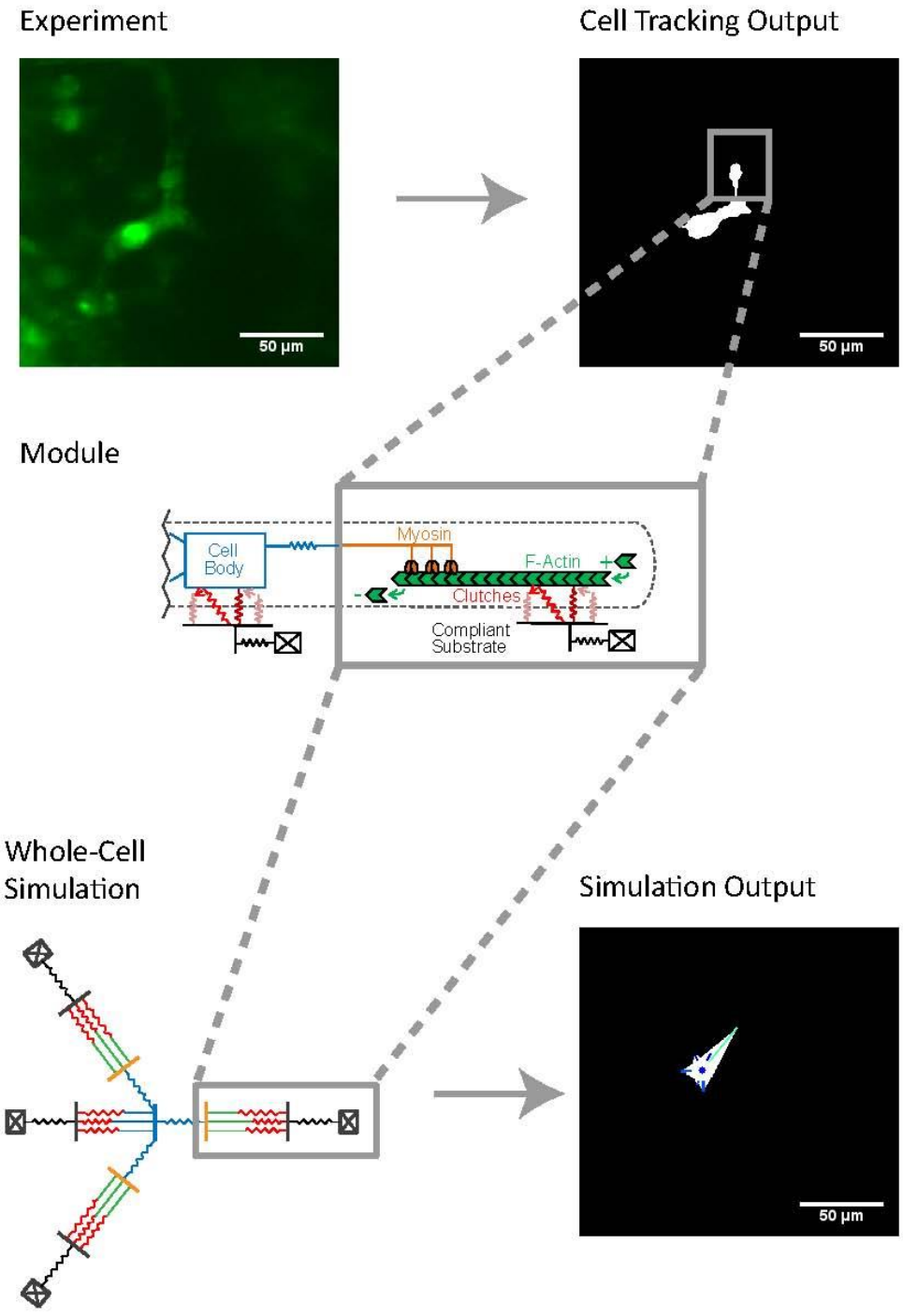
A. Continuum plot of survival time vs. Rab14-normalized CD44 expression of deceased patients from the TCGA database. Black curve shows the statistically significant best-fit quadratic equation (p=0.018). Similar to figure 4, a linear fit of the data was not statistically significant (p=0.694). B. Overlaid histograms of CD44 expression for deceased and living patients available from the TCGA database. F-test comparisons of the distributions revealed that living subjects have a significantly broader distribution (p=0.021), indicative of higher survival potential at the high/low expression extremes and the potential bias that may be enlisted from analyzing incomplete survival datasets. It is also evident that living subjects are more likely to be high or low expressors, while deceased subjects are more likely to be intermediate expressors.


BLI Signal
10^9
10^8
10^7
10^6
10^5
10^4
10^3
**
***
**
**
19-25
26-32
33-39
40-46
47-53
54-60
61-67
68-74
75-81
82-88
89-95
96-102
103-109
110-116
117-123
Days
KO
WT
KO+CD44
WT+CD44
Significant animal mortality

Supplemental Figure 2. Bioluminescence imaging

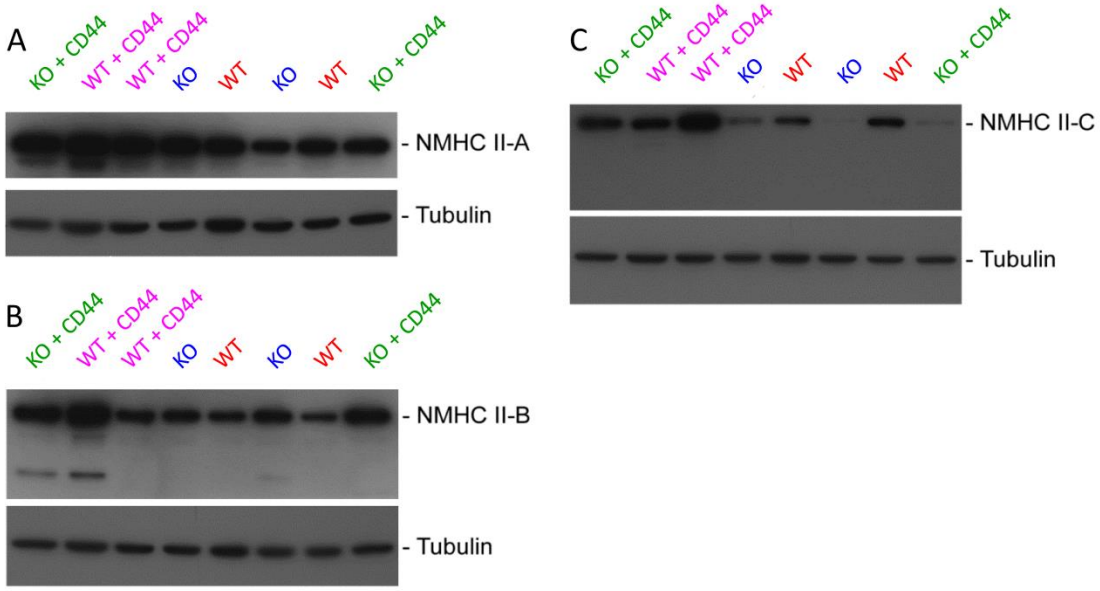
BLI data for each animal at each time point is plotted. Means are indicated by black bars (corresponding to points in figure 5B). Significant animal mortality occurs after week 4, indicated by dashed gray line, so this data was not analyzed for differences in tumor growth rate.

Fig. S3


Experiment
Cell Tracking Output
50 µm
50 µm
Module
Cell Body
Myosin
F-Actin
Clutches
Compliant Substrate
Whole-Cell Simulation
Simulation Output
50 µm

Supplemental Figure 3. Whole-cell model schematic

Upper left panel shows confocal experimental output from a WT animal brain slice experiment. A semi-automated Matlab algorithm takes the input images from the experiment and outlines the cell shape (white) and tracks it through time (Upper right panel). A protrusion in the experiment is regarded as a "module" in the whole-cell simulations (Middle panel/gray box). Each module is mechanically connected to the cell body and anchored to the compliant substrate (black) via molecular clutches (pink = unbound, red = bound), which are modeled as Hookean springs that reversibly engage and disengage. Myosin motor activity (orange) causes F-actin (green) to flow retrogradely, which displaces clutches and builds up tension. A clutch under tension deforms the underlying compliant substrate until it fails to support the load and unbinds from the F-actin, redistributing the load amongst the remaining bound clutches. Module lengths are determined by the F-actin length in each protrusion. As in an actual cell, F-actin undergoes a treadmilling effect, where G-actin monomers are added to the F-actin barbed ends located near the leading edge plasma membrane, and F-actin undergoes net disassembly toward the cell center. Multiple modules are attached to a single cell body, which is itself anchored to the substrate through molecular clutches (Bottom left panel). The total number of motors, clutches, and actin subunits is conserved throughout the entire cell. Cell location and estimated cell shape based on module positions are tracked throughout time. A sample output is shown in the bottom right panel. The white is the cell outline, the blue dot represents the cell body location, and the blue/green lines represent the individual modules.



Supplemental Figure 4. Western blots of non-muscle myosin heavy chain isoforms

Western blotting of non-muscle myosin heavy chain isoforms was conducted on our animal model cell lines to test the model assumption that motor numbers were unchanged between conditions. Whole cell lysates were prepared from the indicated cell lines and analyzed by western blot for expression of isoforms of Nonmuscle Myosin Heavy Chain (NMHC). A. NMHC II-A. B. NMHC II-B. C. NMHC II-C. Each western blot was analyzed separately for tubulin expression as a loading control. While blotting is not constant between all cell lines, variations in myosin content do not appear to directly correlate with CD44 concentrations, and do not vary to the same extent as CD44 (up to 14x WT). The results justify our model assumption that the number of motors is not changed as the CD44 level, i.e. number of clutches in the model, changes.

Supplemental Movie Captions

Supplemental Movie 1. Whole-cell simulations of changing clutch number

This movie shows the output from the whole-cell migration simulations. The clutch number is increased from left to right (N_c = 50, 250, 750, 1000). Each panel displays a representative cell (white) and its trajectory (red) over 20 simulated hours. Scale bars = 50 μm.

Supplemental Movie 2. Tracked cells in a WT brain slice experiment

This movie shows an example of a brain slice experiment over 7.5 hours. Green shows eGFP-actin emission from cells. Red dots indicate the cell centroid as tracked by the Matlab algorithm.

Acknowledgments

There were many contributors to this study who need to be acknowledged for their work. The study of the effects of CD44 on mouse glioma progression was initiated by Stacy Decker Grunke, John Ohlfest, and James McCarthy. The mouse tumor model was designed and carried out by the co-first author of the study, Stacy Decker Grunke, and Karen SantaCruz. Benjamin Bangasser developed and tested the 2D motor-clutch migration simulator. Colleen Forster and Peter Canoll conducted immunohistology of tumor sections. Additionally, Peter Canoll provided his expert opinion of the different histological specimens. Matthew Price conducted the western blots to measure protein

levels in our animal models. Thomas Odde, under the guidance of David Odde and me, analyzed the human survival data from the Phillips dataset and TCGA. My specific contributions included conducting brain slice experiments on the confocal microscope and the subsequent analysis, and suggesting the hypothesis that a biphasic relationship between CD44 level and animal survival/cell migration. Steven Rosenfeld, Eva Turley, James McCarthy, John Ohlfest, and David Odde were advisors to this work and provided unspecified intellectual contributions.

The study was supported by the U.S. National Institutes of Health grants RC1-CA145044 and R01-CA172986 to DJO and SRR; University of Minnesota Institute for Engineering in Medicine Group Grant to DJO; U.S. National Institutes of Health grant R01-CA138437, American Cancer Society grant RSG-09-189-01-LIB, State of Minnesota, Minnesota Partnership for Biotechnology and Medical Genomics, and Children's Cancer Research Fund to JRO; U.S. National Institutes of Health grants T32-GM008471, T32-DA022616, and F31-NS67937 to SAD; National Science Foundation Graduate Research Fellowship 00006595 to BLB, and the Minnesota Supercomputing Institute.

Chapter2: Brownian dynamics simulation of tumor progression and application to glioblastoma

Authors: Rebecca L. Klank[1], Steven S. Rosenfeld[2], David J. Odde[1]

Affiliations

[1]Department of Biomedical Engineering, University of Minnesota, Minneapolis, Minnesota, USA

[2]Burkhardt Brain Tumor Center, Department of Cancer Biology, Cleveland Clinic, Cleveland, Ohio, USA

Summary

Tumor progression modeling offers the potential to accurately predict tumor growth and invasive behavior to improve prognostic accuracy and guide therapy development. Common simulation methods include continuous reaction-diffusion (RD) differential equation approaches that capture mean spatio-temporal spreading behavior and discrete agent-based (AB) approaches which capture individual cell events such as proliferation or cell migration. The brain cancer glioblastoma (GBM) is especially appropriate for such proliferation-migration modeling approaches because tumor cells seldom metastasize outside of the central nervous system, and cells are both highly proliferative and invasive, contributing to inevitable patient fatality over 1-2 years. In glioblastoma research, current RD approaches rely on estimates of proliferation and

migration parameters derived from computed tomography or magnetic resonance images and estimated based on Fisher's approximation, which does not provide a unique solution. Estimates of glioblastoma cell migration rates, modeled with a diffusion coefficient are approximately 1-2 orders of magnitude larger than single-cell measurements in animal models of this disease. To identify possible sources for this discrepancy, we evaluated the fundamental RD simulation assumptions that cells are point-like, i.e. are volume-less and can therefore overlap. To give cells physical size (~10 μm), we used a Brownian dynamics approach that simulates individual single-cell diffusive migration, growth, and proliferation activity via an agent-based approach where cells can be prohibited from overlapping each other. We found that, for realistic single-cell parameter growth and migration rates, a non-overlapping model gave rise to a jammed configuration in the center of the tumor and a biased outward diffusion of cells in the tumor periphery, creating a quasi-ballistic advancing tumor front. The simulations demonstrate that a fast-progressing tumor can result from minimally diffusive cells, but at a rate that is still dependent on single-cell diffusive migration rates. Thus, this small change in assumptions to a more realistic volume-conserved model can account for the apparent discrepancy between estimated and measured diffusion of GBM cells and provide a new framework for simulating the mechanisms of tumor growth that naturally links single-cell growth and migration dynamics to tumor-level progression.

Introduction

Tumor modeling is a rapidly developing approach to understanding and predicting cancer progression. A primary goal of computational tumor progression models is to develop accurate predictions about the extent and time scale of tumor spread that can assist clinicians in treating cancer patients. In addition, such models can potentially be used to inform possible surgical interventions, aid in patient stratification, and guide therapeutic strategy development. Ideally, a tumor progression model would be consistent across a wide range of spatial-temporal scales (*85*) to predict how molecular-level perturbations, resulting from either mutation or therapeutic intervention, would affect tumor-level progression and, ultimately, patient outcome.

Many computational cancer models of progression have been established, each with variable predictability and applications. The simplest model describing tumor progression is an exponential growth model. While this type of model is helpful for predicting proliferative growth of benign tumors (*86*), it lacks spatial information and therefore does not capture the invasive behavior of malignant tumors.

