# Invasion of homogeneous and polyploid populations in nutrient-limiting environments

**Gregory J. Kimmel[1], Mark Dane[2], Laura Heiser[2,1], Philipp M. Altrock[1] and Noemi Andor[1*]**

*[1]Department of Integrated Mathematical Oncology, Moffitt Cancer Center, Tampa, FL, USA [2]Department of Biomedical Engineering, OHSU Center for Spatial Systems Biomedicine, Knight Cancer Institute, Oregon Health & Sciences University, Portland, OR, USA*

## Financial support

This work was supported by the National Cancer Institute R00CA215256 awarded to NA. PMA acknowledges support through the National Cancer Institute, part of the National Institutes of Health (NIH), under grant number P30-CA076292. LH acknowledges support through NIH research grants 1U54CA209988, U54-HG008100, Jayne Koskinas Ted Giovanis Foundation for Health and Policy, and Breast Cancer Research Foundation.

## Conflict of Interest

The authors declare no potential conflicts of interest.

## Keywords

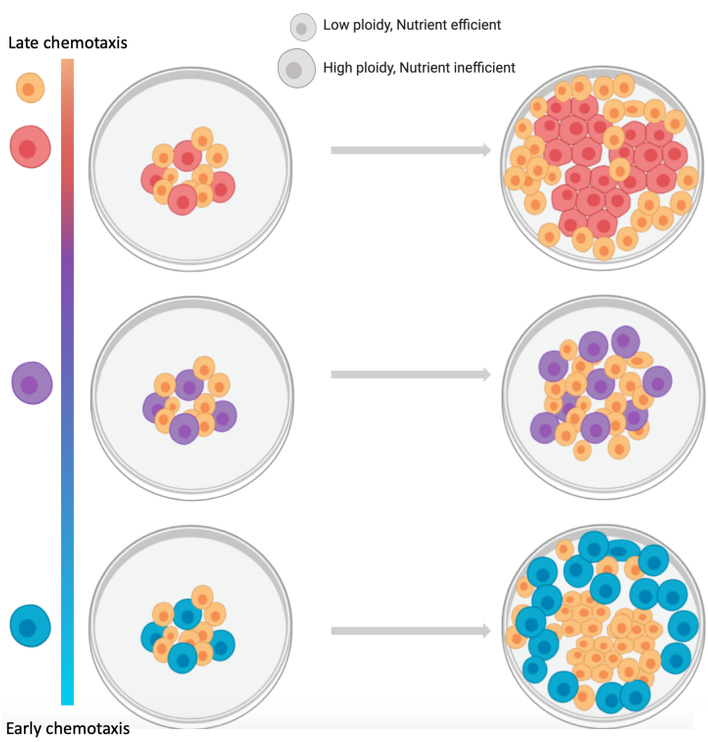
Evolutionary tradeoffs, Aneuploidy, Polyploidy, Breast cancer, Mathematical Modeling, Whole genome doubling, Partial differential equations, Traveling waves

[*]Noemi Andor; Moffitt Cancer Center, SRB 23224 B2 3011 Holly Drive, 33612 Tampa; noemi.andor@moffitt.org

**Abstract**

Breast cancer progresses in a multistep process from primary tumor growth and stroma invasion to metastasis. Progression is accompanied by a switch to an invasive cell phenotype. Nutrient-limiting environments promote chemotaxis with aggressive morphologies characteristic of invasion. It is unknown how co-existing cells differ in their response to nutrient limitations and how this impacts invasion of the metapopulation as a whole. We integrate mathematical modeling with microenvironmental perturbation-data to investigate invasion in nutrient-limiting environments inhabited by one or two cancer cell subpopulations. Hereby, subpopulations are defined by their energy efficiency and chemotactic ability. We estimate the invasion-distance traveled by a homogeneous population. For heterogeneous populations, our results suggest that an imbalance between nutrient efficacy and chemotactic superiority accelerates invasion. Such imbalance will spatially segregate the two populations and only one type will dominate at the invasion front. Only if these two phenotypes are balanced do the two subpopulations compete for the same space, which decelerates invasion. We investigate ploidy as a candidate biomarker of this phenotypic heterogeneity to discern circumstances when inhibiting chemotaxis amplifies internal competition and decelerates tumor progression, from circumstances that render clinical consequences of chemotactic inhibition unfavorable.

**Significance:** A better understanding of the nature of the double-edged sword of high ploidy is a prerequisite to personalize combination-therapies with cytotoxic drugs and inhibitors of signal transduction pathways such as MTOR-Is.




# Introduction

Invasion and infiltration are hallmarks of advanced cancers, including breast cancer, and accumulating evidence suggests that invasive subclones arise early during tumor evolution [1].

Infiltrating and invasive phenotypes are often observed among high-ploidy cells. Converging evidence from different cancer types, including colorectal-, breast-, lung- and brain cancers, suggests a strong enrichment of high ploidy cells among metastatic lesions as compared to the primary tumor [2, 3]. Even in normal development: trophoblast giant cells - the first cell type to terminally differentiate during embryogenesis - are responsible for invading the placenta and these cells often have hundreds of copies of the genome [4]. Coexistence of cancer cells at opposite extremes of the ploidy spectrum occurs frequently in cancer and is often caused by whole genome doubling (WGD). Similar to infiltration, the timing of WGD is early in tumor progression across several cancer types [5, 6], including breast cancer. Tetraploid cells resulting from WGD often lose regions of the genome, giving rise to poly-aneuploid cancer cells (PACCs). Multiple studies have described a minority population of PACCS with an unusual resilience to stress [7–9]. A very recent investigation of evolutionary selection pressures for WGD suggests that it mitigates the accumulation of deleterious somatic alterations [10]. However, it is not clear what costs cells with a duplicated genome pay for this robustness.

To address this question, we developed a mathematical model of co-evolving high- and low-ploidy clones under various energetic contingencies. We calibrate the model to recapitulate doubling times and spatial growth patterns measured for the HCC1954 ductal breast carcinoma cell line via MEMA profiling [11]. This includes exposure of HCC1954 cells to HGF in combination with 48 extracellular matrices (ECMs), followed by multi-color imaging [12]. Our results show that long-term coexistence of low- and high-ploidy clones occurs when sensitivity of the latter to energy scarcity is well-correlated to their chemotactic ability to populate new terrain. Higher energy uniformity throughout population expansion steers selection in favor of the low-ploidy clone, by minimizing the fitness gain the high-ploidy clone gets from its chemotactic superiority. Better understanding of how these two phenotypes co-evolve is necessary to develop therapeutic strategies that suppress slowly-proliferating, invasive cells before cytotoxic therapy favors them.

# Materials and Methods

We first introduce the conceptual framework that lead to the model, formulate the model equations and then we derive analytical and numerical solutions. Finally we describe drug-sensitivity and RNA sequencing data analysis of cell lines with different ploidies.

## Overall model design

The use of partial differential equations (PDEs) over a stochastic-based approach such as agent-based modeling permits us to make predictions based on analytical results derived from the subsequent PDEs and an increase in computational efficiency.

