**Quantitative analysis of gradient sensing: towards building predictive models of chemotaxis in cancer**

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1016/j.ceb.2012.01.001">http://dx.doi.org/10.1016/j.ceb.2012.01.001</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Elsevier</td>
</tr>
<tr>
<td>Version</td>
<td>Author's final manuscript</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sat Apr 06 23:40:03 EDT 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/99475">http://hdl.handle.net/1721.1/99475</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a></td>
</tr>
</tbody>
</table>
Quantitative Analysis of Gradient Sensing: Towards Building Predictive Models of Chemotaxis in Cancer

Shannon K. Hughes-Alford and Douglas A. Lauffenburger

1Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139

Abstract

Chemotaxis of tumor cells in response to a gradient of extracellular ligand is an important step in cancer metastasis. The heterogeneity of chemotactic responses in cancer has not been widely addressed by experimental or mathematical modeling techniques. However, recent advancements in chemoattractant presentation, fluorescent-based signaling probes, and phenotypic analysis paradigms provide rich sources for building data-driven relational models that describe tumor cell chemotaxis in response to a wide variety of stimuli. Here we present gradient sensing, and the resulting chemotactic behavior, in a ‘cue-signal-response’ framework and suggest methods for utilizing recently reported experimental methods in data-driven modeling ventures.

Introduction

Although significant progress has been made in the cancer biology field, the mechanisms by which primary tumor cells metastasize to distant sites in the body are widely unknown. One important aspect of carcinoma invasion is the presence of multiple stimuli within the tumor microenvironment, including a wide variety of biochemical factors [1], biophysical effects of the extracellular matrix [2], and interstitial flow [3]. The spatiotemporal distributions of these cues are dynamic and demand rigorous analysis for interpretation. Tumor cells must sense a cue, most often through ligand binding, integrate signals, through activation of multiple intracellular networks, and respond through mobilization of cytoskeletal machinery and concomitant modification of surrounding matrix. Often it is only through simplification of experimental assays and implementation of mathematical and engineering principles can one begin to understand mechanisms of gradient sensing and response.

Directed cell migration, or chemotaxis, results from the ability of cells to process an extracellular cue through a complex intracellular signaling network to produce a coordinated and robust response. This inherent ‘cue-signal-response’ process represents a convenient paradigm for deconstructing the processes by which cells utilize gradients of chemical and mechanical stimuli within their microenvironment. We and others have taken advantage of this framework to build quantitative, data-driven, predictive models of cell response, analyzing cancer-relevant cell motility behaviors in two- and three-dimensional in vitro settings [4–7]; these types of relational models (see Box 1) have recently been extended to in...
vivo applications [8]. In this mini-review we concentrate on the quantitative cell biological and biochemical measurements that can serve as inputs (cues, signals, and responses) to modeling efforts with a focus on previously underutilized parameters that hold potential for improving predictive ability in application to chemotaxis in cancer. Recent advancements in experimental strategies for obtaining multi-dimensional and information-rich datasets will be emphasized. Particular mathematical techniques will not be discussed here; rather, readers are directed to reviews focused on the modeling approaches [9,10].

**Box 1**

**Relational modeling for predicting chemotactic behavior**

Relational modeling can predict cell motility behaviors based upon measurement of intracellular signaling network state in vitro [4,5,7,74]. Modeling efforts begin through stimulation of motility through a set of ‘cues’, which can be addition of extracellular ligand, intracellular perturbation (siRNA or chemical inhibitor), and/or changes to the extracellular matrix. Highly multidimensional data is collected through biochemical or imaging techniques (panel 1). Representation of this type of data typically requires data-reduction techniques such as principal components analysis to reveal relationships between the measured ‘signals’ and the input ‘cues’ (panel 2). To gain insight to motility responses, signals and/or cues can be regressed against cellular ‘response’, often using techniques such as partial least squares or multilinear regression. These analyses correlate the clustered ‘signal’ and/or ‘cue’ data to a phenotypic output. For example, specific signals may be more representative of cell speed versus cell persistence (panel 2). Relational models are best utilized to generate hypotheses, such as specific ‘signals’ or ‘cues’ that may be inhibited to modulate cell motility (panel 3).