A common simulation method that captures spatial information is the reaction-diffusion (RD) partial differential equation approach. In some cases, this approach also includes convection to account for chemotactic gradients, i.e. reaction-diffusion-convection simulations. RD models are based on the continuity equation that describes mass transfer in a differential control volume and can be used to predict changes in

particle concentration over space and time. The RD equation describing the model is typically written as

$$\frac{\partial C}{\partial t} = D\nabla^2 C + pC \qquad \text{(Eq. 14)}$$

where C is the concentration of cells, D is the single-cell random motility coefficient (units: m^2/s) modeled as random Fickian diffusion, and p is the net proliferation rate of cells (units: 1/s). A fundamental limitation of this approach is that it allows cancer cells to achieve unrealistically high concentrations, higher than the physical dimensions of a cell would allow (~10 μm). For example, in the extreme case of non-migrating cells (D=0), the tumor will grow exponentially without limit. To address this problem, the Fisher-Kolmogorov partial differential equation has been be used to describe tumor growth, which assumes a logistic growth term with a maximum carrying capacity, *k*, to prevent proliferation when local concentration is too high to permit the creation of a new cell, so that Eq. 14 becomes

$$\frac{\partial C}{\partial t} = D\nabla^2 C + p\left(1 - \frac{C}{k}\right)C \qquad \text{(Eq. 15)}$$

More sophisticated versions can also account for the normal cells in the tissues, as well as cells that could potentially be recruited by growth factors secretion such as that described in (*15*). Importantly, the combined effects of migration (D) and proliferation (p) with a local carrying capacity (k) generate clinically-realistic tissue-level spreading behavior which is often estimated as a traveling wave with linear radial velocity (*86–88*).

The RD approach has been used in glioblastoma research to validate the therapeutic benefit of surgical resection (*89*) or chemotherapy (*90*) by comparing expected tumor growth in the absence of treatment based on repeated computed tomography (CT) images to actual patient outcomes. Similarly, Wang *et al.* showed that they could prove efficacy of various therapeutic treatments by assuming continued linear growth based on pretreatment magnetic resonance (MR) images of patients' glioblastoma tumors in the absence of treatment and comparing to patients' progression with treatment (*91*). While many have found such simulations to have predictive clinical power, the predictions are of mean cellular behavior and the physical dimensions of single cells, and their presumed inability to overlap, is not strictly enforced, potentially leading to a mechanistic disconnect between actual cell movements at the single-cell level and the tumor-level progression dynamics.

Modeling approaches that account for single-cell behaviors are termed discrete particle or agent-based (AB) simulations (*92*). These simulations track the spatio-temporal behavior of individual cells and can include complex cellular-level dynamics such as acquisition of a new mutation or a phenotypic change. There are many different approaches to discrete simulations including lattice-based cellular automata models (*87, 93*), Potts models (*94–96*), or agent-based models (*97, 98*)(reviewed in (*91*)), but all have the common attribute of simulating individual cells or cell clusters. Since simulating individual cells at realistic tumor quantities ($N \approx 10^9$)(*93*) is computationally demanding,

lattice model are often used to limit computational needs (though this is not a requirement (*99, 100*)). Lattice models, while containing spatial information on the locations of cells, seldom actively incorporate the effects of cell migration, but permit tumor spreading only through proliferation events when a new cell is placed in an adjacent vacant lattice site. While AB approahes that include migration do exist, such as the recent study by Waclaw *et al.*(*93*), the addition of a probabilistic switching to a migratory state in this simulation, where cells randomly "jump" out of the tumor mass, is empirical in nature and not based on physical first principles.

More complex yet is a hybrid continuum-discrete model that incorporates environmental variables such as tissue oxygenation or extracellular matrix concentration as a continuum (often modeled with RD equations) into the discrete cell framework (*91, 92*). Some of these models are quite comprehensive and simultaneously account for several variables that contribute to tumor progression. For instance, Alarcon *et al.* used a hybrid model to incorporate the effects of vascularization, blood flow, growth factors, and cellular interaction of normal and cancerous cells into a discrete model of tumor cell growth (*85*). Similarly, in a model of breast cancer growth, Kim *et al.* simulated four distinct cell types – cancerous epithelial cells, normal epithelial cells, fibroblasts, and myofibroblasts – at single-cell resolution and incorporated growth factor secretion and mechanical interactions between the tumor and surrounding tissues (*100*). These hybrid models, though highly complex and able to link cellular, or even subcellular, events to

gross macro-level behavior, again, often ignore fundamental single-cell migration dynamics. Discrete models that include migration typically incorporate a probabilistic switching from proliferative to migratory phenotype, enlisting the so-called "go-or-grow" dichotomy(*93*, *98*, *101*). Exception to this are the model by Hatzikirou *et al.* which includes cell motility modeled as a random walk on a lattice (*87*), and the model by Khain et al. that used a discrete model to simulate in vitro scratch wound assays of glioma (*102*).

In simulating the specific case of the brain cancer glioblastoma (GBM), which is characterized by the invasive migration of primary tumor cells, it is especially important to consider the effects of cell migration on progression. This highly invasive grade IV glioma is infamous for its poor survival outcomes, which are often attributed to the presence of invasive cells which almost invariably lead to tumor recurrence and patient death (*88*, *89*, *103*, *104*). Even with current standard treatment of combined surgery, radiotherapy, and chemotherapy, invasive cells, which are not eliminated by these interventions, typically cause tumor to recurrence in 3-6 months (*25*, *105*) and patients with GBM only survive approximately 15 months post-diagnosis (*36*). Previous studies that have focused on simulating GBM progression recognize the importance of modeling its invasive capabilities. As such, nearly all models of GBM have used an RD framework to capture invasive cell behavior (*91*, *103*, *106*, *24*, *107*). An exception is the agent-based

model by Zhang et al. that simulates go-or-grow dynamics by including sub-cellular EGFR signaling in their lattice model (*98*).

Input parameters to the RD modeling approach such as proliferation rate and diffusion coefficient for GBM are commonly estimated from sequential CT or MRI using Fisher's approximation (*104*, *107*). While there are limitations with MRI as it is an indirect measure of glioma cell location, discussed previously (*108*), it is currently the best tool available for tracking *in vivo* tumor progression in humans, and it is not unreasonable to use these parameter estimates in order to simulate human GBM (Table 1). However, more recently, *ex vivo* measurements of single-cell migration in animal models have emerged. While not realistic for an unlabeled human specimen, in animal models of GBM, measurements of cell migration behavior have been made by tracking the trajectories of single fluorescently tagged cells within *ex vivo* brain slice cultures and calculating cell diffusion coefficient (also known as random motility coefficient) based on the mean squared displacement (MSD) of cell trajectories over time using the equation

$$D = \frac{MSD(t)}{2dt} \qquad \text{(Eq. 16)}$$

where D is the diffusion coefficient, t is the time interval, and d is dimensionality of migration. A summary of experimentally measured values for D in the literature is given in Table 2.

Table 2. Bulk tumor-based theoretical estimates of diffusivity in the literature using and RD model

Study	Modeling	D (μm^2/hr)
Anderson *et al.* 2005 (*92*)	Generic tumor growth	3600
Jiang *et al.* 2014 (*109*)	Generic tumor growth	3.6
Tracqui *et al.* 1995 (*90*)	Anaplastic Astrocytoma (estimated by CT)	43000-46000
Woodward *et al.* 1996 (*89*)	Glioma	1800-5400
Burgess *et al.* 1997 (*110*)	High grade glioma (Fisher's approximation)	5400
	Low grade glioma (HG Glioma/10)	540
Swanson *et al.* 2000 (*103*)	GBM gray matter (Fisher's approximation)	5400
	GBM white matter (Fisher's approximation)	27000
Rockne *et al.* 2009 (*107*)	Glioma (estimated by MRI)	600-37000
Wang *et al.*, 2009 (*91*)	GBM (preoperative MRI)	34-36000
Eikenberry *et al.* 2009 (*106*)	Motile "go" GBM cells (estimated by Stokes-Einstein relationship)	4200
	Non-motile "grow" GBM cells	80
Harko *et al.* 2015 (*86*)	GBM (estimated from Rockne 2015)	16300
Hathout *et al.* 2015 (*111*)	GBM	16700

Table 3. Single cell experimental measurements of diffusivity in the literature

Study	Condition	D ($\mu m^2/hr$)	Species	Transforming method	Driver
Chapter 1	WT Glioma	9.1	Mouse	PEI/Sleeping Beauty	SV40LgT/NRAS
	CD44 -/- Glioma	2.1	Mouse	PEI/Sleeping Beauty	SV40LgT/NRAS
Ivkovic *et al.* 2012 (*112*)	Gliomatosis Cerebri	100	Rat	Retrovirus	EGFR
	Gliomatosis Cerebri + blebbistatin	10	Rat	Retrovirus	EGFR
	GBM	80	Rat	Retrovirus	PDGF/EGFR
	GBM + blebbistatin	10	Rat	Retrovirus	PDGF/EGFR
Massey *et al.* 2012 (*15*)	Glioma	250	Rat	Retrovirus	PDGF

Curiously, the values of glioma migration estimates based on tissue-level observations (Table 2) are 1-2 orders of magnitude larger than those measured at the single-cell level in animal models (Table 3). Such high diffusion values used in the RD models have been noted in the literature before. In their simulation of *in vitro* glioma growth, Stein *et al.* found that more realistic input values for diffusion (40-800 $\mu m^2/hr$) resulted in poor agreement between *in vitro* growth and simulated growth, and, in fact mention that their model outputs would be in much better agreement with *in vitro* data if they increased diffusion coefficient by a factor of 10 (*113*). Moreover, it has been noted that estimating D and p parameters based on Fisher's approximation does not make

independent predictions, but accuracy of the approximation improves as the number of cells is increased (*114*) (i.e. estimates made from bulk behavior). Therefore, we sought to understand the cause of this quantitative difference between parameter estimation and direct measurements by examining the assumptions of the RD approach. Particularly, we considered the unrealistic assumptions of the basic mass continuity equation which assumes particles are present at low concentrations and occupy no volume.

To evaluate these assumptions, we used a Brownian dynamics (BD) approach to simulate migration, growth, and division of individual cells in three different models:

1. An overlapping model in which individual cells can diffuse such that they occupy the same space and proliferate independently of the presence of nearby cells, consistent with many RD models (*89–91*, *103*, *104*, *107*, *110*)
2. An overlapping model with a carrying capacity in which cells can diffuse freely, but proliferation is inhibited if cells are currently overlapping, consistent with other RD models (*15, 91, 106, 115*)
3. A physically realistic non-overlapping model in which cells may neither diffuse nor grow if the movement would result in an overlapping configuration.

Using these models, we explore how the inclusion of non-zero cell volume affects overall tumor progression rates, such that relatively low diffusion coefficients enable realistic tumor margin velocities, and are sufficient to explain the discrepancies between estimated and measured cell diffusivity.

Results

Comparison of glioma migration parameter estimates (Table 2) to animal model measurements (Table 3) showed a common disconnect between single-cell behavior and RD model inputs describing tumor cell motility. Therefore, our first task was to determine if the estimates for cellular diffusion used in RD models were physically realistic. Among the fastest cells are human embryonic mesenchymal stem cells, which have been shown to move along a one dimensional adhesive track at a velocity of 5.2 μm/min and are highly directional in their movements (*116*). From this velocity and the relevant time scale of typical directional switching (~10 min), the theoretical diffusion coefficient of the "world's fastest cell" is approximately 8000 μm^2/hour. Moreover, the cell speed of 5.2 μm/min is approximately equivalent to the speed of actin protrusion, 0.21 μm/s (or 12.6 μm/min)(*117*), and therefore is likely to be close to the upper limit for possible cell motility. Consequentially, any values of diffusion beyond this theoretical upper limit likely do not represent actual single cell behavior. While some of the estimates in Table 2 are within this limit, many are extremely large, particularly for estimates of white matter migration. Therefore, we were motivated to re-examine the assumptions of the RD models of tumor spread used to estimate such large values.

One issue we noticed with adapting an RD framework to estimate tumor progression parameters was the assumption that the diffusing species (i.e. the cells) is present at low concentrations (i.e is dilute). However, characteristically, tumor cell

density is very high, with cells closely packed together. Therefore, we derived the diffusion coefficient for a concentrated tumor environment. At high concentrations, we assume that the gradient in chemical potential dominates cell movements instead of the concentration gradient. Estimating the chemical activity as increasing linearly with mole fraction with an intercept at 1 (to be consistent with the low concentration equation), we found that the apparent diffusion due to high cell concentration is at most 2-times the intrinsic diffusion (see full derivation in the supplement).

Since the estimates of diffusion were much greater than 2x the measured single-cell values and could not be explained by the low concentration assumption, we next examined the RD assumption that cells can be treated as points in space which occupy no volume. Because an analytical solution to this problem does not yet exist, we used a BD approach to simulate how single-cell behavior relates to overall tumor progression with and without this assumption. Since the BD approach simulates single stochastic events, this method is computationally intensive compared to the RD approaches, and therefore one-dimensional (1D) simulations were used to evaluate the validity of the assumptions and explore basic dynamic features and regimes. Moreover, the primary issue with current diffusion estimates is the use of uncharacteristically large values for migration along white matter tracts, which can be described as a quasi-1D route of migration (*118*), therefore we believe the 1D model is potentially relevant to glioma disease progression along these routes.

Starting from a single initiating cell, *in silico* linear tumors were grown for a duration of time long enough to generate the sixth generation of cells (assuming perfect, unimpeded division) and positions of all of the cells in the simulation were tracked over time. In total, three 1D models were analyzed. The first model, termed the simple overlapping model, captures the assumptions of the mass conservation equation (Eq. 14) by allowing cells to migrate, grow, and divide without regard for the presence of neighboring cells occupying the same space. The second model, the overlapping model with carrying capacity (overlapping+cc), exhibits similar rules to the simple overlapping model in terms of diffusive and growth behavior, but does not allow cell division to occur if the cell in question is overlapping with another cell (Figure 7A), thus enlisting a local carrying capacity akin to the logistic growth term in Fisher-Kolmogorov equation (Eq. 15). The final model, the non-overlapping model, does not allow cells to occupy the same space at any point in their life-cycle, as diagramed in Figure 7B. A full description of these models is available in the methods section.