We modeled growth dynamics in polyploid populations of various subpopulation compositions. Appealing to a continuity description and assuming a continuum approximation of the cellular and energy concentration is valid, we derived a system of coupled PDEs. Each compartment in the PDE describes the spatio-temporal dynamics of the quantity of interest (e.g. energy or cellular dynamics). At the core of our model lies the assumption that chemotactic response to an energy gradient is a function of the cell's energetic needs. This trade-off implies that heterogeneous populations will segregate spatially, with higher energy-demanding cells leading the front of tumor growth and invasion. In contrast, for an energy-rich environment we expect the cells to grow in a similar way as they will have no need to search for places of higher energy density.

We model competition for energy in a heterogeneous population, consisting of goer and grower subpopulations, to predict their behavior during plentiful and energy sparse conditions. We assume both goer and grower have the same random cell motility coefficient $\Gamma$, the same chemotactic coefficient $\chi$ and maximal growth rate $\lambda$. Sensitivity to the available energy is modeled via a Michaelis-Menten type equation with coefficients $\Phi_i$ that determine the amount of energy that population $i$ needs for a half-maximal growth rate. Chemotactic motion is asymmetrically sensitive to the amount of energy available. This is accounted for by $\Xi_i$. The goers ($U$) are more motile and require more energy compared to the growers ($V$). This manifests itself mathematically via the parameter relations $\Phi_U > \Phi_V$ and $\Xi_U < \Xi_V$.

Quantitative estimates of how a cell's growth rate and motility depends on energy availability have been described [13–15]. Energetic resources come in various forms, and the identities of the limiting resources that ultimately drive a cell's fate decision vary in space and time. We used MEMA profiling to investigate what likely is only a narrow range of that variability – 48 HGF-exposed ECMs [12]. HGF stimulates both growth and migration of epithelial and endothelial cells in a dose-dependent manner, whereby maximal growth-stimulating effects have been reported at concentrations twice as high as concentrations that maximize migration [16]. In line with these reports, our model demonstrates a shift from proliferation to migration as resources get depleted.

Mathematical models of a dichotomy between proliferation and migration are numerous [17–19], but whether the two phenotypes are indeed mutually exclusive remains controversial [20]. Our efforts to use mathematical modeling to inform what cost high-ploidy cells (goers) pay for their robustness builds upon these prior works. We extend it by accounting for differences in the rate at which cells consume energy ($\delta$) and differences between media in the rate at which energy diffuses (coefficient $\Gamma_E$). For mid-range energy diffusion coefficients our model describes directed cell motility in response to a gradient of a soluble attractant, i.e. chemotaxis. By contrast, small values of $\Gamma_E$ approximate cell motility towards insoluble attractants, i.e. haptotaxis. As such, the chosen value for $\Gamma_E$ sets where along the continuum between haptotaxis and chemotaxis directed cell movement resides. A special case applies when the energy diffusion coefficient is very large relative to cell movement, in which case neither chemotaxis nor haptotaxis occurs. All these energetic contingencies determine whether phenotypic differences between goers and growers manifest as such, and explain why non-proliferative arrested cells can have the same motility as cycling cells [20].

## Quick guide to equations

Our model assumes that the energy diffusion coefficient $\Gamma_E$ depends on the type of media or surface upon which the cells grow. We also suppose that energy is consumed in proportion to the amount of cells present. For cell motility, we assume it is driven both by random cell motion and chemotaxis. This leads to our general coupled system:

$$\frac{\partial \tilde{E}}{\partial t} = \underbrace{\Gamma_E \nabla^2 \tilde{E}}_{\text{Nutrient diffusion}} - \underbrace{\delta \tilde{E}(U+V)}_{\text{Nutrient consumption}}, \tag{1a}$$

$$\frac{\partial U}{\partial t} = \underbrace{\Gamma \nabla^2 U}_{\text{Random motion}} \underbrace{- \chi \nabla \cdot \frac{U \nabla \tilde{E}}{\Xi_U + \tilde{E}}}_{\text{Chemotaxis}} + \underbrace{\lambda \frac{\tilde{E}}{\Phi_U + \tilde{E}}}_{\text{Energy-dep. growth}} U \underbrace{\left(1 - \frac{U+V}{K}\right)}_{\text{Contact inhibition}}, \tag{1b}$$

$$\frac{\partial V}{\partial t} = \underbrace{\Gamma \nabla^2 U}_{\text{Random motion}} \underbrace{- \chi \nabla \cdot \frac{V \nabla \tilde{E}}{\Xi_V + \tilde{E}}}_{\text{Chemotaxis}} + \underbrace{\lambda \frac{\tilde{E}}{\Phi_V + \tilde{E}}}_{\text{Energy-dep. growth}} V \underbrace{\left(1 - \frac{U+V}{K}\right)}_{\text{Contact inhibition}}. \tag{1c}$$

The system (1) is defined on a dish of radius $R$ subject to no-flux boundary conditions. We assume the goers ($U$) and growers ($V$) are initially concentrated at the center with radius $\rho_0 \leq R$ and initial concentrations $U_0, V_0$. Further, we assume that the energy density is uniformly distributed on the plate with initial value $E_0$. All parameters except otherwise stated are independent of any of the state variables $(\tilde{E}, U, V)$. The energy is consumed at rate $\delta$. Both cells can divide at the maximal rate $\lambda$, but are restricted by the energy density $\tilde{E}$. The cells can locally grow to a local maximal density given by $K$. This parameter is often cell line-dependent and is related to contact inhibition and a cell's ability to grow on top of each other.

We now convert our system to dimensionless form that was used for all subsequent simulations and analysis via appropriate re-scaling (Supplementary methods),

$$\frac{\partial E}{\partial \tau} = \gamma_E \nabla^2 E - aE(u+v), \tag{2a}$$

$$\frac{\partial u}{\partial \tau} = \nabla \cdot \left( b \nabla u - u \frac{\nabla E}{\xi_U + E} \right) + \frac{E}{\phi_U + E} u(1-u-v), \tag{2b}$$

$$\frac{\partial v}{\partial \tau} = \nabla \cdot \left( b \nabla v - v \frac{\nabla E}{\xi_V + E} \right) + \frac{E}{\phi_V + E} v(1-u-v), \tag{2c}$$

Hereby re-scaling simplified the system from thirteen (ten parameters and three initial conditions) to nine (seven parameters and two initial conditions), with dimensionless variables $\gamma_E = \Gamma_E/\chi,\ a = \delta K/\lambda,\ b = \Gamma/\chi,\ \phi = \Phi/E_0$ and $\xi = \Xi/E_0$.

## Analytical estimates of infiltration

The desire for cells to move is inherently tied to the availability of nutrients and space. To this end we define $\Psi(\tau) := [\rho(\tau) - \rho_0]/\rho_0$

where $\rho(0) = \rho_0$, the initial radius of cell seeding density. $\Psi$ can be thought of as a non-dimensional measure of infiltration attained after time $\tau$. This dimensionless measure has the added benefit of being scale-independent. An inherent difficulty with random cell motility and calculating infiltration is that the system always reaches the boundary of the dish in finite time. Instead we will define the maximum degree of infiltration to be given by the time needed for the total energy to be below a threshold $\varepsilon \ll 1$.