**Cues**

A major challenge of data-driven modeling is the need for generation of appropriate multivariate datasets, sufficiently broad in scope to comprehend a significant range of response behaviors driven by a spectrum of potentially germane cue conditions. More specifically, when studying a heterogeneous response, such as chemotaxis, how does one collect relevant data from enough cells that are exposed to a consistent set of cues? In contrast to macrogradients that persist over long distances, such as those that direct organogenesis [11], control lymphocyte egress [12], or are encountered in some wound healing models [13], most tumor cells encounter local gradients of signaling molecules arising from autocrine or paracrine secretion of chemokines and growth factors. This situation has been observed in vivo using intravital microscopy in murine mammary carcinoma, where tumor-associated macrophages induce tumor cell migration via secretion of Epidermal Growth Factor (EGF). In turn, tumor cells secrete Colony Stimulating Factor-1 (CSF-1), which is a potent chemoattractant for macrophages [14,15]. In vivo collection of
stromal and carcinoma cells from ectopic and spontaneous tumors in response to growth
factor diffusion from a Matrigel-based solution within a hollow-bore needle [16] has
confirmed chemotaxis towards a variety of stimuli [14,17,18]. Several important points arise
from these studies, including the demonstration of a saturable chemotactic response (thus
setting boundary conditions on relevant in vitro studies), differential ligand requirement
based upon receptor expression (lending insight into in vivo response across clinical
subtypes), and demonstration of relative independence of biomechanical properties of the
tumor (as the peak invasion across studies is roughly equivalent, but the position of
collection is variable and does not account for tumor mechanics). These insights must be
considered when designing experiments informative for model construction.

Close Encounters: Presentation of Chemoattractants in Tumors

Although it is clear that the mechanical properties of the tumor microenvironment play a
role in tumor progression, we will focus here on cell-secreted ligands whose gradients could
be influenced by their binding of extracellular matrix components. The most commonly
studied in carcinoma are the EGF receptor family ligands (EGF, TGFα, HB-EGF,
amphiregulin, NRG1), the angiogenic factor VEGF, the macrophage-motility factor CSF-1,
the ‘scatter factor’ HGF, the antiproliferative factor TGFβ, and the cytokines CCL19,
CCL21, and CXCL12 or SDF-1. Several methods of ligand presentation have been utilized
to study chemotaxis of tumor cells in vitro, with Boyden or transwell chambers being the
most common end-point assays [19]. Despite providing a convenient screening approach for
evaluation of chemoattractant potency [20], there are several drawbacks to these assays,
predominantly gradient instability [21,22]. To overcome this limitation several groups have
developed microfluidic devices capable of generating stable gradients of chemoattractants
that mimic the nonlinearity of early time points found with diffusion-based approaches [23–
28]. These devices allow direct visualization of chemotaxis within well-defined gradients in
both 2D- and 3D-culture formats and in some cases can be utilized for high-throughput data
collection [29]. Alternatively, passive release of ligands from synthetic or natural hydrogels
can also be used to create chemoattractant gradients [30,31].

ECM-binding properties of cytokines and growth factors present an additional challenge for
re-creating and characterizing chemoattractant gradients in vitro. For example, the heparin-
associated matrix binding properties of CCL21, a chemokine that binds CCR-7, promotes
formation of autologous autocrine gradients that are amplified in the presence of interstitial
flow [3,32]. Creating appropriate CCL21 gradients in vitro is difficult due to the time
required to reach steady-state gradient formation [26,27]. To address this issue, Haessler et
al. created a microfluidic device incorporating a CCL21-containing agarose overlay. By
allowing steady-state gradient formation to occur across the agarose, injection of a cell-
containing ECM yields minimal system perturbation and leads to rapid gradient formation
across the area of interest (see Figure 1) [26]. Such a device could be employed along with
high-resolution imaging to study real-time dynamics of events during early stages of the
chemotactic response in 3D culture.