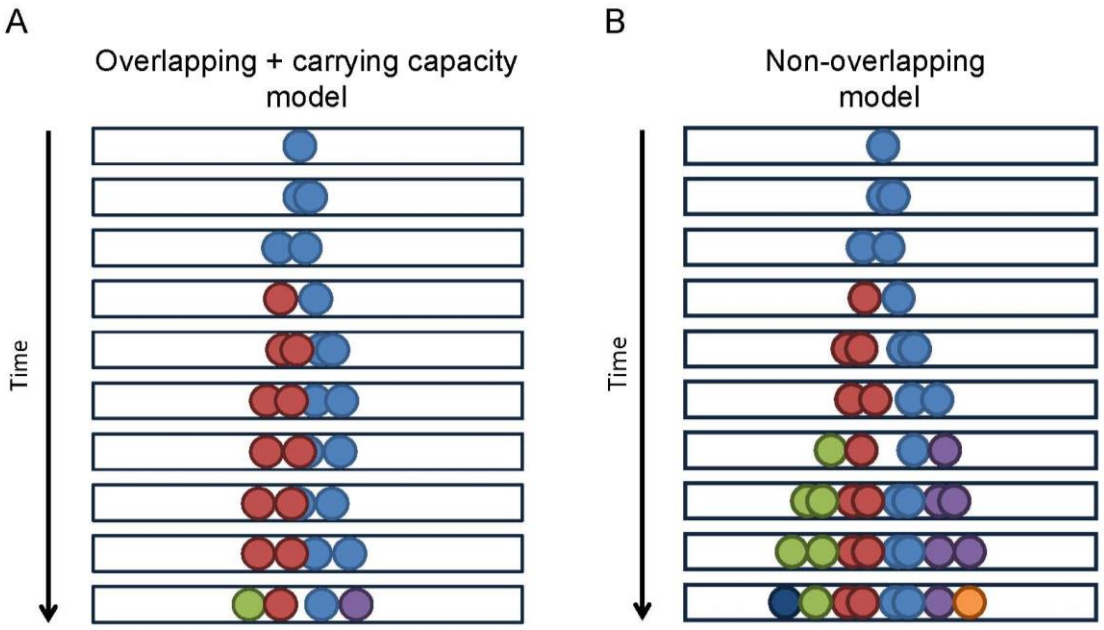


Figure 7. Diagram of overlapping + carrying capacity and non-overlapping models

A. Overlapping + carrying capacity model diagram. Black boxes indicate the one-dimensional line in which cells were permitted to occupy. The initial cell, colored light blue, is able to diffuse and grow, indicated by the two spheres of the same color moving gradually apart. As time advances, the cell undergoes division, indicated by the arrival of a new cell color. When multiple cells are present (each distinct cell is represented by two spheres of the same color), cells are unable to divide if they are currently overlapping. This is how an effective local carrying capacity is carried out in simulations. B. Non-overlapping model diagram. Similar to A, the same rules apply, except that individual cells are unable to overlap each other and occupy the same space at any time. Not diagramed, the simple overlapping model permits cell movements and cell division events without regard for the presence of other cells.

By directly comparing the outcomes of the three models, it is evident that the conservation of volume used in the non-overlapping model enhances tumor spreading, as shown in Figure 8A (see also Supplemental Figures 5-7 and Supplemental Movies 3-5).

Moreover, proliferation is localized to the outer edge in the non-overlapping simulations (indicated by red dots) which is consistent with histological observations (*87*, *93*). The distribution of final cell positions in each simulation exhibits a Gaussian (normal) distribution in the overlapping model, consistent with a diffusive model, whereas the final cell positions in the non-overlapping model exhibit an approximate uniform distribution indicating non-diffusive behavior (Figure 8B). Moreover, the curvature in the cell MSD plots in Figure 8C is close to linear in the overlapping models (ignoring curvature at long times where error is large) indicating diffusive cell movements, while the non-overlapping model MSD plot has upward curvature indicating superdiffusive behavior. A plot of tumor radius over time shows that the overlapping model advance linearly with time. The non-overlapping model also advances linearly with time until it reaches a jammed configuration. After jamming, speed of radial growth is increased (slope is greater) (Figure 8D), consistent with previous AB modeling approaches of glioma migration and proliferation (*98*). Together these data indicate that the non-overlapping model enhances tumor spreading compared to the overlapping models with the same input parameters.

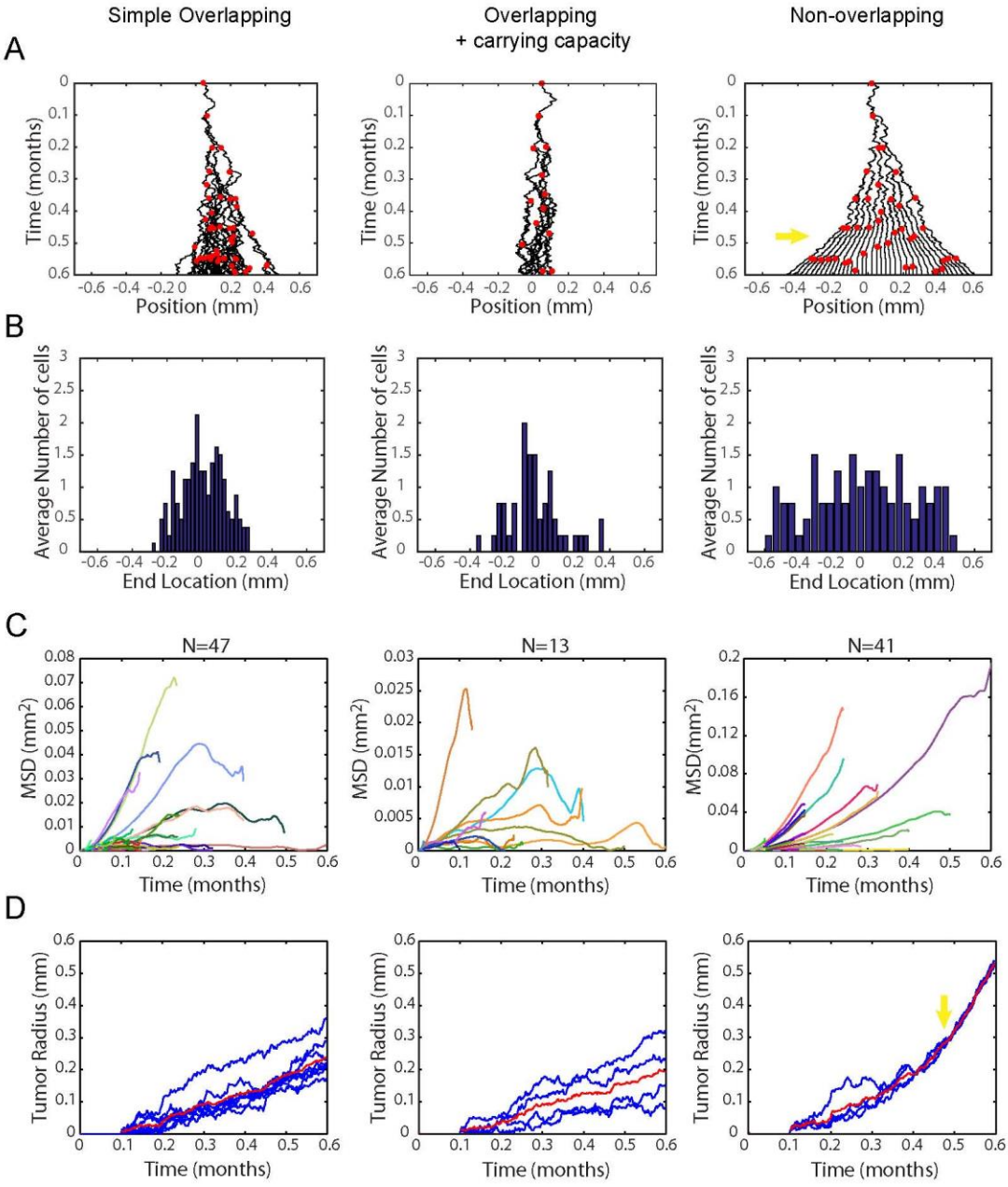

Simple Overlapping
Overlapping
+ carrying capacity
Non-overlapping
A
Time (months)
Position (mm)
B
Average Number of cells
End Location (mm)
C
N=47
N=13
N=41
MSD (mm²)
MSD(mm²)
Time (months)
D
Tumor Radius (mm)
Time (months)

Figure 8. Spatio-temporal dynamics of the three BD models

A. Progeny trees of the three models using the same input parameters where $D_0 = 20\ \mu m^2$/hour and $p_0 = 10$ divisions/month. Black curves indicate the trajectories of cells. Red dots indicate the time and position of a division event. On this time scale, unimpeded cell proliferation would generate exactly 6 generations of cells. Yellow arrow on the non-overlapping model indicates where jamming begins. B. Distributions of cell locations at the end of a simulation (t = 0.6 months). The Gaussian distributions in the overlapping models are characteristic of diffusion, while the uniform distribution of the non-overlapping models indicates non-diffusive behavior. C. Plots of mean squared displacement (MSD) over time for individual cells within each simulation. Linear behavior indicates a purely diffusive process, asymptotic behavior (i.e. concave dowm) indicates a subdiffusive process, and upward concavity (as in the non-overlapping model) indicates a superdiffusive process. D. Tumor radius over time for multiple simulations evaluated at 20 μm^2/hour and 10 divisions/month in each model show the overlapping models progress linearly with time. The non-overlapping model progresses linearly, similar to the overlapping models, then the slope of the linear progression increases after the simulation reaches a jammed configuration, as indicated by the yellow arrow.

In each of the three models, the input parameters affect the progression characteristics differently. As shown in Figure 9A, in both of the cell-cell interaction models (overlapping+cc and non-overlapping) the apparent proliferation rate, p_{app}, normalized to nominal proliferation, p_0, is decreased as p_0 is increased. As expected, the simple overlapping model shows no correlation with increased proliferation. In the cell-cell interaction models, this effect is essentially a negative feedback on proliferation. Interestingly, the slope of normalized proliferation rate in the non-overlapping model is similar between different diffusion coefficient values, but this is not the case for the overlapping+cc model, especially at the lowest diffusivity values. This is likely because the small diffusivity values do not allow cells to move away from each other to generate

sufficient space for a division event on the timescale of proliferation in the overlapping+cc model, while the non-overlapping model generates space to allow division because movements which result in overlapping configurations are rejected.

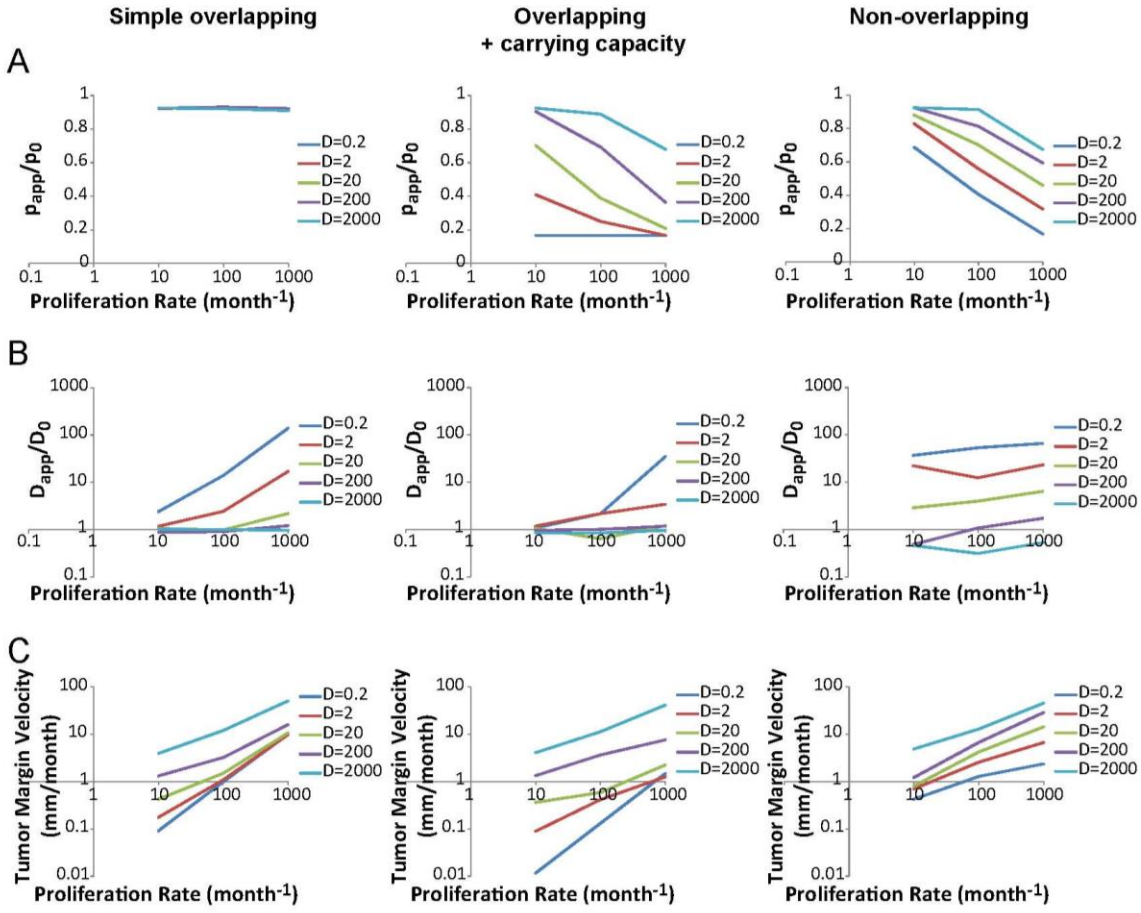


Figure 9. Analysis of overall model behavior and parameter sensitivity

A. Effect of proliferation and diffusion input values on normalized apparent proliferation (p_{app}/p_0) in the three models. Colors indicate the input value for diffusion coefficient (D_0) in μm^2/hour. The negative trend in the models that account for cell-cell interactions (middle and right columns) shows that proliferation rate has a negative feedback on itself. B. Effect of proliferation and diffusion input values on normalized

apparent diffusion coefficient (D_{app}/D_0) in the three models. Colors indicate the input value for diffusion coefficient (D_0) in μm^2/hour. The overlapping models (left and middle columns) show some dependence on D_0 and p_0, particularly at low diffusion coefficient and high proliferation rate. This is due to the division of cells which generates an artifactual apparent movement, which appears as a significant boost in motility when the nominal diffusion coefficient is small. The non-overlapping model shows little dependence of apparent diffusion on proliferation, but does have a true boost in apparent diffusion due to jamming effects, particularly at low D_0. C. Tumor velocity as a function of both input parameters. In most cases, the overlapping+cc model progresses much slower than the non-overlapping model.

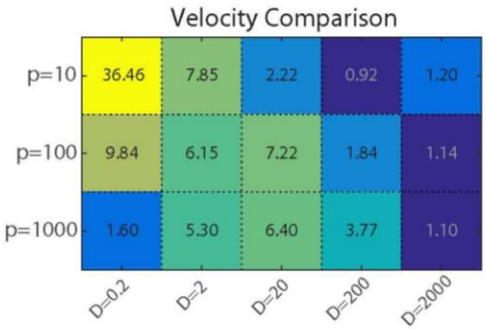


Figure 10. Comparison of velocity values between the overlapping+cc and non-overlapping models
Velocities of the non-overlapping model and overlapping+cc model with the same set of parameters are compared. Values in the heatmap are calculated as the quotient of the average velocity of non-overlapping and the average velocity of overlapping+cc.