In general, the maximum degree of infiltration is difficult to predict analytically, so we only considered the single subpopulation case when obtaining our analytical estimates. We also made use of the simplification that most energy-type molecules (e.g. glucose) have a diffusion coefficient that is very large, relative to cell movement. This allows us to write a reduced model which has energy homogeneous in space (see (S3a)-(S3b) in Supplementary Methods). We estimate infiltration analytically through two different approaches. First, we extract an ODE system that couples nutrient consumption to the wave front location. Second, we derive estimates of the infiltration achieved by using the cell concentration after it has become uniform.

**No chemotaxis infiltration estimate**

We can derive an estimate for infiltration in the absence of chemotaxis by appealing to (S3a)-(S3b). These logistic growth-reaction-diffusion models often exhibit complex dynamics. One such example is that of a traveling-wave solution, where one state, typically the stable state, travels (infiltrates) through the domain. The canonical example of this phenomenon is the Fisher-KPP. In contrast to the Fisher-KPP and other traveling-wave problems typically studied, our model has a decaying growth rate and so the magnitude of the non-linearity that caused the traveling wave is tending to zero. Therefore, in the classic sense, our system does not admit a traveling wave. We here extend the theory by assuming a separation of time scales between consumption of energy (e.g. decay of energy-dependent growth rate) and the speed of the traveling wave.

To begin, we make the assumptions that the wave speed is a slow function of $r$ and $\tau$. The solution obtained will verify that these assumptions are valid for our system. Our ansatz takes the form $u(r,\tau) = U(r - \eta\tau) = U(z)$. Note that in spatial equilibrium, $u = 1$ is stable and $u = 0$ is an unstable steady state. If the unstable state is what governs the wave speed, then the wave is said to be "pulled", otherwise it is "pushed" [21, 22]. The resulting analysis yields a coupled system of ODEs that govern the speed of the front (Supplementary methods):

$$\frac{dE}{d\tau} = -a\left(\frac{\rho}{R}\right)^2 E, \tag{3a}$$

$$\frac{d\rho}{d\tau} = 2\sqrt{\frac{bE}{\phi + E}} - \frac{b(n-1)}{\rho}. \tag{3b}$$

We see that our assumptions on the behavior of the traveling wave are verified, since $\rho$ is assumed much larger than 0 and $0 \ll \rho \ll R$. This shows that $E$ is a slowly varying function of time and $\eta = d\rho/d\tau$ is a slowly varying function of time and its current distance from the center. In other words, a traveling wave will only form when the initial seeding radius is relatively large ($\rho_0 \sim 1$ was sufficient in most simulations).

**Estimating the degree of infiltration from equilibration**

The previous section yields a system that can be integrated to track the evolution of the cell front over time. However, we may be more interested in how far it will ultimately travel before energy exhaustion and not the speed at which it gets there. An interesting alternative to tracking the wave over time is to only assume it travels as a wave, but only record the density after the system has reached uniformity. This is possible if the death rate (which has been neglected) is much smaller than the time it would take the cells to spread uniformly. If this is the case, we can bound the degree of infiltration from only knowing the uniform value $\bar{u}$ at the end of the experiment (see Supplementary Methods for details):

$$\rho(\tau = T) = R\sqrt{\frac{2\bar{u}}{1+\Lambda^2}}, \tag{4}$$

where $\Lambda$ is the transition width, i.e. the length scale on which cell concentration goes from $u = 1$ to $u = 0$, and $\bar{u}$ it is the concentration at the end of the experiment (at equilibration). Solving (4) gives us the estimated infiltration as:

$$\Psi = \frac{R}{\rho_0}\sqrt{\frac{2\bar{u}}{1+\Lambda^2}} - 1. \tag{5}$$

Since we assume that the transition width is unknown, we can bound $\Psi$ (or $\rho(T)$) by considering the lower and upper bounds $\Lambda = 0, 1$, respectively.

## Numerical estimates of infiltration

When directed cell movement is not negligible, analytical approximations are more difficult to obtain, and numerical simulations are preferred to estimate the degree of infiltration. For simulations where we measured infiltration, we took the 1-norm $\|E\|_1 = \sum_i |E_i| < \varepsilon = 1e-4$ as the threshold defining minimum energy requirements. We calibrate the model to recapitulate doubling times and spatial growth patterns measured for the HCC1954 ductal breast carcinoma cell line via MEMA profiling. The dataset includes exposure of HCC1954 cells to HGF in combination with 48 ECMs in a DMSO environment (i.e. no drug was added to the media). Between 13 and 30 replicates of each ECM are printed on a rectangular MEMA array as circular spots (Supplementary Fig. 1A), adding up to a total of 692 spots [12].

**MEMA data analysis**

An average of 31 cells (90% confidence interval: 23-41) were seeded on each 350 $\mu$m spot and grown for three days. The median number of cells imaged at day three was 121 – which falls within the range expected from the estimated seeding density and a doubling time of approximately 43.81 hours reported for this cell line (72-128 cells). Confluence at seeding was calculated from the ratio between the cumulative area of cells and the area of the spot (see Supplementary Fig. 1B-C).

Quantification of segmented, multi-color imaging data obtained for each spot three days post-seeding was downloaded from Synapse (synID: syn9612057; plateID: LI8C00243; well: B03). Cells tended to grow in a symmetric, toroidal shape (Supplementary Fig. 1D), albeit considerable variability was observed across the ECMs. We binned cells detected within a given spot according to their distance to the center of the spot and calculated the confluence of each bin (Supplementary Methods). This was then compared to the confluence obtained from the simulations as described below.

### Simulation environment

Simulations were ran to recapitulate logistics and initial growth conditions on the MEMA array (Supplementary Fig. 1A): the spatial domain was circular with radius $R = 1750$ $\mu$m and the temporal domain was exactly three days. Cells were seeded uniformly at the center of this domain along a radius $\rho_0 = 175$ $\mu$m at 36% confluence. To recapitulate the configuration of the MEMA profiling experiment, cells leaving the $\rho_0$ domain can no longer adhere to the ECM and die. This was implemented by having the carrying capacity $K(x)$ rapidly approach zero when $x > \rho_0$. This setup can have energy attract cells to the periphery of a MEMA spot and beyond. We ran 161,000 simulations at variable energy consumption rates, chemotactic/haptotactic coefficients, energetic sensitivities, and diffusion rates of the growth-limiting resource (i.e. ECM-bound HGF; Table 1). 92% of these simulations completed successfully. For each simulation/ECM pair, we compared spatial distributions of in-silico and in-vitro confluence using the Wasserstein metric [23].

The model was implemented in C++ (standard C++11). The armadillo package (ARMA version: 9.860.1) [24] was used for simulation of the PDEs. Simulations were run on a Intel Core i7 MacBook Pro, 2.6 GHz, 32 GB RAM. The source code is available at the github repository for the Integrated Mathematical Oncology department: GoOrGrow.