Active Release: Accounting for Extracellular Proteases

Several families of extracellular proteases, including matrix metalloproteinases (MMPs), A
Disintegrin and Metalloproteinases (ADAMs), serine proteases (such as matripase/
MTSP-1), and cysteine proteases (such as cathepsins) are secreted by tumor cells (see [33]
for a recent comprehensive review). Although their contributions to cancer progression and
metastasis are well appreciated, therapeutic targeting of proteases in cancer has been
unsuccessful [34]. The lack of systematic studies identifying and validating protease targets,
which is a direct consequence of the complexity of the network and the paucity of specific,
quantitative tools, likely underlies failure of protease-targeted drugs. In the context of gradient sensing, extracellular proteases -- often membrane-bound or secreted from tumor and stromal cells (Figure 1) -- can generate free ligand through cleavage of ECM-binding domains [35,36], enhance gradient formation through cleavage of ECM-binding ligands from the cell surface [37], and contribute to unmasking of cryptic ligand-like binding sites through direct cleavage of ECM proteins. Therefore, activities of extracellular proteases provide cues for gradient sensing, but have been largely neglected in data-driven relational models for cell migration.

Protease activity probes, based upon a decrease in fluorescence resonance energy transfer (FRET) upon substrate cleavage, have been employed to study protease activities in vitro [38,39]. However, identification of which protease is responsible for cleavage is problematic because of broad promiscuity for multiple substrates by many proteases [40] and substrate cross-recognition among several protease family members [38]. Miller et al. reported a multi-variate approach to deconvolving protease activity networks using combinatorial measurements coupled to a mathematical analysis permitting elucidation of specific protease activities [41]. The method involves construction of a matrix of kinetic rate parameters determined from purified enzymes with a library of FRET-based probes (see Figure 1). This parameter matrix is readily expandable to include additional proteases as desired depending on reagent availability. Although this method has to date been applied to two-dimensional culture systems, it can be extended to study of cells cultured in a range of ligand and matrix conditions -- thus permitting quantification of changes in protease activities during tumor progression.

**Signals**

Currently, most data-driven, relational models utilize population-level measurements to represent the signaling state of cells at various time-points after stimulation, often in the presence of perturbations such as RNAi or chemical inhibition [4–7,42]. Given the heterogeneity characteristic within tumors and tumor cell lines, this simplification may fail to capture signaling events occurring in the most dangerous cells, which may be rare within any given population. Development of more robust reporters and effective analysis tools [43] should facilitate building predictive models for cell gradient sensing that include quantitative measures of receptor distribution, protein localization, and spatiotemporal evolution of signaling states.

**Initiation of Signal Propagation: Distribution of Cell Surface Receptors**

Regardless of cue presentation, cells begin information processing at the cell membrane. In the case of most biochemical ligands (e.g., growth factors or cytokines), cell surface receptors initiate signaling flux upon ligand binding, so that the original spatial distribution of unligated receptor could influence initial gradient sensing. This may not be the case for cytokine sensing in mammalian cells as G-protein coupled receptors are uniformly distributed along the plasma membrane regardless of chemokine gradient [44]. Similarly, receptor tyrosine kinases (RTKs), such as the EGF receptor, are typically distributed evenly along the cell perimeter in vitro in the absence of signaling gradients [45]. However, despite a uniform total receptor distribution as visualized using epifluorescence microscopy techniques, bias towards increased receptor clustering at the cell periphery has been inferred using fluorescently-tagged receptor or single-particle tracking and TIRF microscopy [46,47]. Therefore, relevant receptor distribution studies might be those describing redistribution of receptors in the first minutes after exposure to a graded cue. The spatiotemporal parameters of receptor distribution, trafficking, and ligand binding have not yet been incorporated into relational models of cell migration. Nor have these parameters been combined with large proteomic or genomic datasets to obtain a systems view of how the biophysical presentation...
of cytokine and growth factor receptors directly impact specific cell responses. Increased accessibility of live-cell, high resolution imaging modalities should enable correlation of near-single molecule dynamic data to downstream signaling pathways.