When comparing the effects on apparent diffusion, as expected, the overlapping models allow cells to diffuse freely, so the apparent diffusion, D_{app}, is equal to the nominal diffusion coefficient D_0 at large values of D_0. (Figure 9B). For small values of D_0, apparent diffusion is increased as p_0 increases because the generation of a new cell artifactually moves the cell centroid the distance of one cell radius (10 μm), which is a significant movement in this regime. What we see in the non-overlapping model is that

central cells at high concentrations are unable to move because they are prevented from overlapping. This results in a jammed configuration where interior cells nearly stop diffusing. At the same time, exterior cells have freedom to move and are biased toward outward movements, making apparent diffusion larger for those cells. In this model, normalized D is not dependent on p_0, but is strongly dependent on D_0. At low D_0, especially, the prevention of cell overlap can generate apparent diffusivity that reaches values up to 65 times that of D_0, showing that the effect of jamming-induced ballistic behavior can largely account for the discrepancy between single cell measurements and estimates of diffusion used in literature.

Analysis of the tumor margin velocity of the three models shows a dependence of velocity on both proliferation and diffusion (Figure 9C). However, the overlapping models show a reduced tumor margin velocity, especially at lower diffusion coefficients. Comparisons of the plots for the different models shows how a smaller input diffusion coefficient in the non-overlapping models can generate a tumor velocity similar to the overlapping+cc model at larger input diffusion (i.e. compare D_0=20 μm^2/hr and p_0=100/month of the non-overlapping model to D_0=200 μm^2/hr and p_0=100/month of the overlapping+cc model).

Using an agent-based approach allows motion statistics to be computed for individual cells as a function of position in the tumor, relative to the tumor edge, which makes testable predicitons about the spatial dependence of cell trajectories. As shown in

Figure 11A, the overlapping models show no pattern in instantaneous velocity, while the non-overlapping model shows a distinct quasi-ballistic front with a zero velocity core (Figure 11A). Similarly, single cells in the overlapping models all have normalized instantaneous diffusion measurements centered around the value 1 that do not vary with spatial position. In the non-overlapping model, cells on the ballistic front display superdiffusive behavior (with values over 1), while cells at the core display subdiffusive behavior (Figure 11B). The length of the convective front in the non-overlapping model varied with input parameters, as shown in Figure 11C. Ballistic front length was largest for low values of p_0 and high values of D_0, because as p_0 increases, more cells become jammed. Likewise, the time to reach a jammed configuration depends on the input parameters, as shown in Figure 11D. Jamming occurs more quickly with high proliferation inputs, and depends only slightly on diffusion inputs. In the p_0=10/month situation, some simulations did not reach a jammed configuration in the time allotted for simulation, but would likely jam given sufficient time.

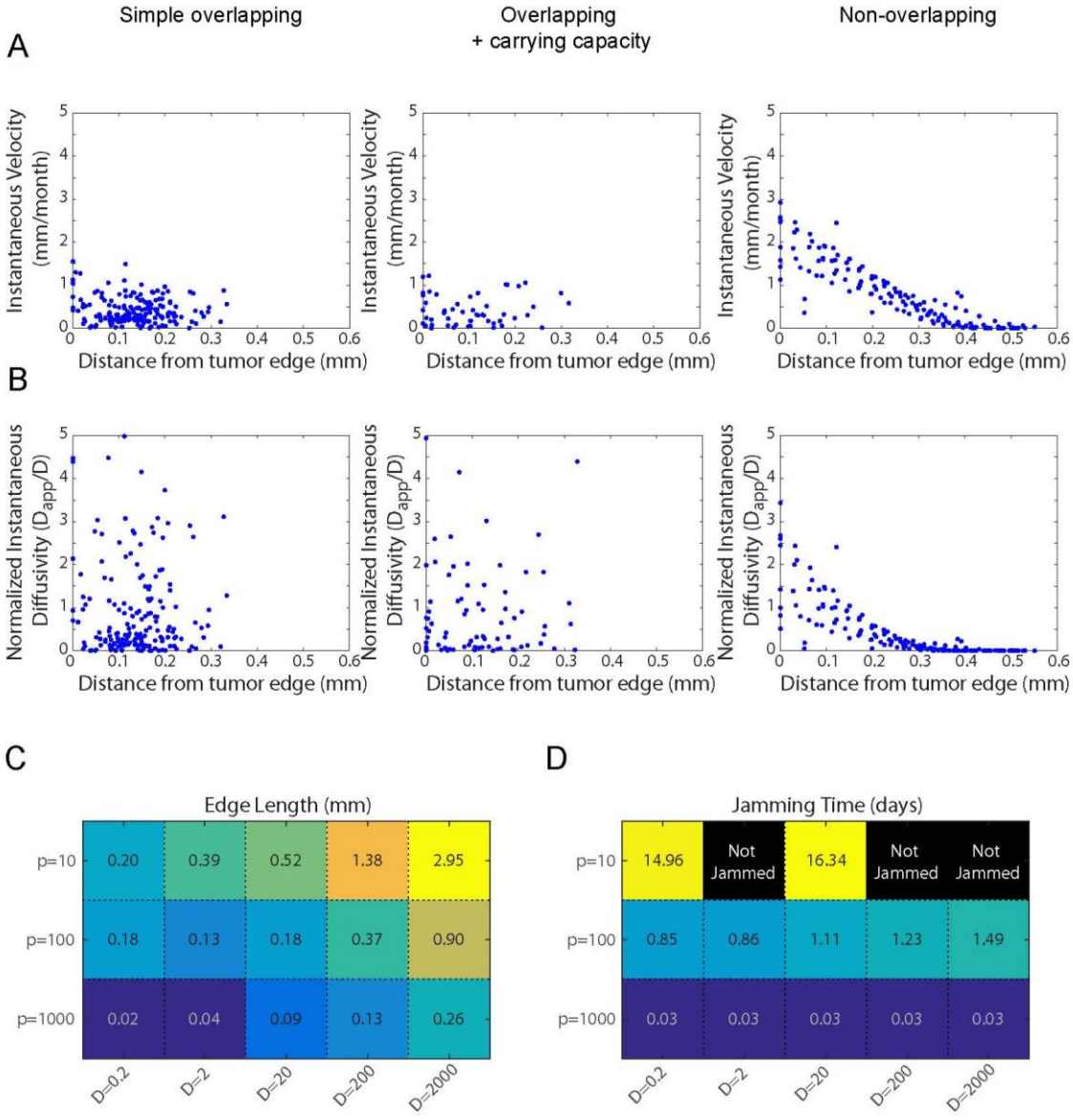


Figure 11. Intratumoral spatial variation of single cell motion behavior

A. Instantaneous velocities of individual cells within 4 simulations for each model evaluated after the sixth generation of cell division (total time = 432 hours) with D_0=20 μm^2/hour and p_0=10 divisions/month. Single-cell velocities are plotted against location in the tumor, starting from the most distal cells. B.

Instantaneous diffusivity of individual cells within 4 simulations for each model evaluated under the same conditions as A. Diffusion coefficients are plotted against location in the tumor, starting from the most distal cells. C. The length of the ballistic front in the non-overlapping model simulations are displayed for all parameter combinations in the heatmap. D. Jamming times for all parameter combinations are displayed in the heatmap. Black boxes indicate cases where jamming was not achieved within the allotted time (time to the sixth generation of cells if growth is unimpeded).

Overall, we find that the non-overlapping model requires much smaller diffusion coefficient values to reach the same rate of tumor spreading as the overlapping models. Thus, this suggests that the discrepancy of diffusion coefficient values between bulk tumor estimates (Table 2) and actual single-cell measurements (Table 3) could possibly be explained by the fact that the current RD approahces do not account for the effect of incompressible tumor cell volume on spreading.

Discussion

Several models have been developed to describe GBM tumor spreading. However, many of these models, particularly those based on RD approaches, require unrealistically large values of individual cellular diffusion coefficients which are estimated from apparent spreading based on CT or MRI (compare Table 2 and Table 3). In fact, single-cell values for animal glioma migration modeled as a random diffusive process are 1-2 orders of magnitude smaller than estimates used in many of these simulations. To understand the source of this discrepancy in diffusion coefficients

between estimates and measurements, we evaluated the assumptions of the RD approach. We first tested the RD assumption that particles are present at low concentrations and showed that a high-concentration correction based on thermodynamic activity coefficients only resulted in a 2-fold larger apparent diffusion coefficient. Therefore, we also evaluated the assumption that cells can be modeled as volume-less points in space. Using a 1D BD modeling approach that accounts for single-cell stochastic movements, we found that the inclusion of volume has a significant impact on overall tumor spreading behavior. Specifically, not allowing cells to occupy the same space led to the development of a jammed configuration that generated a non-motile tumor core and a ballistic tumor periphery. Additionally, the non-overlapping correction allows realistic diffusion input parameters to generate faster tumor spreading velocities, and thus we hypothesize that cell jamming contributes to overall tumor spreading *in vivo*.

What is interesting about the non-overlapping model is that it generates increased spreading behavior without the inclusion of chemotaxis terms (*109*) or the "go-or-grow" dichotomy (*106*), in which cells are either proliferative or migratory. All cells in the non-overlapping model have the capacity to both proliferate and invade, and a ballistic tumor front develops naturally from physical volume-exclusion principles. The advantage of this approach is that it is unnecessary to determine a third parameter which is the probabilistic switching between the proliferative and migratory phenotypes which has frequently been added empirically to generate realistic tumor growth (*93*).

In future studies, we plan to test the prediciton that jamming has a significant impact on tumor spreading by comparing animal model behavior to computational model predictions. If jamming does significantly contribute to GBM spreading, the non-overlapping model predicts that there would be a difference in motility between cells at the tumor edges and in the tumor core, which can be tested by measuring single-cell trajectories as well as cell location relative to the central tumor. Furthermore, the non-overlapping model predicts that the migration of cells at the edges would be biased in the outward radial direction. These predictions can potentially be assessed in *ex vivo* animal systems such as those used to measure single-cell diffusion coefficients (*31*) (also see Chapter 1).

One potential limitation of the current model is that we characterize the effects of conserved volume in 1D, while real tumors progress in 3D. When exploring the behaviors of our model in 3D, we expect that jamming will contribute to tumor progression in that it will bias migration radially outward; however, since cells are able to move in multiple directions to escape jamming, it will become more difficult to achieve a jammed configuration. Interestingly, our 1D model presumably has application to GBM tumor dissemination in that GBM cells have been observed to migrate preferentially along linear tracts within the brain such as perivascular spaces(*119*) or white matter tracts (*8*, *27*, *28*, *31*, *118*).. Therefore this could be a method of tumor spreading in GBM that

contributes to its extensive invasiveness, and the current model could describe the migration at this meso-scale.

In terms of simulating GBM tumors, we also recognize that our model currently lacks the characteristic isolated invasive cells that blur the tumor margin (*4*, *113*). We anticipate future models could rectify this issue by including a mixture of low and high diffusion coefficients, similar to what had been done previously to simulate differences in migration between white matter and gray matter (*103*). Additionally, our model lacks the characteristic necrotic regions that are typical of GBM tumors. This could easily be incorporated into the model by the addition of oxygen consumption where more dense regions have to compete for resources and thus are more likely to become necrotic (*92*). Furthermore, other models have demonstrated how the inclusion of oxygen consumption generates *in silico* tumors with more invasive, finger-like morphology (*92*, *25*), frequently observed for GBM tumors.

In comparison to RD modeling approaches to GBM, the non-overlapping model presented here has the advantage that it can simulate single-cell events. This is important because predictive models of tumor behavior should be physically and biologically consistent across multiple spatio-temporal scales, and RD approaches have been previously shown to fail to capture cell-cell interaction and clustering effects, particularly when proliferation rate is significant compared to diffusion coefficient (*97*, *120*), which is precisely when we would expect jamming to occur.

Although the BD approach used here is based on single cell events, it is perhaps not the optimal approach for generating patient predictions. A major disadvantage of BD approach is that it is computationally expensive, which may explain why the field has previously relied largely on RD approaches. For future studies we propose implementation of the concepts discovered here into more high-throughput computational models, perhaps by the inclusion of a convective term to an RD framework to simulate jamming effects, or by the addition of a concentration-dependent diffusion coefficient such as that used by Harko *et al.* (*86*). However, the model by Harko *et al.* used a concentration-dependent D such that, as cellular concentration approaches c_{max}(a carrying capacity), D approaches infinity. In our model we see the opposite - as local concentration reaches its maximum, cells become jammed and can no longer move. Therefore, rather than applying the change empirically, it may be possible to derive an analytical solution (*60*) to the jamming dynamics observed in our simulations so that the model retains its relevance at both the cell and tissue scales. In addition, our simulations reveal an early onset of the jammed-ballistic behavior, which can be regarded essentially as a steady state that will remain locked-in indefinitely, making longer simulations unnecessary (Supplemental Figure 8).

Conclusions

Many different approaches have been taken to the problem of predicting tumor growth outcomes. One common method has been to adapt the RD framework, used to

solve molecular transport problems, to the realm of cancer cell proliferation and migration via a diffusion-like process. However, since *in vivo* single-cell measurements of diffusion and proliferation do not yet exist for the human disease, the parameters for this type of model have been estimated using bulk tumor characteristics of tumor progression from *in vivo* imaging modalities. Within animal models of the disease, however, direct measurements of single cell diffusivity are now available. Despite the different species, we would expect that cell-level measurements between animals and humans would be similar, as cell migration generally employs conserved acto-myosin-adhesion machinery (*121*), yet there is a large discrepancy between measurements and estimation. Therefore, we re-evaluated the underlying assumptions of the RD framework and found a possible origin of this discrepancy, namely that the RD framework assumes individual particles do not occupy volume and therefore are permitted to overlap, while real cells clearly do occupy volume, and therefore overlapping configurations are non-realistic.

In implementing this adjustment to the RD framework, we found that non-overlapping cells rapidly enter a jammed regime which produces slow-growing interior cells with subdiffusive migration characteristics and fast-growing peripheral cells with superdiffusive, ballistic, migration characteristics. The increased tumor spread allows cells with nominally low diffusive characteristics to generate a fast progressing tumor that still depends strongly on diffusive cell movement. Thus the non-overlapping model

helps to explain the apparent discrepancy between measured and estimated diffusion and provides a realistic *in silico* tumor based on measurable parameters. In summary, we suggest that non-overlapping behavior should be implemented in future models of tumor progression to generate physically realistic and quantitatively accurate models that are predictive at multiple spatio-temporal scales.