## Ploidy as biomarker of phenotypic divergence

We identified 44 breast cancer cell lines of known ploidy [25] and with available RNA-seq data in CCLE [26] and analyzed their drug sensitivity- and expression profiles as follows.

### Drug sensitivity analysis

We used Growth rate inhibition (GR) metrics as proxies of differences in drug sensitivities between cell lines. Unlike traditional drug sensitivity metrics, like the $IC_{50}$, GR curves account for unequal division rates, arising from biological variation or variable culture conditions – a major confounding factor of drug response [27]. Previously calculated GR curves and metrics were available for 41/44 breast cancer cell lines. A total of 46 drugs had been profiled on at least 80% of these cell lines and their $GR_{AOC}$ drug response metric [28] was downloaded from GRbrowser. For each drug we calculated the z-score of $GR_{AOC}$ across cell lines in order to compare drugs administered at different dose ranges. Of these 46 drugs, 39 could be broadly classified into two categories as either cytotoxic (25 drugs) or inhibitors of signaling pathways (14 drugs) (Supplementary Table 2). We then evaluated a cell line's ploidy as a predictor of its $GR_{AOC}$ value using a linear regression model. Since molecular subtype of breast cancer cell lines is known to influence drug sensitivity we performed a multivariate analysis, including the molecular subtype as well as an interaction term between ploidy and drug category into the model.

### RNA-Seq analysis

The molecular subtype classification of all cell lines was available from prior studies [29–32]. Of these 44 cases, four were suspension cell lines and excluded from further analysis. Of the remaining 40 cell lines, 20 originated from primary breast cancer tumors and were the focus of our analysis. Gene expression data was downloaded from CCLE. We used gene set variation analysis (GSVA) to model variation in pathway activity across cell lines [33]. Pathways for which less than ten gene members were expressed in a given cell lines were not quantified. The gene membership of 1,417 pathways was downloaded from the REACTOME database [34] (v63) for this purpose.

# Results

## High-ploidy breast cancer cell lines have increased metabolic activity and cell motility

To better understand the phenotypic profile of high-ploidy cells, we compared the ploidy of 41 breast cancer cell lines with their response to 46 drugs. As drug response metric, we used the integrated effect of the drug across a range of concentrations estimated from the 'area over the curve' ($GR_{AOC}$) [27, 28]. We observed that cytotoxic drugs and drugs inhibiting signal transduction pathways were at opposite ends of the spectrum (Fig. 1A). Namely, ploidy was negatively correlated with the $GR_{AOC}$ for several cytotoxic drugs and positively correlated with the $GR_{AOC}$ of various mTOR inhibitors, suggesting high ploidy breast cancer cell lines tend to be resistant to DNA damaging agents, while sensitive to drugs targeting nutrient sensing and motility.

We built a multivariate regression model of drug sensitivities to test the hypothesis that the relationship between ploidy and $GR_{AOC}$ was different for cytotoxic drugs than for inhibitors of cell signaling pathways. Molecular subtype alone (Fig. 1B), could explain 0.4% of the variability in $GR_{AOC}$ z-scores across cell lines (adjusted R-square = 0.0044; p = 0.026). Including ploidy into the model did not improve its predictive accuracy (adjusted R-square = 0.0037; p = 0.058). However, an interaction term between ploidy and drug category (cytotoxic: 27 drugs vs. signaling: 16 drugs) increased accuracy to explain 2.6% of variability in drug sensitivity across cell lines (adjusted R-square = 0.026; p < 1e-5; Fig. 1C). The same improvement from an interaction term between ploidy and drug category was observed in an independent dataset of maximal inhibitory concentration ($IC_{50}$) values of 34 cytotoxic drugs and 51 signaling inhibitors obtained from the GDSC (Genomics of Drug Sensitivity in Cancer) database [35] (Supplementary Fig. 3).

We then focused on a subset of aforementioned 41 cell lines, namely those that had been established from primary breast cancer

tumors as adherent cells (20 cell lines; Fig. 1D) and we quantified their pathway activity (see Methods). A total of 27 pathways were correlated to ploidy at a significant p-value (| Pearson r | $\geq$ 0.44; p $\leq$ 0.05; Supplementary Table 1). The strongest correlations were observed for metabolic pathways such as hyaluronan metabolism and metabolism of vitamins (Fig. 1E-F). Hyaluronic acid is a main component of the ECM and its synthesis has been shown to associate with cell migration [36, 37].

These results support a model that connects high ploidy with both, the chemotactic ability and metabolic energy deficit of a cell.

## Infiltration of homogeneous populations

Model design was guided by the goal to describe growth dynamics along two axes: from random to directed cell motility and from homogeneous to heterogeneous cell compositions (Methods equations (2a) - (2c)). While the last section will step into the second axis, the following two subsections distinguish scenarios along the first axis: (i) "homogeneous nutrient environments" are environments in which random cell motility dominates throughout population growth; (ii) "heterogeneous nutrient environments" imply the formation of gradients which cause cells to move in a directed fashion, as is the case during cellular growth on an ECM.

### Homogeneous nutrient environments

When the diffusion of the nutrient occurs on a much faster time scale than the actions taken by cells, we can assume that at the time scale of cells, the nutrient is essentially uniform in space. This simplification allows us to neglect chemotactic/haptotactic motion and consider only random cell motility as the driving force that spreads the cell density throughout the dish.

We obtain analytical estimates for the degree of infiltration in a homogeneous environment that lays the groundwork for new predictions. To arrive at (3a)-(3b), we employed the approximations that diffusion of energy molecules (e.g. glucose) is fast relative to cell movement and that the cells' movement through the dish can be approximated by a traveling wave [38]. These assumptions were verified by comparing the front estimates with results from the full numerical model (equations (2a) - (2c); Fig. 2A,B).

An interesting alternative to tracking the wave over time is to assume it travels as a wave, but only record the density after the system has reached uniformity. If the death rate is much smaller than the time needed for the cells to spread uniformly, we can bound the degree of infiltration that occurred from only knowing the uniform density of cells (eq. (4); Fig. 2C).

These analytical solutions point to scaling relationships for the speed of the moving front. For highly efficient energy-using cell lines ($\phi \ll 1$) (Fig. 2D), the front will evolve at a speed nearly independent of energy. In contrast, for large $\phi \gg 1$, the speed of the front falls off as $1/\sqrt{\phi}$. These predictions of the behavior of infiltration on parameters can be investigated experimentally.

### Heterogeneous nutrient environments

Assumptions made in the prior section apply to standard cell cultures of adhesive cells in a typical cell culture dish, where energetic resources diffuse so fast that gradients do not form. These assumptions break down during cellular growth on an ECM. Binding to the ECM can cause soluble factors (like HGF) to act and signal as solid-phase ligands [39, 40]. Proteolytic degradation of these ECMs then creates haptotactic gradients. Fig. 1 includes HCC1954 – a near-tetraploid breast cancer cell line whose growth on various ECMs has been measured via MEMA profiling. We analyzed HCC1954 and looked to determine if our mathematical model can explain its spatial growth patterns. MEMA profiling resulted in considerable variability of growth patterns across different ECMs (Fig. 3A,B).