**Choosing a direction: Receptor proximal signaling in single cells**

In the simplest sense, chemotaxis is driven through coordinated cycles of actin polymerization at sites of ligand-bound receptor and substrate adhesion [48]. Our knowledge of the molecules involved in the chemotactic response has increased significantly in the last decade, and a review of many of the known mechanisms underlying chemotaxis in cancer was recently published [1]. Here, we will concentrate on novel experimental and quantitative methods for studying the underlying signaling cascades.

An early step in the chemotactic response of mammary carcinoma cells is the activation of PLC\(\gamma\), which facilitates initial actin polymerization through a local increase in filament severing activity by PI(4,5)P2-associated proteins [49]. This activity of PLC\(\gamma\) is required for directional migration in response to a chemotactic cue [50,51]. Generally, phosphorylation of PLC\(\gamma\) is used as a proxy for hydrolysis of PI(4,5)P2, but van Rheenen et al. showed through FRET-based assays that local hydrolysis of PI(4,5)P2 leading to spatially-restricted cofilin-actin interaction sets the directional compass of carcinoma cells responding to gradients [52]. The amplitude of this response likely varies within heterotypic cultures, especially considering that cells from the same tumor can exhibit dramatically different responses to growth factor stimulation [53].

In lower organisms as well as tumor cells, a second wave of polymerization-competent actin barbed end generation occurs and is dependent on signaling through the PI3K pathway [49]. The activation of PI3K, and the generation of PI(3,4,5)P3, is recognized as a primal step in cell polarization. A recent series of reports from Haugh and colleagues demonstrates a technique to monitor motility parameters as a function of PI3K activity on an individual-cell basis [54–56]. The evolution of spatiotemporal PI3K-mediated PI(3,4,5)P3 generation was tracked using a fluorescently tagged Akt-PH domain in randomly migrating and chemotaxing cells. A custom algorithm revealed significant positive correlation between a motility parameter (the chemotactic index) and a PI3K-specific parameter (the signaling index) that is a measure of the directionality of signal strength based upon localization and fluorescence intensity (see Figure 1). This correlation was independent of growth factor gradient, suggesting that the location of PI3K signaling is indicative of directional migration regardless of a spatiotemporal cue context [54]. These results agree with the polarized role of PI3K in previous models of motility [57], but extend the insight to include mechanistic coupling between stochastic generation of signal and the subsequent migration directionality in mammalian cells [55].

Assessing signaling activities using ectopic expression of biosensors remains compromised by potential artifactual effects. Another, perhaps more generally reliable and accessible, approach to obtain information-rich, single-cell, spatially relevant data for relational models, is the employment of high-content imaging [58]. Two examples illustrate the usefulness of this approach for studying directed migration. First, the localization of active (phosphorylated) cofilin is inversely related with its ability to sever actin filaments and promote early membrane protrusion, rendering cell-averaged measures of phospho-cofilin uninformative about fast cell migration responses [59,60]. Second, chemotaxis is dependent on detachment of the cell rear from the underlying substrate [48], and cleavage of adhesion components by m-calpain is integral to this process [61,62]. Localization of m-calpain to PI(4,5)P2 at the cell membrane that regulates this activity, instead of phosphorylation level
of either the inhibitory or activating phospho-serine sites [63]. Therefore, localization of at least certain signaling proteins can produce increased insights concerning chemotaxis.

**Responses**

The final input variable for cue-signal-response relational models, and the most important for connection to physiological applications, is the cell phenotypic response(s) elicited by imposed signaling gradients. One consequence of building relational models that result in identification of important signals (or cues) is the issue of context dependency. When measuring motility phenotypes it is important to consider the relevance of experimental conditions to the physiological question that is being interrogated. Classic two-dimensional chemotaxis assays exhibit release profiles and kinetics that are predictable based upon diffusive transport phenomenon [64,65], which usually are inappropriate for tumor cell chemotaxis due to fast dissipation of the imposed gradient (min) in comparison to the relatively slow rate of tumor cell migration *in vitro* (hours). Thus we raise here recently developed methods for quantifying directed motility parameters arising from acute cell response to chemokine and long-term chemotaxis.