Methods

One-Dimensional Brownian Dynamics Simulation

Three distinct Brownian dynamics models were created using a variable time step Monte-Carlo framework, which we refer to as the simple overlapping model, the overlapping plus carrying capacity model, and the non-overlapping model. Simulations were seeded with one initial "cell" modeled as two overlapping, eclipsed spheres. Time (t) was incremented based on the nominal time increment (τ= 1 min) divided by the current number of cells, $N_{cells}(t)$ as shown in equation 3. This was to account for the fact that the probability of selecting any one particular cell is decreased proportionally by the current number of cells.

$$t(i+1) = t(i) + \frac{\tau}{N_{cells}(i)} \qquad \text{(Eq. 17)}$$

At every iteration (i) of the simulation, a *uniformly* distributed random number was used to select a single cell j for movement.

$$j-1 < rand_1 * N_{cells}(i) < j \qquad \text{(Eq. 18)}$$

A second *normally* distributed random number was generated to determine the magnitude and direction of the movement of cell *j* based on the nominal diffusion coefficient, *D*, where $x_j(i)$ is the cell centroid position.

$$x_j(i+1) = x_j(i) + rand_2 * \sqrt{2dD\tau} \qquad \text{(Eq. 19)}$$

In the non-overlapping model, movements would be rejected if the location of the centroid at time $t(i+1)$ caused the cell volume to overlap with any neighboring cell volumes. The simulation would then attempt up to 100 times to find a suitable location to move, and if it failed to do so, would return to its previous position at time *t*. In the overlapping model, however, this overlap check was not performed and cells were permitted to occupy the same space when diffusing.

Another uniformly distributed random number was used to select a cell for growth using Equation 5 (above). To simulate cell growth, the two overlapping spheres, defining the volume of a single cell, gradually pushed apart such that total cell volume increased linearly and volume was exactly doubled at time $1/p_0$, where p_0 is the nominal proliferation rate. In the non-overlapping model, cells were only able to grow if the growth step did not result in overlap with any neighboring cells. In both the non-overlapping and overlapping+cc models, after the cell reached a volume of $2(4/3)\pi r^3$, where r = 10 μm, a check was performed to ensure the resulting daughter cell would not

overlap any neighboring cells, and if the check was passed, the cell would divide and form 2 individual sets of completely overlapping spheres. This was essentially enlisting a local carrying capacity limit where the probability of cell division decreases with increased local cell density.

Data Analysis

A minimum of 4 simulations per input parameter set was used for data analysis. Simulations were carried out to the number of iterations necessary for the initial cell to generate its sixth division event based on the nominal proliferation rate. Apparent proliferation rate was calculated based on the actual number of division events that occurred within the given time. Apparent diffusion was calculated based on the mean squared displacement of individual cells over time, using equation 3, and MSD was calculated using the overlapping method (*65*). Tumor velocities were calculated based on the distance of the most distal cells to the tumor center over time. Single-cell velocities were calculated as $(x(t)-x(t-\Delta t))/\Delta t$. The Δt value used to calculate instantaneous velocities was 8.3 hours, consistent with typical brain slice imaging experiments. (See Supplemental Figure 9 for the effects of varying Δt.) Similarly, instantaneous diffusion coefficients were calculated as $(x(t)-x(t-\Delta t))^2/(2*\Delta t)$. The length of the ballistic front in the non-overlapping model was estimated based on the linear fitting of the instantaneous velocity plots. Values less than 10% of the maximum were eliminated from the fit to remove jammed cells and the ballistic front lengths were taken as the x-intercept of the

fit. Time to jamming was estimated based on when the center cell reached an instantaneous velocity of 0 in the non-overlapping model.

Supplemental Materials

Derivation of diffusion within a concentrated solution

For a dilute solution, flux of a solute in the x-direction is proportional to the diffusion coefficient and the gradient of concentration, as defined by Fick's law

$$J_{i,x} = -D_{ij}\frac{dC_i}{dx} \quad \text{(Eq. S1)}$$

In a concentrated solution, solute-solute interactions can no longer be ignored, therefore the driving force for flux is not concentration gradient, but the gradient in chemical potential.

$$J_{i,x} = -\frac{D_{ij}C_i}{RT}\frac{d\mu_i}{dx} \quad \text{(Eq. S2)}$$

where μ_i is the chemical potential described by

$$\mu_i = \mu_i^0 + RT\ln(\gamma_i\chi_i) \quad \text{(Eq. S3)}$$

where μ_i^0 is the reference chemical potential, γ_i is the activity coefficient for the solute-solute (cell-cell) interactions, and χ_i is the mole fraction of the cells (*122*). μ_i^0 can be arbitrarily set to zero and *RT* is constant, therefore

$$J_{i,x} = -\frac{D_{ij}C_i}{RT}RT\left(\frac{d\ln(\gamma_i\chi_i)}{dx}\right) = -D_{ij}C_i\left(\frac{d\ln(\gamma_i\chi_i)}{dx}\right) \quad \text{(Eq. S4)}$$

By the chain rule

$$J_{i,x} = -\frac{D_{ij}C_i}{\gamma_i \chi_i}\left(\frac{d(\gamma_i \chi_i)}{dx}\right) \quad \text{(Eq. S5)}$$

By the product rule

$$J_{i,x} = -\frac{D_{ij}C_i}{\gamma_i \chi_i}\left(\chi_i \frac{d\gamma_i}{dx} + \gamma_i \frac{d\chi_i}{dx}\right) \quad \text{(Eq. S6)}$$

Assuming γ_i is a function of χi

$$\frac{d\gamma_i}{dx} = \frac{d\gamma_i}{d\chi_i}\frac{d\chi_i}{dx} \quad \text{(Eq. S7)}$$

Therefore

$$J_{i,x} = -\frac{D_{ij}C_i}{\gamma_i \chi_i}\left(\chi_i \frac{d\gamma_i}{d\chi_i} + \gamma_i\right)\frac{d\chi_i}{dx} \quad \text{(Eq. S8)}$$

$$C_i = \chi_i C \quad \text{(Eq. S9)}$$

where C is the total concentration of all elements, then

$$J_{i,x} = -D_{ij}C\left(\frac{\chi_i}{\gamma_i}\frac{d\gamma_i}{d\chi_i} + 1\right)\frac{d\chi_i}{dx} \quad \text{(Eq. S10)}$$

Using the fact that the derivative of the natural log of y is 1/y we can simplify the problem to

$$J_{i,x} = -D_{ij}C\left(\frac{d\ln\gamma_i}{d\ln\chi_i} + 1\right)\frac{d\chi_i}{dx} \quad \text{(Eq. S11)}$$

$$C\frac{d\chi_i}{dx} = \frac{dC_i}{dx} \quad \text{(Eq. S12)}$$

$$J_{i,x} = -D_{ij}\frac{dC_i}{dx}\left(\frac{d\ln\gamma_i}{d\ln\chi_i}+1\right) \quad \text{(Eq. S13)}$$

$$D_{app} = -\frac{J_{i,x}}{\left(\frac{dC_i}{dx}\right)} = D_{ij}\left(\frac{d\ln\gamma_i}{d\ln\chi}+1\right) \quad \text{(Eq. S14)}$$

Assuming a simple linear relationship between γ_i and χ_i where the intercept is 1 (dilute solutions are assumed to be ideal with an activity coefficient of 1)

$$\gamma = \beta\chi_i + 1 \quad \text{(Eq. S15)}$$

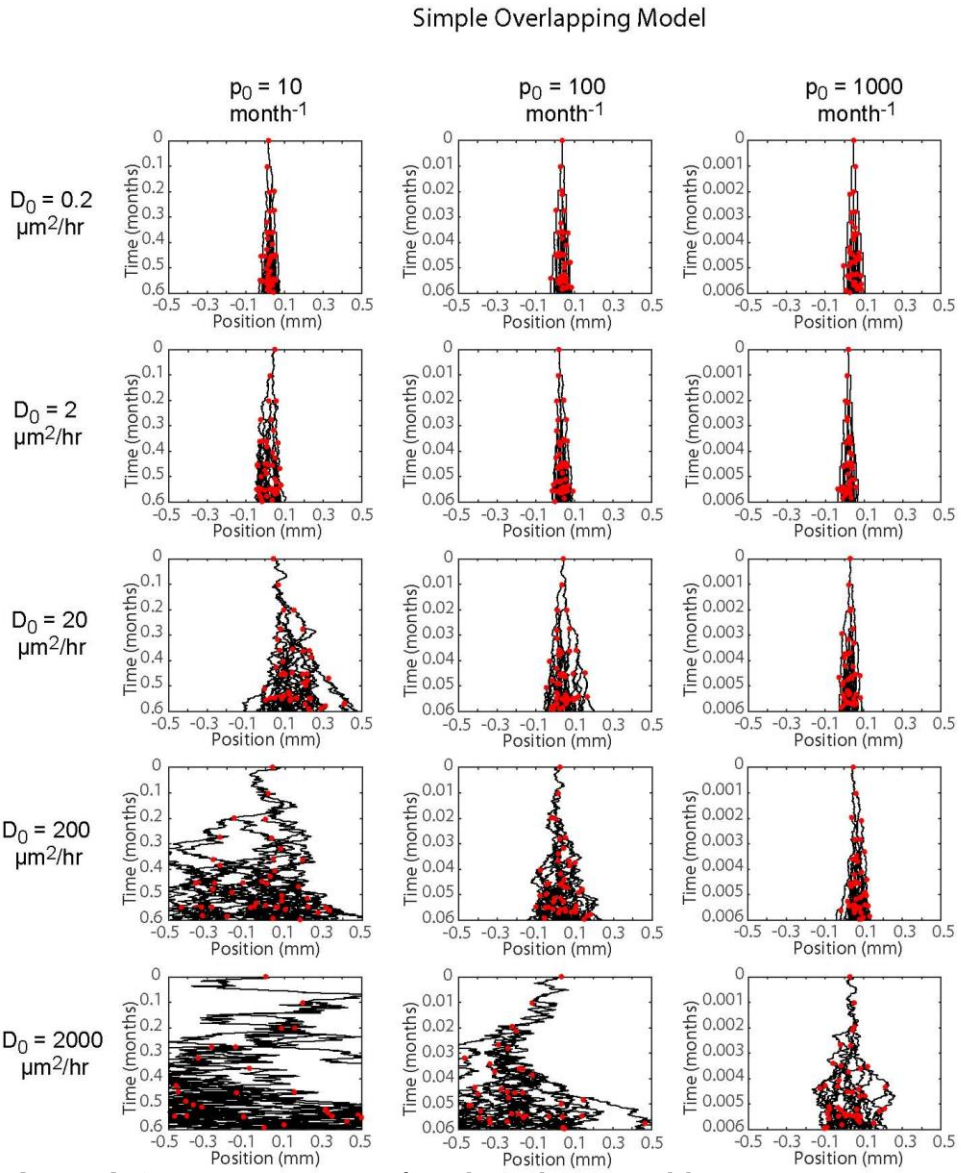
and simplifying with the chain rule, we get

$$D_{app} = D_{ij}\left(\frac{\beta\chi_i}{\beta\chi_i+1}+1\right) \quad \text{(Eq. S16)}$$

The maximum interaction would occur when mole fraction χ_i equals 1. As β tends to infinity we approach the upper limit of $D_{app}$where

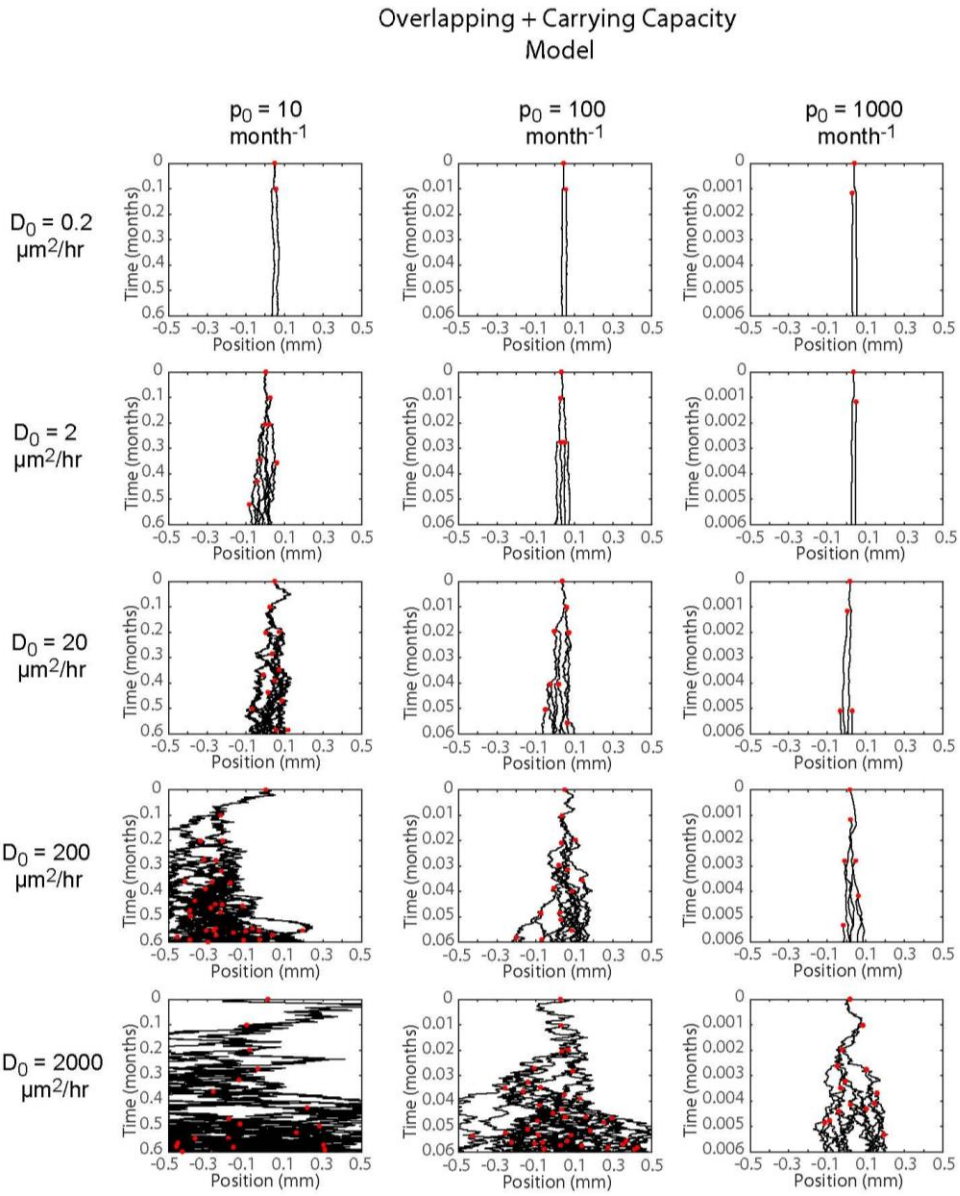
$$\begin{aligned}\beta &\to \infty \\ D_{app} &\to 2D_{ij}\end{aligned} \quad \text{(Eq. S17)}$$