We projected 2D spatial distributions measured on the MEMA array onto one dimension (Fig. 3C), rendering them comparable to those obtained from our simulations (Fig. 3D-E). For each simulation/ECM pair, we calculated the distance between in-silico and in-vitro spatial cell distributions using the Wasserstein metric (Fig. 3C,F) and ranked simulations by their minimum distance across ECMs. The top 2.3% simulation parameters were then stratified by the ECM whose spatial pattern they best resemble and compared to uniform prior parameter distributions (Fig. 3D,E). 75% of the ECMs accounted for only 1.55% of the top simulations. The tendency of cells on these ECMs to grow in a toroidal shape was strong, suggesting it may be the consequence of non-uniform printing of ECMs onto the array. We concluded that our model cannot explain the growth patterns on those ECMs well and focused our attention on the remaining 12 ECMs, represented by 3,429 simulations. We refer to the parameters of these simulations as inferred parameter space. Principal component analysis (PCA), Uniform Manifold Approximation and Projection (UMAP) [41] and density clustering [42] of the inferred parameter space revealed three clusters (Fig. 3G), with different ECMs segregating mainly into different clusters (Fig. 3H).

The two largest clusters differed mostly in their chemotactic/haptotactic- and energy diffusion coefficients; while the small cluster stood out by a high sensitivity to low energy and fast chemotactic/haptotactic response (Fig. 3I). Overall, all five model parameters showed significant differences between the three clusters, suggesting they all contribute to distinction between ECM growth patterns (Fig. 3I). This was further affirmed when looking at the % variance explained per principal component per parameter (Fig. 3J). To formalize parameter sensitivity analysis independent of ECMs, we also calculated the Sobol index [43] of each parameter. The Sobol Index quantifies how much of the variability in spatial cell concentration is explained by each parameter, while accounting for all its interaction effects. Each parameter contributed to significant variance (Sobol Index > 0.02; [43]) in at least one of three spatial statistics (Fig. 3K): skewness, confluence and gradient near the edge of the ECM.

We observed substantial differences in chemotactic/haptotactic coefficients and energy consumption rates between ECMs (Supplementary Fig. 4). To query the biological significance of this variability we quantified the expression of the 12 ECMs in the HCC1954 cell line (Methods). Two of the five inferred ECM-specific model

parameters were correlated with RNA-seq derived expression of the corresponding ECM: energy consumption rate (Pearson r = -0.657, p = 0.028) and sensitivity to low energy (Pearson r = 0.562, p = 0.071) – though latter fell short above significance (Supplementary Fig. 5).

In summary, the posterior distributions of model parameters represent a substantial departure from the uniform priors and could explain a significant proportion of growth conditions on the HGF-exposed MEMA array. This approach identified regions of interest in the parameter search space, allowing us to focus further simulations on biologically relevant chemotactic/haptotactic coefficients and energy diffusion rates.

### Infiltration of heterogeneous, chemotactic populations

Growth of cells in a given ECM environment was measured across 13-30 replicates on the MEMA platform. While our model – when calibrated to the corresponding ECM environment – could explain the observed growth pattern in the majority of these replicates, a substantial fraction could not be explained by fixed choices of sensitivity to low energy and directed cell motility (Supplementary Fig. 4). One possibility that may explain this is that HCC1954 is a heterogeneous cell line, with clones of variable phenotypes co-evolving. Representation of these clones among the 31 cells that were on an average sampled for each replicate may vary (Supplementary Fig. 2). This hypothesis is supported by a bimodal distribution of DNA content observed among replicating HCC1954 cells on individual ECM spots (Fig. 4A,B). If the HCC1954 population was homogeneous, we would expect a unimodal distribution of DAPI intensity among S-phase cells of this cell line. The observation of a bimodal distribution among S-phase cells suggest that HCC1954 is likely a polyploid cell line, i.e. clones of variable ploidies co-exist in this cell line.

To better understand the growth dynamics in a polyploid population, we used the two-subpopulation version of our model, whereby variable chemotactic abilities and energetic sensitivities of goer- and grower subpopulations compete with one another (equations (2a) - (2c)). We used fixed values for energy diffusion- and consumption rates as informed by model calibration (Fig. 3B) and varied sensitivity to low energy and chemotactic ability of both goer and grower, subject to equations (2a) - (2c) (Table 1). We initially used the same spatial and temporal domains as during model calibration, but concluded that the implied duration of the experiment (3 days) was too short for dynamics between the two populations to manifest. Each MEMA spot has a low capacity, whereby confluence is reached at no more than a few hundreds of cells. Such a small number of cells will not exhibit wave-like behavior and therefore will not suffice for spatial structure to emerge. We therefor extended temporal and spatial domains of our simulations, seeding cells at a lower confluence and letting them grow onto the entire energy domain until they consume all available energy (average of 127 days; Table 1).

We observed a non-monotonic relation between the goer's chemotactic ability and the speed with which the metapopulation invades the dish, with intermediate values being the least beneficial to its growth and spread (Fig. 4C). Temporal analysis of the simulations (Supplementary Data 1-3), revealed that if the goer's chemotactic motility is too high, it will leave the center of the dish too soon, leaving room for the grower to expand locally (Fig. 4D). By contrast, if the goer's motility is too low, it will miss the time-window of opportunity to ensure its dominance further away from the center of the dish while energy is still abundant. As a consequence, it will be outgrown by the grower at the edge of the dish once energy becomes sparse (Fig. 4E). Only when the goer has an intermediate motility, does the grower persistently coexist with it, both at the center and edge of the dish (Fig. 4F).

## Discussion

Models of infiltration are typically formulated under two critical assumptions. First, that energy production and consumption are non-uniform, leading to the formation of an energy gradient [44–46]; or second, that energy consumption is very slow compared to production, leading to an essentially infinite energetic resource [47]. Here we formulate a generalized model of infiltration when energy is finite and investigate its behavior along a spectrum of scenarios, from permanent energy uniformity to scenarios where this uniformity is gradually lost. The model derivation does not assume a particular dimension (e.g. 2D *in vitro* experiments vs. *in vivo* or 3D spheroid experiments). Many parameters that were valid in 2D will also extend naturally to 3D. For example, we would not expect a difference in consumption rates or half-maximal growth rates of the cells. However, energy diffusion ($\Gamma_E$) or random cell motion ($\Gamma$) will be higher due to the increased degree of freedom [48]. When energy is uniformly distributed at all times and the time scale for cell death is substantially longer than that of cell motility and birth, our results suggests that the degree of infiltration can be approximated using the cells' density at equilibration of movement and growth (Fig. 2C).

With an energy gradient that becomes steeper over time, our analytical approximations no longer hold, as directed cell movement becomes non-negligible. For this scenario, we leveraged MEMA profiling to inform regions of interest in the parameter search space. These regions of interest are relevant for cellular growth on a variety of HGF exposed ECM proteins. We observe correlations between inferred model parameters and RNA-seq derived signatures, even though the latter were not used during parameter inference. A potential explanation for the negative correlation between ECM-specific energy consumption and expression is that our model does not account for the possibility that cells can replace the ECM they degrade. The slower the rate of this replacement is, the higher the consumption rate appears to be. On the other hand, the more dependent cells are on a ECM for growth, the faster they must replace it, potentially explaining a positive correlation between ECM-specific expression and sensitivity to low energy.