**Measuring the output: New methods to quantify gradient sensing phenotype**

Among the first quantifiable cell responses to a chemotactic cue is formation of a lamellipodial protrusion. Membrane protrusion response to paracrine factors occurs *in vivo* (see Supplemental Movie 3 from [15]) and has been examined in three-dimensional culture following perturbation in adhesion-related proteins [66]. Unbiased characterization of lamellipodial dynamics in response to growth factors and cytokines has proven challenging, especially in unmodified cells (those not expressing a fluorescently tagged protein), due to the heterogeneous shape of the extending cell protrusion in most cell types. The manual nature of these analyses has limited their applicability in correlative studies due to small sample size. Recently, morphodynamic profiling [67] (see Figure 1) approaches for quantification of leading edge dynamics in resting cells and in response to various perturbations have provided significant insight to the roles of several proteins important in the chemotactic response [59]. Thus far only the localized dynamics of signaling proteins occurring near the leading edge have been analyzed with respect to the detailed membrane activity maps [43,59,67–70]. These data could also be analyzed in combination with network-wide signaling measurements and measurements of protein secretion and proteolytic activity (on a population-level as a first pass) to construct network maps informative of changes in lamellipodial dynamics. We have observed that the initial membrane protrusion response across a panel of triple negative breast cancer cell lines is highly predictive of subsequent 3D migration in extracellular matrix (Meyer, A.S., Hughes-Alford, S.K., Lauffenburger, D.A., unpublished work). Therefore, insight gained from using morphodynamic profiling techniques applied to cells in a gradient of chemoattractant might provide predictive power in regards to paracrine and autocrine sensing in 3D culture.

It is attractive to hypothesize that the intracellular signaling pathways underlying random and directed migration are similar [54,71], and that migration mode is governed by signal initiation that could result from cue gradients, receptor presentation [46] or stochastic variability [55,57,71]. Indeed, there is substantial evidence for a set of common pathway participants, although the utilization of individual components may differ among cell types. Gruver et al. recently proposed a simple analysis technique based upon distinct changes in cellular morphology that characterize cell motility, random and directed, using a single scaling law derived from bimodal analysis [72]. Migration can be parsed into the time spent moving in a persistent manner versus that spent reorienting the cell polarity. When
compared, these characteristic times are inversely correlated and fit an exponentially decreasing curve. The authors show that the model is robust to abrupt changes in cell direction, which offers an advance over the commonly used ‘persistence time’ parameter found by fitting migration paths to a persistent random-walk model [73].

Conclusions and Future Perspectives

Data that can be generated with single-cell experimental methods are optimal for estimation of molecular-level mechanistic process parameters. However, the realization of fully mechanistic network models encompassing multiple processes and their regulatory pathways remains a future hope. Accordingly, our notion of integrating these kinds of data into relational multi-variate network models can begin to bridge the spatiotemporal knowledge gap that currently exists in the era of high-throughput genomic and proteomic analysis. How does one incorporate single cell measurements with population level data to make predictions about cell phenotype? One advantage of single-cell measurements is the realization of a distribution of response. It is then possible to use measures of error, or distribution width, to characterize the heterogeneity within a specific signaling or response measurement. These data may allow identification of ‘rare cell’ populations, such as those that respond in the bottom or top 5% of the distribution, and might represent the most aggressive tumor cells.