Supplemental Figures



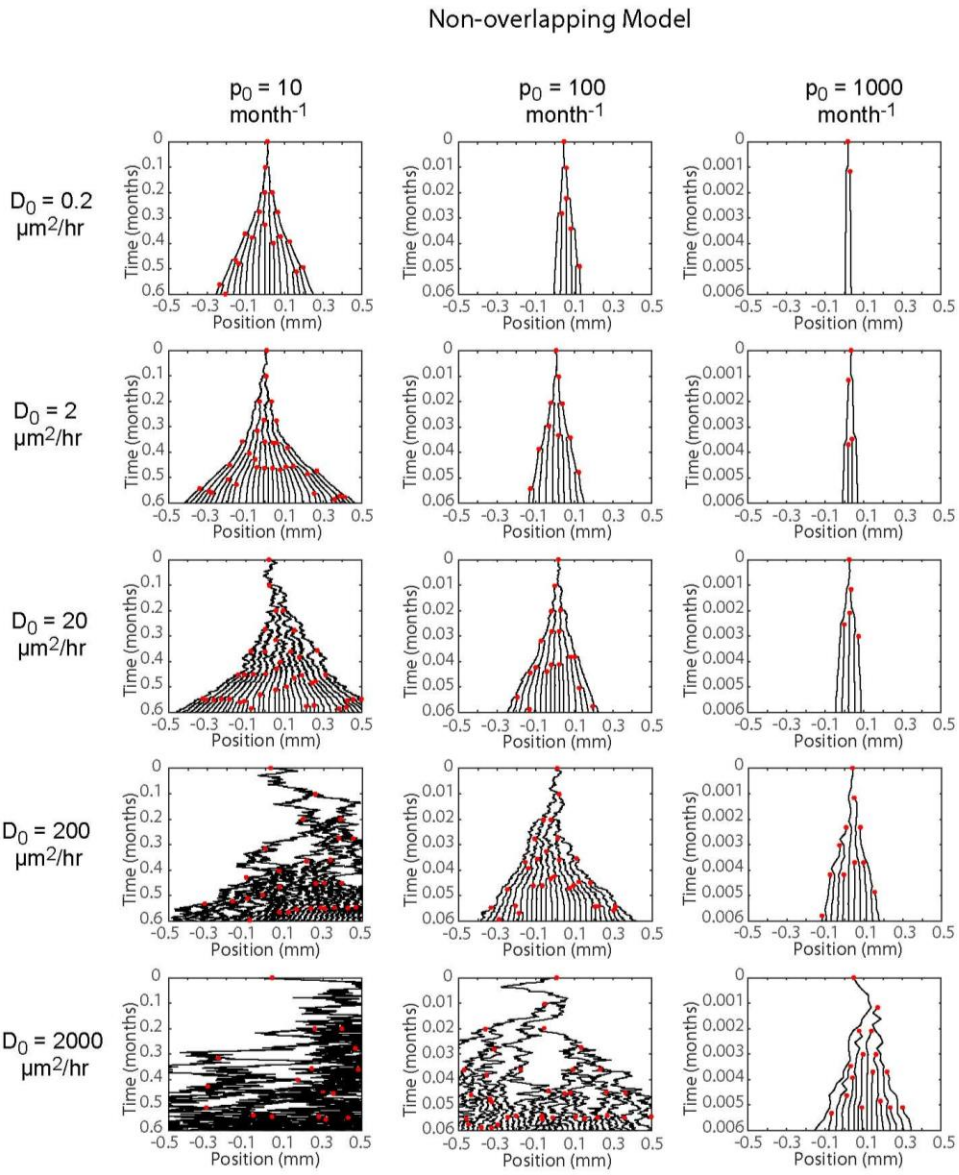
Supplemental Figure 5. Progeny trees of simple overlapping model

Sample outputs of simulations for all 15 parameter sets tested in the simple overlapping model. Rows display identical input diffusion coefficients. Columns display identical input proliferation rates. The time scale of each column is adjusted to display the time needed to achieve six generations of cell division for that particular nominal proliferation rate. Black curves indicate individual cell centroid locations, and red dots indicate the time and location of a division event.



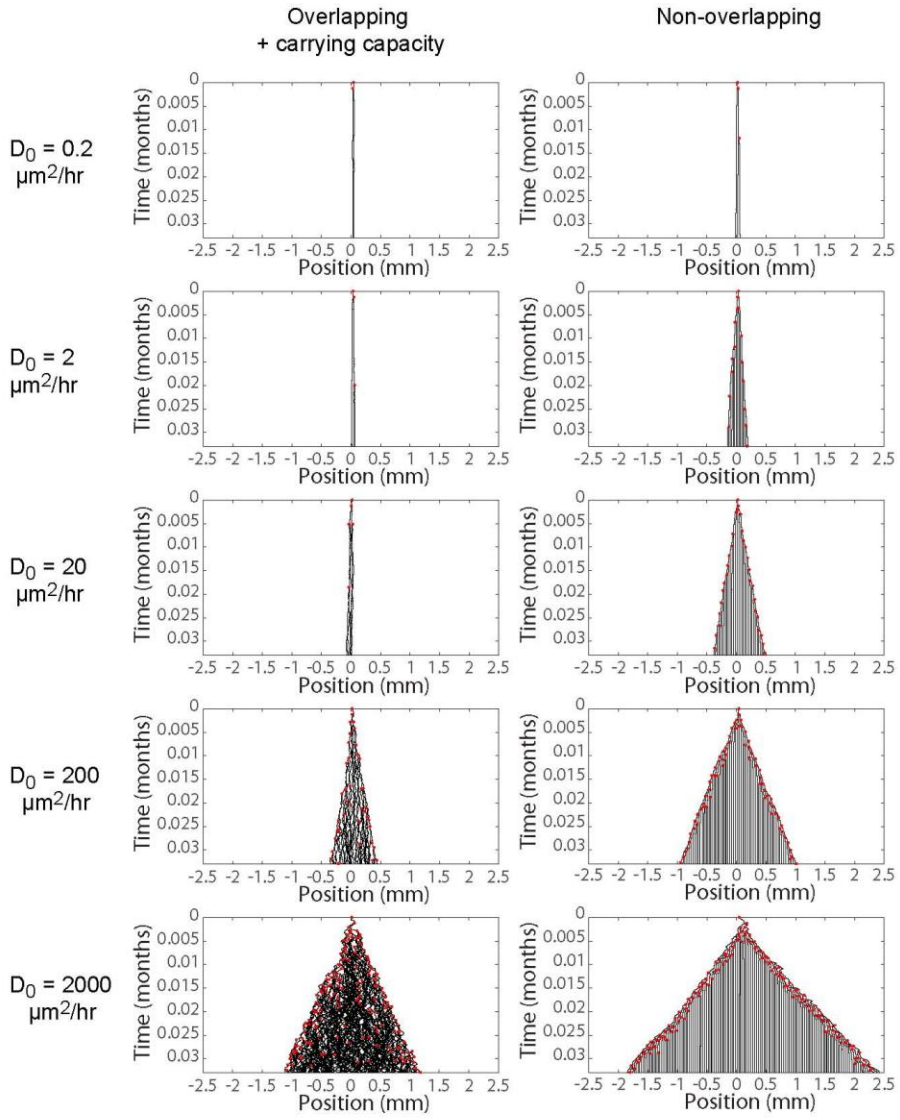
Supplemental Figure 6. Progeny trees of the overlapping + carrying capacity model

Sample outputs of simulations for all 15 parameter sets tested in the overlapping + carrying capacity model. Rows display identical input diffusion coefficients. Columns display identical input proliferation rates. The time scale of each column is adjusted to display the time needed to achieve six generations of cell division for that particular nominal proliferation rate. Black curves indicate individual cell centroid locations, and red dots indicate the time and location of a division event.



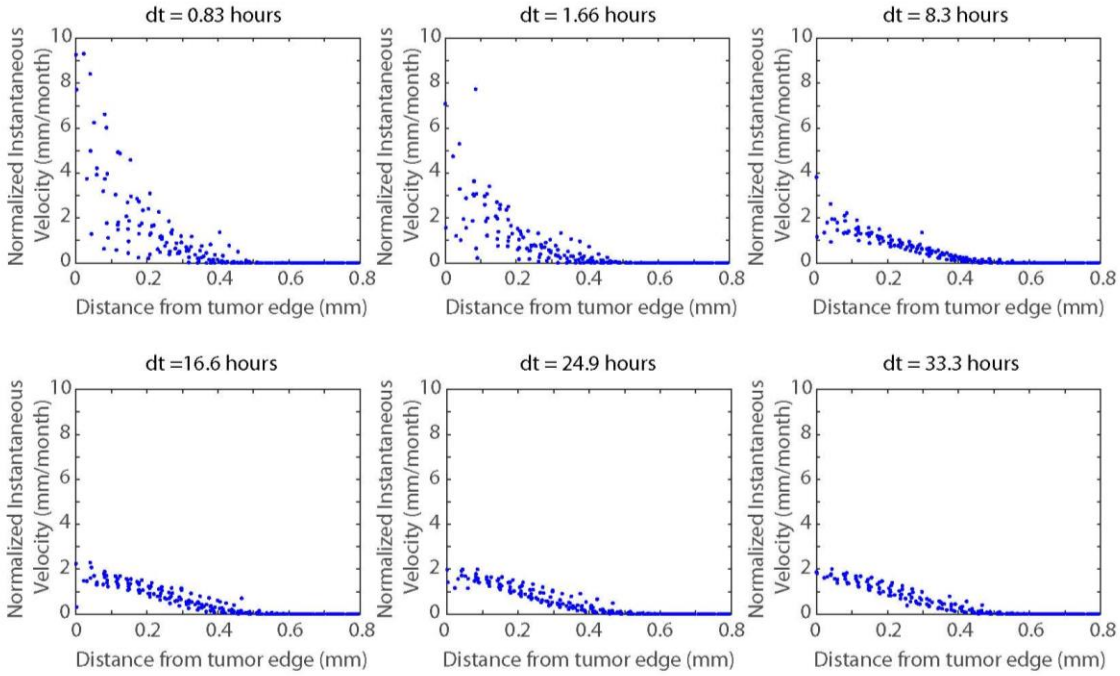
Supplemental Figure 7. Progeny trees of the non-overlapping model

Sample outputs of simulations for all 15 parameter sets tested in the non-overlapping model. Rows display identical input diffusion coefficients. Columns display identical input proliferation rates. The time scale of each column is adjusted to display the time needed to achieve six generations of cell division for that particular nominal proliferation rate. Black curves indicate individual cell centroid locations, and red dots indicate the time and location of a division event.



Supplemental Figure 8. Long term behavior of simulations

Progeny trees for the overlapping+cc model and the non-overlapping model evaluated at p_0=1000 divisions/month with varying diffusion values as indicated on the figure. Jammed behavior continues at long times for the non-overlapping model and an effective convective velocity allows cells to spread faster than in the overlapping+cc model. In the non-overlapping model, proliferating cells (red dots) are only present at the tumor edges, while they are present throughout the tumor in the overlapping+cc model.



Supplemental Figure 9. Effect of cell time interval on cell velocity estimates

Plots of velocity versus distance from the tumor edge evaluated at different time intervals. All values are measured from the same set of simulations for the non-overlapping model with input parameters D_0 = 20 μm^2/hour and p_0 = 10 divisions/month. Only Δt is varied for the calculation of instantaneous velocity. As Δt is increased, the data become less noisy and a ballistic tumor front becomes apparent.

Supplemental Movie Captions

Supplemental Movie 3. Simulation of the simple overlapping uninhibited model

Movie of a sample simulation output of the simple overlapping model evaluated at D_0=20 μm^2/hour and p_0=10 divisions/month. Each cell is represented by two spheres of the same color that gradually separate as the cell grows. A cell division event is indicated by a change in color.

Supplemental Movie 4. Simulation of the overlapping + carrying capacity model

Movie of a sample simulation output of the overlapping + carrying capacity model evaluated at D_0=20 μm^2/hour and p_0=10 divisions/month. Each cell is represented by two spheres of the same color that gradually separate as the cell grows. A cell division event is indicated by a change in color.

Supplemental Movie 5. Simulation of the non-overlapping model

Movie of a sample simulation output of the non-overlapping model evaluated at D_0=20 μm^2/hour and p_0=10 divisions/month. Each cell is represented by two spheres of the same color that gradually separate as the cell grows. A cell division event is indicated by a change in color.

Conclusions

The goal of this work was to investigate the mechanisms involved in GBM tumor dissemination to advance the basic scientific understanding of the disease. In Chapter 1, we investigated how molecular changes in CD44 expression affected cellular migration behavior and organism survival outcomes in a mouse model of glioma. In Chapter 2, we used migration measurements from Chapter 1 to simulate glioma tumor progression and hypothesized why our measured values of glioma migration were so different from estimates in the literature. Both of these studies expanded the current thinking on glioma progression by relying on quantitative approaches and physical principles.

Chapter 1 was focused on the effects of the transmembrane cell adhesion molecule CD44 on glioma progression. Since CD44's primary ligand, HA, is abundant in the brain microenvironment (*27*), and CD44 itself is commonly overexpressed in tumor cells (*52, 76*), we hypothesized that CD44 would significantly contribute to the migratory behavior of GBM cells. Using a *de novo* mouse model of glioma, we showed that CD44 protein expression level affects cell migration in a biphasic manner, in which intermediate expression levels resulted in the highest migration rates. Survival of mice was also biphasically correlated with CD44 expression level, albeit with the opposite curvature to that of migration, with intermediate expression level corresponding to the shortest survival times. This demonstrates that migration and survival are anti-correlated in our model and suggests migration is a possible mechanism for biphasic survival. Using

expression data from previous publications (*49*) and The Cancer Genome Atlas database, we also showed that a similar biphasic survival phenomenon with respect to CD44 expression is present in the human disease. This is in contrast to previous studies which have implicated a negative linear trend with increasing CD44 expression (*42–45*).

This investigation demonstrates many important contributions to the field of glioma research and cell biology, in general. Previous biophysical theory has hypothesized the existence of a biphasic relationship between cell adhesion to the substrate and cell motility (*16*). While this has been previously demonstrated using *in vitro* systems (*17, 18*), our study is the first to demonstrate such a behavior in an animal model and in an *ex vivo* culture system that retains the real tissue ECM. In glioma research, the correlation of CD44 expression to survival outcomes has been disputed, and our finding that CD44 has biphasic relationships with both migration and survival provides a probable rationale for previous conflicting results that assumed a monotonic trend. Moreover, the biphasic relationship that we discovered for CD44 suggests that other molecules may exhibit similar non-linear trends and calls for a paradigm change in the way genomic/proteomic correlations are analyzed to include the possibility of higher order relationships.