When calibrating our model to a given ECM environment, growth patterns of a substantial fraction of replicates of that ECM could

not be explained by fixed choices of sensitivity to low energy (Supplementary Fig. 4). A potential explanation for this are variable cell compositions across experimental replicates. An alternative explanation is that this variability stems from artifacts that arise during non-uniform printing of ECMs onto the array–the so called ring effect. However, a bimodal distribution was also observed in the DNA content of replicating cells, which is not affected by potential printing artifacts. The second peak of this bimodal distribution was wider, consistent with the fact that high-ploidy cells with more DNA need longer to replicate.

The cell line HCC1954 is described as a hyper-tetraploid cell line with an average DNA-content of 4.2 [25]. However, this average value may be misleading, as suggested by stark variability in nuclei sizes (Fig. 4A). Despite a wealth of genomic information generated for this cell line [25], to the best of our knowledge no prior reports indicate whether or not the cell line is polyploid. We and others have found that high ploidy is an aneuploidy-tolerating state that accompanies intra-tumor heterogeneity *in vivo* and *in vitro* [5, 49, 50]. Our results suggest that HCC1954 is likely polyploid.

One event that could have led to this polyploid state is WGD. In contrast to cell lines, WGD events in primary tumors are mostly clonal, not subclonal [5, 6, 10] – clones carrying a doubled genome often sweep over the population, such that by the time the tumor is detected, the diploid ancestor no longer exists. A related scenario are advanced, therapy-exposed tumors shown to revert back to genomic stability, potentially bringing a WGD population back to a genomic state that more closely resembles its diploid ancestral state [51]. The model presented here can investigate how dynamics between the two subpopulations unfold in both of these scenarios – early, shortly after the WGD or late, after therapy exposure. This would characterize what circumstances prevent the WGD carrying clone from becoming dominant or from retaining its dominance and could help explain WGD incidence in primary and recurrent tumors.

If spatial and temporal domains were to be extended beyond the configuration of MEMA spots, our simulations predict that spatial segregation of two co-existing subpopulations according to their ploidy is a likely scenario and depends on the energy consumption rate. Our model can easily be extended to more than two subpopulations, for example to include a subpopulation of normal cells. For each additional cell type, a new compartment can be added to the model, with growth- and motion-related parameters that are specific to the corresponding cell type. However, the model currently does not incorporate mutations, i.e. the process of generating new clonal lines. A next step will be to extend our model to include mutation events, specifically chromosome mis-segregations that contribute extensively to diversify ploidy of a population [52, 53]. The additional DNA content of high-ploidy cells, though energetically costly, brings a masking effect against the deleterious consequences of chromosome losses [10]. This duality may explain the higher sensitivity to glycolysis inhibitors of high-ploidy cells and their lower sensitivity to cytotoxic drugs previously reported in Glioblastoma [54].

In line with prior reports we find that increased resistance of breast cancer cell lines to cytotoxic drugs is associated with high ploidy. In contrast, high ploidy breast cancer cell lines were sensitive to inhibitors of signal transduction pathways, including EGFR and especially MTOR signalling. A commonality among those pathways is their contribution to a cell's chemotactic response [55–57], suggesting opportunities to tune chemotaxis. MTOR inhibitors (mTOR-I), such as Rapamycin, significantly decrease migration of breast cancer cells in a dose-dependent manner [58, 59]. Rapamycin inhibits cell motility by a mechanism similar to that by which it inhibits cell proliferation [60], suggesting that the mTOR pathway lies at the intersection of a cell's decision between proliferation and migration. If high ploidy is indeed a characteristic specific to goer-like cells, then mTOR-Is are likely affecting this cell type (Fig. 1A,E,F), and could be used to inhibit its chemotactic response, thereby moving the population up the $x$-axis of Fig. 4C. Delaying chemotactic response of highly chemotactic cells could slow down invasion by maximizing competition within a polyploid population. If on the other hand chemotactic response of high ploidy cells is already at an intermediate level, our simulation suggest that further reduction may accelerate invasion of low ploidy cells. For such scenarios therapeutic strategies that include an mTOR-I may not be successful. Experiments will be needed to verify these in-silico results in-vitro. Knowing how co-existing clones with differential drug sensitivities segregate spatially can offer opportunities to administer these drug combinations more effectively.

## Acknowledgements

We thank the researchers in the Integrated Mathematical Oncology department at Moffitt Cancer Center for fruitful discussions and early feedback on parts of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the H. Lee Moffitt Cancer Center and Research Institute.

# Tables

| Experimentally measured parameters | Symbol | Homogeneous cells on any ECM | Heterogeneous cells | Unit |
|---|---|---|---|---|
| Maximal growth rate | $\lambda$ | 0.56 | 0.56 | $day^{-1}$ |
| Goer's seeding density | $u_0$ | 36.0 | 3.8 | % confluence |
| Grower's seeding density | $v_0$ | | 1.3 | % confluence |
| Simulation time | $t$ | 3 | 127 | days |
| Radius of domain | $\check{R}$ | 1750 | 1750 | µm |
| Seeding radius | $\rho_0$ | 175 | 175 | µm |
| Growth beyond $\rho_0$ | | no | yes | |

| Inferred parameters | Symbol | Homogeneous cells on GAP43 | Heterogeneous cells | Unit |
|---|---|---|---|---|
| Energy diffusion coefficient on ECM | $\Gamma_E$ | 9500 | 8982 | $\mu m^2\ day^{-1}$ |
| Energy consumption rate | $\delta$ | 57 | 23 | $cells^{-1}\ day^{-1}$ |
| Energy needed for half-maximal growth | $\phi_u$ | [0.01, 3] | 0.310 | |
| | $\phi_v$ | | [0.05, 0.16] | |
| Chemotactic coefficient | $\chi$ | 5500 | 1919 | $\mu m^2\ day^{-1}$ |
| Energy deficit inducing chemotaxis | $\xi_u$ | [2, 300] | [1, 8] | |
| | $\xi_v$ | | 140 | |

**Table 1.** Measured and inferred simulation parameters for one-population model (second column); and two-population model (third column). Parentheses indicate ranges with equally good fits. Corresponding non-dimensional parameters are shown in Supplementary Table 3. Net chemotaxis of cells is given by the interaction of the chemotactic coefficient ($\chi$) and energy deficit inducing chemotaxis ($\xi_u$), i.e. effective cell chemotaxis is approximately $\chi/\xi_u$. The larger $\xi_u$ the later the cells sense an energy deficit, i.e. the larger the energy gradient must be in order for them to accelerate their movement in response to it.