Although the gradient sensing mechanisms in chemotaxis are productively conceptualized in a ‘cue-signal-response’ framework, in reality the process for building data-driven, predictive models is iterative. For example, challenges in measuring the appropriate ‘responses’ that arise from well-planned ‘cues’ mandate preliminary experiments to determine what experimental conditions should be used when measuring the signaling dataset. The most useful signaling data may not come from experiments that are exactly the same as those used to measure the phenotypic response. For instance, the similarity between the leading edge of a cell in chemotaxis versus chemokinetic stimuli suggests that measurement of the signaling network under the latter conditions is likely be to equally as informative and more tractable on a large scale. Certainly, careful consideration of trade-offs between experimental consistency and physiological relevance will aid in most effective construction of data-driven multi-variate relational models for tumor cell chemotaxis.

Acknowledgments

We thank Aaron Meyer for critical reading and suggestions. Special thanks to Dr. Andrew Darling and Professor Mingming Wu for figure contribution. Funding was provided by DOD BCRP fellowship BC087781 to SKA, and by National Institutes of Health grants U54-CA112967 and R01-GM081336 to DAL.

References


Curr Opin Cell Biol. Author manuscript; available in PMC 2013 April 1.
41. Miller MA, Barkal L, Jeng K, Herrlich A, Moss M, Griffith LG, Lauffenburger DA. Proteolytic Activity Matrix Analysis (PrAMA) for simultaneous determination of multiple protease activities. Integr Biol (Camb). 2010; 3:422–438. This study presents the first systems approach to
accurately deconvolve and quantify protease activity within mixtures of MMPs and ADAMs. [PubMed: 21807711]


54. Melvin AT, Welf ES, Wang Y, Irvine DJ, Haugh JM. In chemotaxing fibroblasts, both high-fidelity and weakly biased cell movements track the localization of PI3K signaling. Biophys J. 2011; 100:1893–1901. References 54, 55, and 56 track the development of Signaling Vector Analysis, a tool that combines TIRF-based imaging of PIP3 binding probes and long-term chemotaxis or random motility measurements. [PubMed: 21504725]


63* Leloup L, Shao H, Bae YH, Deasy B, Stolz D, Roy P, Wells A. m-Calpain activation is regulated by its membrane localization and by its binding to phosphatidylinositol 4,5-bisphosphate. J Biol Chem. 2010; 285:33549–33566. This elegant study provides a mechanism for tail retraction in migrating cells and provides another example of the importance of polarized phosphoinositide signaling within a chemotaxing cell. [PubMed: 20729206]


75. Haessler U, Kalinin Y, Swartz MA, Wu M. An agarose-based microfluidic platform with a
19343497]
Figure 1. New experimental approaches for populating data-driven relational models of chemotaxis in cancer

Growth factor gradients (circles) that arise from paracrine signaling between tumor-associated macrophages (yellow) and breast carcinoma cells (blue) represent one example of chemotaxis within tumors [14]. Relational models for probing the mechanisms underlying tumor cell chemotaxis can be conceptualized using a ‘cue-signal-response’ paradigm. [1] Utilization of recently developed, highly controllable 3D microfluidic platforms can provide physiologically relevant presentation of chemoattractant cues. [75]. Device diagram used with kind permission from Dr. Andrew Darling. [2] Quantification of protease activities utilizing a systems biology approach provides mechanism underlying gradient generation and cue presentation. A matrix of kinetic rate parameters measured with purified enzymes can be utilized to estimate specific protease activity in biological samples. Figure courtesy of M. Miller and D. Lauffenburger [41]. [3] New single cell analyses, including Signaling Vector Analysis [54,55], provide novel signaling metrics, such as the magnitude and direction of Akt-PH domain localization (pseudo-colored such that red is highest intensity). The resulting time-dependent vectors, inset arrows, can be correlated with input cues and phenotypic response. Figure used with kind permission from [55]. [4] Quantification of phenotypic response ultimately determines the applicability of any data-driven, relational model. Automated analyses, such as morphodynamic profiling [67], may provide unbiased...
quantitative data about acute cell response upon exposure to chemoattractant gradients. The metrics derived from analysis of edge dynamics can be correlated to cues, such as protease activity (chemoattractant release), or signals, such as growth factor receptor distribution (See Signals sub-section), to generate hypotheses through application of relational modeling. Figure reproduced with kind permission from [67].