In Chapter 2, we compared our estimates of single-cell diffusion coefficients measured from brain slice cultures in Chapter 1 to that of simulation parameters used in modeling GBM tumor progression (Table 2). We found that our values and values of

similar *ex vivo* studies (Table 3) were 1-2 orders of magnitude smaller than that of the estimated tumor growth parameter values. This led us to investigate the assumptions used to arrive at the tumor growth parameter estimates, namely that cells can be approximated as volume-less points in space. Using a BD approach, we simulated 1D tumors both with and without permitting tumor cell overlap. In our simulations we showed that the more physically realistic non-overlapping model can produce tumor growth characteristics similar to that of the overlapping model using smaller cell diffusion coefficients. Thus, we hypothesize that the reason for the quantitative discrepancy between measurements of single-cell migration and estimates is that the estimates do not appropriately consider the contribution of conserved cellular volume on tumor spreading.

In the field of tumor modeling, a direct BD approach has not been utilized previously to challenge the assumptions of the predominant models in the field. Rather than adding more complexity to current models, we simplified our models down to basic physical growth and migration principles for single cells and discovered novel cellular interactions that were previously unappreciated. Particularly, we found that volume exclusion is sufficient to explain different populations of cell migratory behavior without the *ad hoc* addition of the so-called "go-or-grow" dichotomy. Our findings also advocate for the modification of current RD models of GBM to incorporate realistic cell diffusivity values to be more quantitatively accurate between cell-level and tissue-level behaviors in order to generate better predictive models of the disease.

While we make significant contributions to modify current thinking in the field of oncology, there is also much to be done before we have a complete picture of how molecular level changes affect the tissue-level progression of GBM. For instance, we have identified the molecule CD44 as being important for both single-cell migration behavior and overall tumor progression and speculate that CD44 is acting as a molecular clutch to achieve biphasic migration, but have not verified that CD44 is in fact working as a clutch. To test this directly, it would be ideal to measure actin retrograde flow in *ex vivo* slices and detect if there are any changes to flow behavior with the addition of blocking agents to the CD44-HA interaction. Based on the predictions of the motor-clutch model (*22*, *23*), we would expect that engaged clutches would impede retrograde flow; thus, if the addition of CD44 blocking agents results in increased retrograde flow velocities, then we can be confident that CD44 is the relevant clutch in our model. Furthermore, it would be important to test the effects of blocking CD44 directly on cell migration behavior. Therefore, another study could be conducted to test if serial additions of CD44 blocking agents to an *ex vivo* slice is consistent with biphasic predictions. That is, if CD44 is inhibited in a high-expressing cell, we would predict relatively small additions of blocking agents would result in an increase in overall migration behavior. We predict that further addition of blocking agents would bring motility back down. Similarly, inhibition of an intermediate-expressing cell would result in reduced migration, and a low-expressing cell would likely experience no change in migration. In addition to looking at cellular level effects in response to CD44 inhibition, future studies could also

investigate the use of anti-CD44 therapies in our models to test if there are any effects of CD44 inhibition on survival outcomes.

Regarding survival outcomes with relation to CD44 expression, our current animal model has a slight disadvantage in that the KO models (CD44 -/- and CD44 -/- with exogenous CD44) use a genetically modified mouse that is CD44 deficient in all of its cells. This slightly confounds our results in that these models have CD44 deficiencies in their stromal cells as well as the tumor cells. Since the time of our investigation, newer techniques have been developed which would allow for direct KO of CD44 in targeted cells using a CRISPR-Cas system (*123*). To properly investigate the effects of CD44 in tumor cells, it would be ideal to repeat our study using such a system.

As mentioned in Chapter 2, our current computational model of GBM progression has a few drawbacks, as well. Namely, the current model is only 1D and tumor progression typically occurs in 3D. In the future, we would like to extend this model to three dimensions and use our data of mouse glioma progression (Chapter 1) to validate our simulations. From our animal studies, we have direct *ex vivo* measurements of cell diffusion coefficients, *in vitro* measurements of low-passage cell line proliferation rates, and estimates for the typical tumor size at the time of death. Using the radial tumor velocity as an output parameter, we can directly compare our simulation behavior to a real disease model. Furthermore, preliminary work has been completed to generate different mouse models of glioma under different oncogenes; specifically, models have

been developed for EGFRvIII, PDGFb, and activated AKT, in hopes to capture characteristic growth differences in the GBM subtypes (EGFRvIII = classical subtype, PDGFb = proneural subtype, activated AKT = mesenchymal subtype). In future studies, we plan to measure growth parameters in each of these different animal models. Preliminary studies have shown that the different animal models not only have different migration and proliferation characteristics, but also different characteristic times of progression (data not shown). When proliferation and migration measurements are used as input parameters in our simulations, we anticipate the simulated spreading velocity will be consistent with observations in each of the models.

Beyond simulating mouse gliomas, we also intend to use our non-overlapping approach to simulate human GBMs. One of the primary issues with current RD simulations of GBM is that they use MRI or CT imaging to estimate parameters for their models. While it is advantageous that these techniques are non-invasive, standard protocol for GBM is to perform a biopsy of the tissue to confirm diagnosis after a lesion is suspected . Once the diagnosis has been made, leftover tissues can be used for GBM cell characterization, as demonstrated by our lab in collaboration with neurosurgeons at the University of Minnesota (results unpublished). We hope that in the future we can use these biopsied tissues to measure proliferation and migration parameters directly from a patient's own cells cultured in an engineered *in vitro* environment. The advantage of this approach is that we would have a patient-specific model based on actual cell

measurements rather than estimations. Furthermore, in such an *in vitro* environment, potential therapies could be directly tested on the patient's cell, and parameters to the simulation could be modified to reflect the expected cell behavior in response to treatment. This type of patient-specific model will likely be key to developing accurate predictive models of GBM progression and a useful, unbiased tool for clinicians in suggesting the best course of treatment.

References

1. R. L. Siegel, K. D. Miller, A. Jemal, Cancer Statistics , 2015. **65**, 5–29 (2015).

2. M. B. Kastan, J. Bartek, Cell-cycle checkpoints and cancer. **432** (2004).

3. R. a Weinberg, *Th Biology of Cancer* (Garland Science, New York, New York, 2007).

4. T. Demuth, M. E. Berens, Molecular mechanisms of glioma cell migration and invasion, 217–228 (2004).

5. J. R. Macdougall, L. M. Matrisian, Contributions of tumor and stromal matrix metalloproteinases to tumor progression , invasion and metastasis, 351–362 (1995).

6. P. D. Mourad, L. Farrell, L. D. Stamps, M. R. Chicoine, D. L. Silbergeld, Why are systemic glioblastoma metastases rare ? Systemic and cerebral growth of mouse glioblastoma. **63**, 511–519 (2005).

7. CBTRUS, "CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2004-2007" (2011).

8. D. B. Hoelzinger, T. Demuth, M. E. Berens, Autocrine factors that sustain glioma invasion and paracrine biology in the brain microenvironment. *J. Natl. Cancer Inst.* **99**, 1583–93 (2007).

9. R. Ramnarayan, S. Dodd, K. Das, V. Heidecke, N. G. Rainov, Overall survival in patients with malignant glioma may be significantly longer with tumors located in deep grey matter. **260**, 49–56 (2007).

10. A. C. Bellail, S. B. Hunter, D. J. Brat, C. Tan, E. G. Van Meir, Microregional extracellular matrix heterogeneity in brain modulates glioma cell invasion. **36**, 1046–1069 (2004).

11. M. L. Gardel, I. C. Schneider, Y. Aratyn-schaus, C. M. Waterman, Mechanical Integration of Actin and Adhesion Dynamics in Cell Migration, 315–335 (2010).

12. B. Alberts *et al.*, *Molecular Biology of the Cell* (Garland Science, New York, New York, ed. 5, 2008).

13. C. C. Dufort, M. J. Paszek, V. M. Weaver, Balancing forces : architectural control of mechanotransduction. *Nat. Publ. Gr.* **12**, 308–319 (2011).

14. K. Kessenbrock, V. Plaks, Z. Werb, Review Matrix Metalloproteinases : Regulators of the Tumor Microenvironment (2010), doi:10.1016/j.cell.2010.03.015.

15. S. C. Massey, M. C. Assanah, K. A. Lopez, P. Canoll, K. R. Swanson, Glial progenitor cell recruitment drives aggressive glioma growth : mathematical and experimental modelling, 1757–1766 (2012).

16. P. A. DiMilla, K. Barbee, D. A. Lauffenburger, Mathematical model for the effects of adhesion and mechanics on cell migration speed. *Biophys. J.* **60**, 15–37 (1991).

17. P. A. Dimilla, J. A. Stone, J. A. Quinn, S. M. Albelda, D. A. Lauffenburger, Maximal migration of human smooth muscle cells on fibronectin and type IV collagen occurs at an intermediate attachment strength. *J. Cell Biololgy*. **122**, 729–737 (1993).

18. S. Palecek, J. Loftus, M. Ginsberg, D. Lauffenburger, A. Horwitz, Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness. *Nature*. **385**, 537–540 (1997).

19. R. J. Pelham, Y. Wang, Cell locomotion and focal adhesions are regulated by substrate flexibility. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13661–13665 (1997).

20. C. Lo, H. Wang, M. Dembo, Y. Wang, Cell Movement Is Guided by the Rigidity of the Substrate. *Biophys. J.* **79**, 144–152 (2000).

21. T. Mitchison, M. Kirschner, Cytoskeletal Dynamics and Nerve Growth. *Neuron*. **1**, 761–772 (1988).

22. C. E. Chan, D. J. Odde, Traction Dynamics of Filopodia on. *Science (80-.).* **322**, 1687–1691 (2008).

23. B. L. Bangasser, S. S. Rosenfeld, D. J. Odde, Determinants of maximal force transmission in a motor-clutch model of cell traction in a compliant microenvironment. *Biophys. J.* **105**, 581–92 (2013).

24. K. Pham *et al.*, Density-dependent quiescence in glioma invasion : instability in a simple reaction – diffusion model for the migration / proliferation dichotomy. **3758** (2012), doi:10.1080/17513758.2011.590610.

25. A. M. G. F. Calvo, L. A. Pérez, R. Víctor, Hypoxic Cell Waves Around Necrotic Cores in Glioblastoma : A Biomathematical Model and Its Therapeutic Implications, 2875–2896 (2012).

26. A. Kwiatkowska, M. Symons, in *Advances in experimental medicine and biology* (2013), vol. 986, pp. 209–220.

27. U. Novak, a H. Kaye, Extracellular matrix and the brain: components and function. *J. Clin. Neurosci.* **7**, 280–290 (2000).

28. C. L. Gladson, The Extracellular Matrix of Gliomas: Modulation of Cell Function. *J. Neuropathol. Exp. Neurol.* **58**, 1029–1040 (1999).

29. T. A. Ulrich, E. M. D. J. Pardo, S. Kumar, The Mechanical Rigidity of the Extracellular Matrix Regulates the Structure , Motility , and Proliferation of Glioma Cells, 4167–4174 (2009).

30. A. Giese *et al.*, Migration of Human Glioma Cells on Myelin. *Neurosurgery*. **38**, 775–764 (1996).

31. C. Beadle *et al.*, The Role of Myosin II in Glioma Invasion of the Brain. *Mol. Biol. Cell*. **19**, 3357–3368 (2008).

32. J. Zamecnik, The extracellular space and matrix of gliomas. *Acta Neuropathol.* **110**, 435–442 (2005).

33. B. P. Toole, Hyaluronan-CD44 Interactions in Cancer: Paradoxes and Possibilities. *Clin. Cancer Res.* **15**, 7462–7468 (2009).

34. H. Ponta, L. Sherman, P. a Herrlich, CD44: from adhesion molecules to signalling regulators. *Nat. Rev. Mol. Cell Biol.* **4**, 33–45 (2003).

35. P. Friedl, S. Alexander, Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell*. **147**, 992–1009 (2011).

36. R. Stupp, M. E. Hegi, M. R. Gilbert, A. Chakravarti, Chemoradiotherapy in malignant glioma: standard of care and future directions. *J. Clin. Oncol.* **25**, 4127–36 (2007).

37. F. Lefranc, J. Brotchi, R. Kiss, Possible future issues in the treatment of glioblastomas: special emphasis on cell migration and the resistance of migrating glioblastoma cells to apoptosis. *J. Clin. Oncol.* **23**, 2411–22 (2005).

38. M. Culty *et al.*, The hyaluronate receptor is a member of the CD44 (H-CAM) family of cell surface glycoproteins. *J. Cell Biol.* **111**, 2765–2774 (1990).

39. A. Merzak, S. Koocheckpour, G. J. Pilkington, CD44 Mediates Human Glioma Cell Adhesion and Invasion in Vitro CD44 Mediates Human Glioma Cell Adhesion. *Cancer Res.* **54**, 3988–3992 (1994).

40. R. Breyer *et al.*, Disruption of intracerebral progression of C6 rat glioblastoma by in vivo treatment with anti-CD44 monoclonal antibody. *J. Neurosurg.* **92**, 140–9 (2000).

41. D. Naor, S. Nedvetzki, I. Golan, L. Melnik, Y. Faitelson, CD44 in cancer. *Crit. Rev. Clin. Lab. Sci.* **39**, 527–579 (2002).

42. J. Anido *et al.*, TGF-β Receptor Inhibitors Target the CD44(high)/Id1(high) Glioma-Initiating Cell Population in Human Glioblastoma. *Cancer Cell*. **18**, 655–68 (2010).

43. M. Jijiwa *et al.*, CD44V6 regulates growth of brain tumor stem cells partially through the AKT-mediated pathway. *PLoS One*. **6** (2011).

44. A. Pietras *et al.*, Osteopontin-CD44 signaling in the glioma perivascular niche enhances cancer stem cell phenotypes and promotes aggressive tumor growth. *Cell Stem Cell*. **14**, 357–369 (2014).

45. K. P. L. Bhat *et al.*, Mesenchymal Differentiation Mediated by NF-κB Promotes Radiation Resistance in Glioblastoma. *Cancer Cell*. **24**, 331–346 (2013).

46. S. M. Ranuncolo, V. Ladeda, S. Specterman, CD44 Expression in Human Gliomas. *J. Surg. Oncol.* **79**, 30–36 (2002).

47. K.-C. Wei *et al.*, Evaluation of the prognostic value of CD44 in glioblastoma multiforme. *Anticancer Res.* **30**, 253–9 (2010).

48. S. M. Wiesner *et al.*, De novo induction of genetically engineered brain tumors in mice using plasmid DNA. *Cancer Res.* **69**, 431–9 (2009).

49. H. S. Phillips *et al.*, Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell*. **9**, 157–73 (2006).

50. R. G. W. Verhaak *et al.*, Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. **17**, 98–110 (2010).

51. W.-Y. Cheng, J. J. Kandel, D. J. Yamashiro, P. Canoll, D. Anastassiou, A multi-cancer mesenchymal transition gene expression signature is associated with prolonged time to recurrence in glioblastoma. *PLoS One*. **7**, e34705 (2012).

52. T. Yoshida, Y. Matsuda, Z. Naito, T. Ishiwata, CD44 in human glioma correlates with histopathological grade and cell migration. *Pathol. Int.* **62**, 463–70 (2012).