# Figure Legends

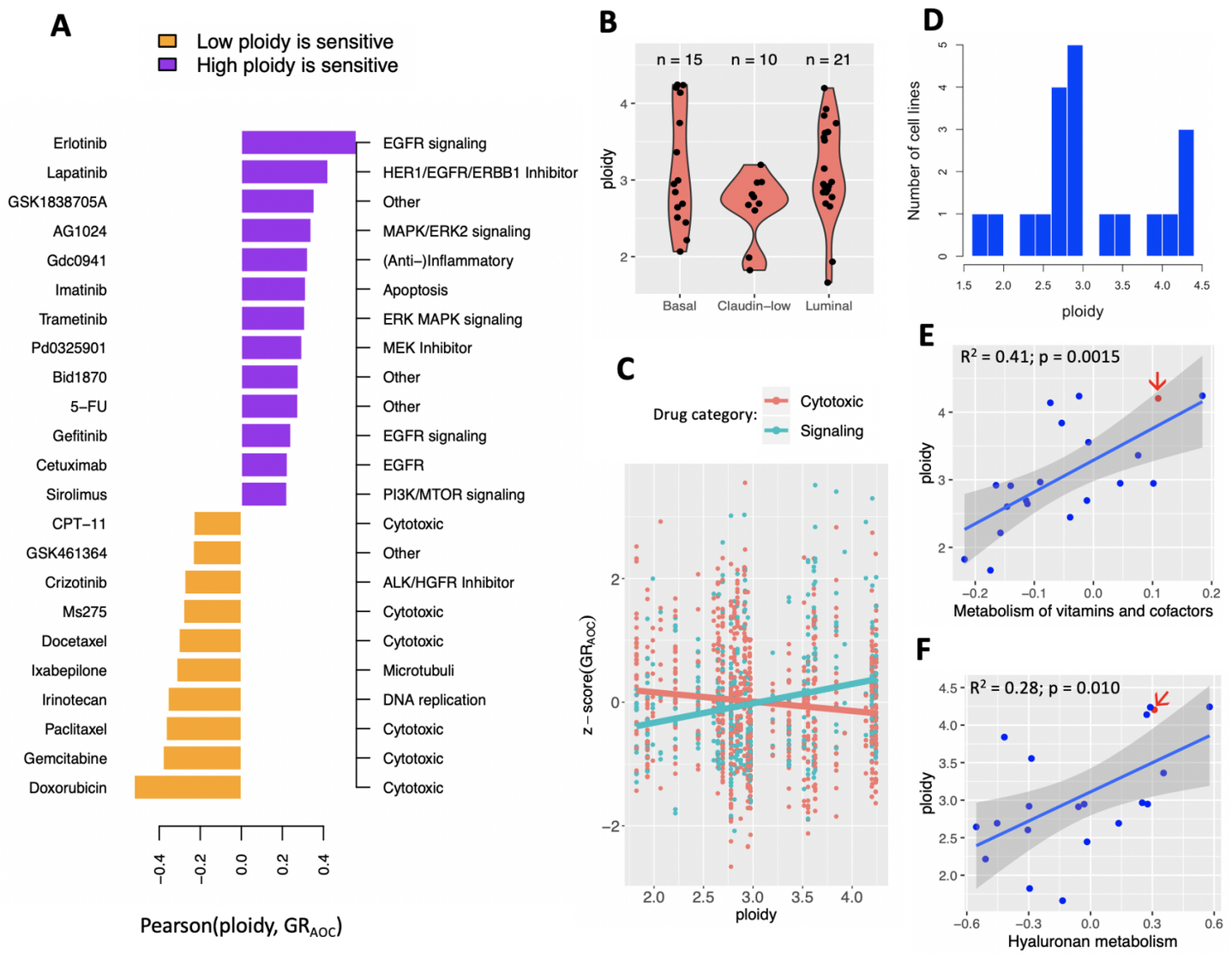



**Figure 1. Ploidy, pathway activity and drug sensitivity across breast cancer cell lines from CCLE**. (A) High-ploidy breast cancer cell lines are resistant to cytotoxic drugs, but tend to be more sensitive to inhibitors of mTOR, EGFR and MAPK signaling pathways. Hereby ploidy is defined as the number of chromosomes in the cell line's consensus karyotype, weighted by chromosome size. Only drugs with a Pearson correlation coefficient at or above 0.2 are shown here. (B) Distribution of ploidy within and across three molecular breast cancer subtypes. (C) Regression coefficient of ploidy as predictor of $GR_{AOC}$ has opposite signs depending on drug category across all subtypes. (D) Distribution of ploidy across 20 primary, adherent breast cancer cell lines from CCLE. (E-F) Ploidy is correlated with the activity of pathways involved in metabolism of vitamins and cofactors (E) and Hyaluronan metabolism (F). One cell line with available MEMA profiling data – HCC1954 – is highlighted (red arrow).