53. D. a Lauffenburger, a F. Horwitz, Cell migration: a physically integrated molecular process. *Cell*. **84**, 359–69 (1996).

54. D. E. Discher, P. Janmey, Y.-L. Wang, Tissue cells feel and respond to the stiffness of their substrate. *Science*. **310**, 1139–1143 (2005).

55. J. E. Molloy, J. E. Burns, J. Kendrick-Jones, R. T. Tregear, D. C. White, Movement and force produced by a single myosin head. *Nature*. **378** (1995), pp. 209–212.

56. T. P. Lele, C. K. Thodeti, J. Pendse, D. E. Ingber, Investigating complexity of protein-protein interactions in focal adhesions. *Biochem. Biophys. Res. Commun.* **369**, 929–934 (2008).

57. G. Jiang, G. Giannone, D. R. Critchley, E. Fukumoto, M. P. Sheetz, Two-piconewton slip bond between fibronectin and the cytoskeleton depends on talin. *Nature*. **424**, 334–337 (2003).

58. S. V. Plotnikov, A. M. Pasapera, B. Sabass, C. M. Waterman, Force fluctuations within focal adhesions mediate ECM-rigidity sensing to guide directed cell migration. *Cell*. **151**, 1513–1527 (2012).

59. D. a. Schafer, P. B. Jennings, J. a. Cooper, Dynamics of capping protein and actin assembly in vitro: Uncapping barbed ends by polyphosphoinositides. *J. Cell Biol.* **135**, 169–179 (1996).

60. B. L. Bangasser, D. J. Odde, Master equation-based analysis of a motor-clutch model for cell traction force. *Cell. Mol. Bioeng.* **6**, 449–459 (2013).

61. G. I. Bell, Models for the specific adhesion of cells to cells. *Science (80-.).* **200**, 618–627 (1978).

62. L. S. Tobacman, E. D. Korn, The kinetics of actin nucleation and polymerization. *J. Biol. Chem.* **258**, 3207–3214 (1983).

63. M. Kasai, S. Asakura, F. Oosawa, The cooperative nature of G-F transformation of actin. *Biochim. Biophys. Acta*. **57**, 22–31 (1962).

64. D. T. Gillesple, Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977).

65. R. B. Dickinson, R. T. Tranquillo, Optimal estimation of cell movement indices from the statistical analysis of cell tracking data. *AIChE J.* **39**, 1995–2010 (1993).

66. Y. Kim, S. Kumar, CD44-mediated Adhesion to Hyaluronic Acid Contributes to Mechanosensing and Invasive Motility. *Mol. Cancer Res.*, 1416–1429 (2014).

67. L. Y. W. Bourguignon, Hyaluronan-mediated CD44 activation of RhoGTPase signaling and cytoskeleton function promotes tumor progression. *Semin. Cancer Biol.* **18**, 251–259 (2008).

68. E. a. Turley, P. W. Noble, L. Y. W. Bourguignon, Signaling properties of hyaluronan receptors. *J. Biol. Chem.* **277**, 4589–4592 (2002).

69. Y. Chuang *et al.*, Role of Synaptojanin 2 in Glioma Cell Migration and Invasion Role of Synaptojanin 2 in Glioma Cell Migration and Invasion, 8271–8275 (2004).

70. M. J. Paszek *et al.*, The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature*. **511**, 319–25 (2014).

71. J. Schlessinger, I. Lax, M. Lemmon, Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell*. **83**, 357–60 (1995).

72. Q. Yu, I. Stamenkovic, Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion. *Genes Dev.* **13**, 35–48 (1999).

73. L. Y. W. Bourguignon *et al.*, Cytoskeleton-Mediated Tumor Cell Migration and Matrix Metalloproteinase (MMP-9) Association in Metastatic Breast Cancer Cells. *J. Cell. Physiol.* **215**, 206–215 (1998).

74. C. Chetty *et al.*, MMP-9 induces CD44 cleavage and CD44 mediated cell migration in glioblastoma xenograft cells. *Cell. Signal.* **24**, 549–559 (2012).

75. O. Nagano, H. Saya, Mechanism and biological significance of CD44 cleavage. *Cancer Sci.* **95**, 930–935 (2004).

76. L. Prochazka, R. Tesarik, J. Turanek, Regulation of alternative splicing of CD44 in cancer. *Cell. Signal.* **26**, 2234–2239 (2014).

77. C. Faller, O. Guvench, Terminal sialic acids on CD44 N-glycans can block hyaluronan binding by forming competing intramolecular contacts with arginine sidechains. *Proteins Struct. Funct. Bioinforma.*, 3079–3089 (2014).

78. J. Lesley, N. English, A. Preschl, J. Gregoroff, R. Hyman, Variant Cell Lines Selected for Alterations in the Function of Hyaluronan Receptor CD44 show Differences in Glycosylation. *J. Exp. Med.* **182**, 431–437 (1995).

79. T. P. Skelton, C. Zeng, A. Nocks, I. Stamenkovic, Glycosylation provides both stimulatory and inhibitory effects on cell surface and soluble CD44 binding to hyaluronan. *J. Cell Biol.* **140**, 431–446 (1998).

80. L. Soroceanu *et al.*, Id-1 is a key transcriptional regulator of glioblastoma aggressiveness and a novel therapeutic target. *Cancer Res.* **73**, 1559–1569 (2013).

81. Y. Xu, I. Stamenkovic, Q. Yu, CD44 attenuates activation of the hippo signaling pathway and is a prime therapeutic target for glioblastoma. *Cancer Res.* **70**, 2455–64 (2010).

82. H. Okada, J. Yoshida, M. Sokabe, T. Wakabayashi, M. Hagiwara, Suppression of CD44 expression decreases migration and invasion of human glioma cells. *Int. J. cancer*. **66**, 255–260 (1996).

83. Hoffman-LaRoche, Open-Label Multicenter 2-Arm Phase I Study of RO5429083 With Dose-Escalation and Extension Cohorts, and Imaging Cohorts With RO5429083 and 89Zr-labeled RO5429083, in Patients With Metastatic and/or Locally Advanced, CD44-Expressing, Malignant Solid Tumors, (available at https://clinicaltrials.gov/ct2/show/NCT01358903 NLM Identifier: NCT01358903).

84. A. Kanno *et al.*, Periostin, secreted from stromal cells, has biphasic effect on cell migration and correlates with the epithelial to mesenchymal transition of human pancreatic cancer cells. *Int. J. Cancer*. **122**, 2707–2718 (2008).

85. T. Alarcon, H. M. Byrne, P. K. Maini, A multiple scale model for tumor growth. *Multiscale Model. Simul.* **3**, 440–475 (2005).

86. T. Harko, Travelling wave solutions of the reaction-diffusion mathematical model of glioblastoma growth: an Abel equation based approach. *Math. Biosci. Eng.* **12**, 41–69 (2015).

87. H. Hatzikirou, L. Brusch, C. Schaller, M. Simon, A. Deutsch, Prediction of traveling front behavior in a lattice-gas cellular automaton model for tumor invasion. *Comput. Math. with Appl.* **59**, 2326–2339 (2010).

88. H. L. P. Harpold, E. C. Alvord, K. R. Swanson, C. Cells, S. G. Cells, The Evolution of Mathematical Modeling of Glioma Proliferation and Invasion. **66**, 1–9 (2007).

89. D. E. Woodward *et al.*, A mathematical model of glioma growth : the effect of extent of surgical resection, 269–288 (1996).

90. P. Tracqui *et al.*, A mathematical model of glioma growth : the effect of chemotherapy on spatio-temporal growth. **05** (1995).

91. C. H. Wang *et al.*, Mathematical Oncology Prognostic Significance of Growth Kinetics in Newly Diagnosed Glioblastomas Revealed by Combining Serial Imaging with a Novel Biomathematical Model, 9133–9141 (2009).

92. A. R. A. Anderson, K. A. Rejniak, P. Gerlee, V. Quaranta, Mathematical Biology Microenvironment driven invasion : a multiscale multimodel investigation. *J. ma*. **58**, 579–624 (2009).

93. B. Waclaw *et al.*, turnover limit intratumour heterogeneity (2015), doi:10.1038/nature14971.

94. Y. Jiang, J. Pjesivac-grbovic, C. Cantrell, J. P. Freyer, A Multiscale Model for Avascular Tumor Growth. *Biophys. J.* **89**, 3884–3894 (2005).

95. E. L. Stott, N. F. Britton, Stochastic Simulation of Benign Avascular Tumour Growth Using the Potts Model. **7177** (1999).

96. S. Turner, J. A. Sherratt, Intercellular Adhesion and Cancer Invasion : A Discrete Simulation Using the Extended Potts Model, 85–100 (2002).

97. S. T. Johnston, M. J. Simpson, R. E. Baker, Modelling the movement of interacting cell populations : A moment dynamics approach. *J. Theor. Biol.* **370**, 81–92 (2015).

98. L. Zhang, C. A. Athale, T. S. Deisboeck, Development of a three-dimensional multiscale agent-based tumor model : Simulating gene-protein interaction profiles , cell phenotypes and multicellular patterns in brain cancer. **244**, 96–107 (2007).

99. D. Drasdo, H. Stefan, A single-cell-based model of tumor growth in vitro : monolayers and spheroids. **133**, doi:10.1088/1478-3975/2/3/001.

100. Y. Kim, M. A. Stolarska, H. G. Othmer, The role of the microenvironment in tumor growth and invasion. *Prog. Biophys. Mol. Biol.* **106**, 353–379 (2011).

101. P. Gerlee, S. Nelander, The Impact of Phenotypic Switching on Glioblastoma Growth and Invasion. **8** (2012), doi:10.1371/journal.pcbi.1002556.

102. E. Khain, M. Katakowski, N. Charteris, F. Jiang, M. Chopp, Migration of adhesive glioma cells : Front propagation and fingering. **011904**, 1–8 (2012).

103. K. R. Swanson, E. C. A. Jr, J. D. Murray, A quantitative model for differential motility of gliomas in grey and white matter, 317–329 (2000).

104. K. R. Swanson, C. Bridge, J. D. Murray, E. C. Alvord, Virtual and real brain tumors : using mathematical modeling to quantify glioma growth and invasion. **216**, 1–10 (2003).

105. B. A. Giese, R. Bjerkvig, M. E. Berens, M. Westphal, Cost of Migration : Invasion of Malignant Gliomas and Implications for Treatment (2015), doi:10.1200/JCO.2003.05.063.

106. S. E. Eikenberry, E. J. Kostelich, Y. Kuang, Virtual glioblastoma : growth , migration and treatment in a three-dimensional, 511–528 (2009).

107. R. Rockne, E. Alvord, J. Rockhill, K. R. Swanson, A mathematical model for brain tumor response to radiation therapy. *J. Math. Biol.* **58**, 561–578 (2009).

108. J. W. Henson, P. Gaviani, R. G. Gonzalez, MRI in treatment of adult gliomas. **6**, 167–175 (2005).

109. C. Jiang, C. Cui, L. Li, Y. Shao, The Anomalous Diffusion of a Tumor Invading with Different Surrounding Tissues. **9** (2014), doi:10.1371/journal.pone.0109784.

110. P. K. Burgess, P. M. Kulesa, J. D. Murray, E. C. Alvord, The Interaction of Growth Rates and Diffusion Coefficients in a Three-dimensional Mathematical Model of Gliomas. *J. Neuropathol. Exp. Neurol.* **56**, 704–713 (1997).

111. L. Hathout, B. M. Ellingson, T. Cloughesy, W. B. Pope, A novel bicompartmental mathematical model of glioblastoma multiforme, 825–832 (2015).

112. S. Ivkovic *et al.*, Direct inhibition of myosin II effectively blocks glioma invasion in the presence of multiple motogens. **23**, 533–542 (2012).

113. A. M. Stein, T. Demuth, D. Mobley, M. Berens, L. M. Sander, A Mathematical Model of Glioblastoma Tumor Spheroid Invasion in a Three-Dimensional In Vitro Experiment. *Biophys. J.* **92**, 356–365 (2007).

114. M. J. Simpson *et al.*, Quantifying the roles of cell motility and cell proliferation in a circular barrier assay (2013).

115. R. Rockne *et al.*, Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo : a mathematical modeling approach. *Phys. Med. Biol.* **55**, 3271–3285 (2010).

116. P. Maiuri *et al.*, Supplemental Information The first World Cell Race.

117. I. V Maly, G. G. Borisy, Self-organization of a propulsive actin network as an evolutionary process. **98** (2001).

118. D. Irimia, M. Tone, Spontaneous migration of cancer cells under conditions of mechanical confinement. *Integr. Biol.* **1**, 506–512 (2013).

119. V. A. Cuddapah, S. Robel, S. Watkins, H. Sontheimer, REVIEWS A neurocentric perspective on glioma invasion. *Nat. Publ. Gr.* **15**, 455–465 (2014).

120. D. C. Markham, M. J. Simpson, P. K. Maini, E. A. Gaffney, R. E. Baker, Comparing methods for modelling spreading cell fronts. *J. Theor. Biol.* **353**, 95–103 (2014).

121. G. Danuser, J. Allard, A. Mogilner, Mathematical Modeling of Eukaryotic Cell Migration : Insights Beyond Experiments, doi:10.1146/annurev-cellbio-101512-122308.

122. G. A. Truskey, F. Yuan, D. F. Katz, *Transport Phenomena in Biological Systems* (Pearson/Prentice Hall, Upper Saddle River, NJ, ed. 2, 2004).

123. J. D. Sander, J. K. Joung, review CRISPR-Cas systems for editing , regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014).

Appendix A. Biographical Sketch

Rebecca L. Klank grew up in the suburban city of Walled Lake, Michigan. From an early age, she found her niche as a zealous student. This enthusiasm for learning carried her through her formative years and through her undergraduate tenure at Michigan Technological University. While studying at Michigan Tech, Rebecca also involved herself with other endeavors such as gaining leadership experience from the Delta Zeta sorority, teaching communication skills at the Michigan Tech Writing Center, and conducting research under the direction of Dr. Ryan Gilbert. In 2009, Rebecca graduated Suma Cum Laude from Michigan Tech with a B.S. in Biomedical Engineering. At the culmination of her time at Michigan Tech she received the distinction as the 2009 Woman of Promise from the Biomedical Engineering Department for her dedication and enthusiasm for the field. When her days as a student were nearing their end, she simply decided she was not finished learning, so she decided to continue her studies as a Ph.D. student at the University of Minnesota. Starting in the fall of 2009, Rebecca joined the laboratory of Dr. David Odde and began conducting research in glioma migration and progression. While at the University of Minnesota, Rebecca spearheaded the first Biomedical Engineering colloquium events to allow graduate students opportunities to present their research and practice their presentation skills. Along with research, Rebecca also dedicated time to tutoring middle school, high school, and undergraduate students in math and science curricula while in graduate school. In her next phase of life, Rebecca intends to continue teaching to spread her passion for math and science.