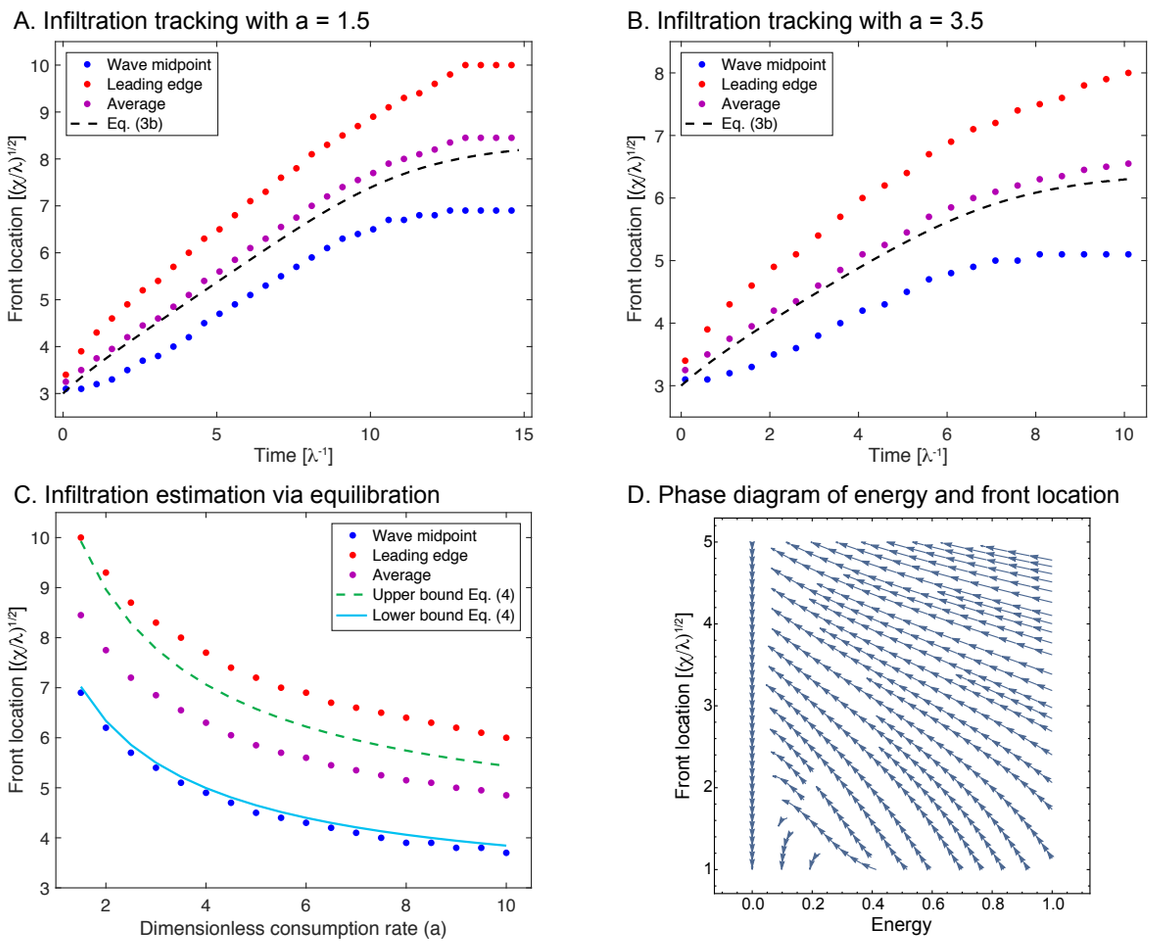



**Figure 2. Comparing analytical approximations of the degree of infiltration with those obtained from simulations**.(A-B) Traveling-wave solutions at energy consumption rates $a = 1.5$ (A) and $a = 3.5$ (B). (C) Upper and lower boundaries of traveling wave solutions estimated from equilibration are shown as a function of consumption rate. Approximation is found by lower and upper bound $\Lambda = 0, 1$ from equation (4). (D) Phase diagram of energy consumption and front location using the derived coupled system (3a)-(3b). (A-D) All approximations and simulations assume energy is uniformly distributed at all times, i.e. chemotaxis does not take place. Parameter values for initial seeding radius $(\rho_0)$, dish radius $(R)$, and sensitivity to low energy $(\phi)$ are set to 3, 10 and 0.05 respectively. Red = leading edge of wave (estimated by finding value of cell concentration closest to 0.01); blue = mid point of wave (estimated by finding value of cell concentration closest to 0.5); purple = average of red and blue; black lines are approximations based on analytical solutions. Time is in units of the maximal growth rate of the given cell line. Front location is given in units of the characteristic length $\sqrt{\chi/\lambda}$.

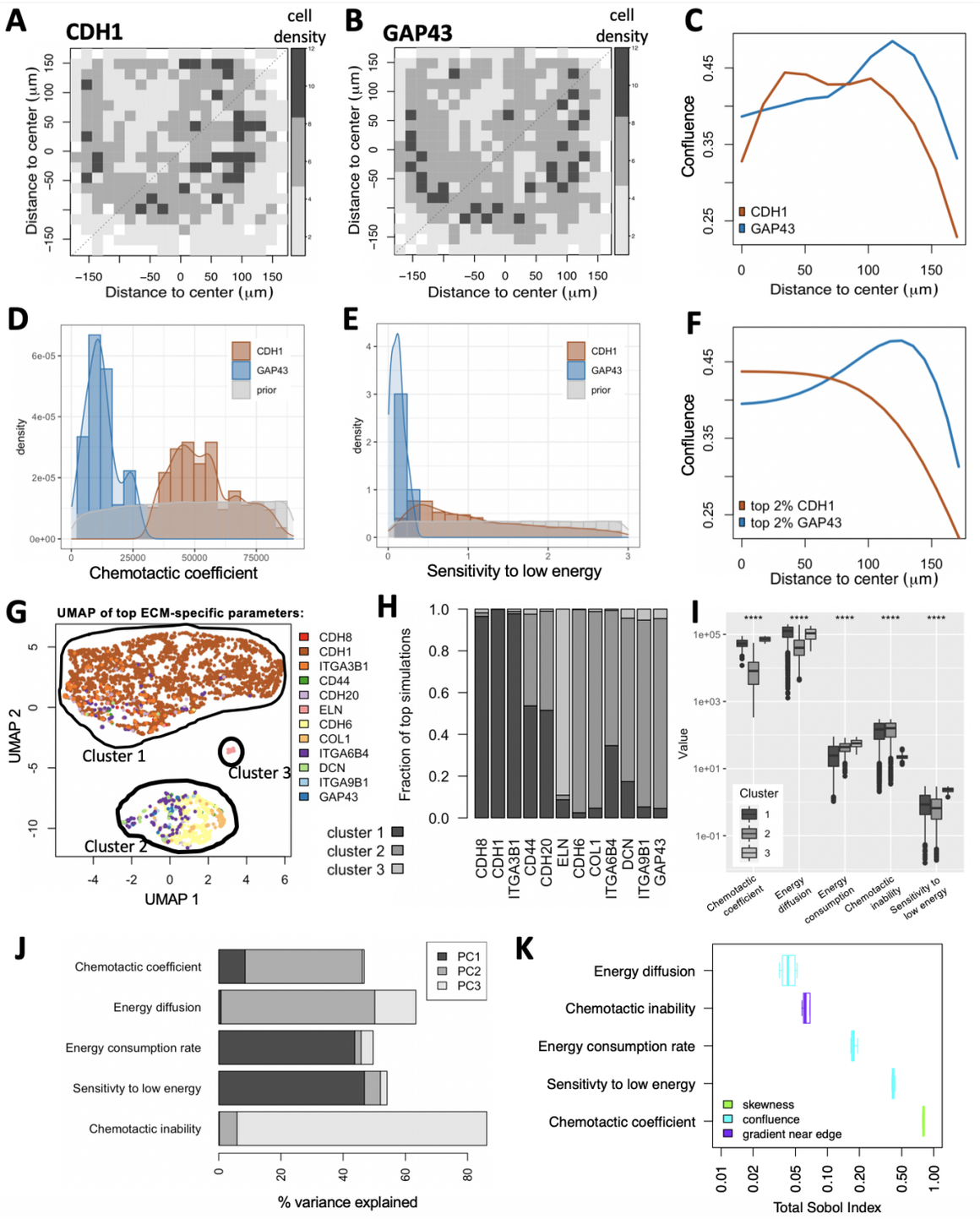



**Figure 3. Model calibration using MEMA profiling of HCC1954 cells**. ( **A-C**) **Experimentally measured data**. Variability in cell growth patterns across ECMs is demonstrated via two example ECMs: CDH1 (A) and GAP43 (B). The average local cell densities are displayed for both (color legend). (C) Projecting their 2D spatial distributions onto 1D reveals enrichment of cells at the edge of the ECM spot for GAP43, but not for CDH1. ( **D-F**) **Simulated data**. Comparing prior-distribution of chemotactic coefficients (D) and sensitivity to low energy (E) to ECM-specific posterior distributions reveals clear differences between CDH1 and GAP43 for both parameters. (F) Maximum-likelihood parameter choices for CDH1 and GAP43 result in distinct spatial distributions between the two ECMs, each of which resemble the measured distributions (C). ( **G-I**) **ECM specific model parameters**. Simulations were compared to each measured ECM-specific growth pattern and ranked by their maximum similarity. (G-H) The five model parameters from the top 2.3% simulations were projected onto UMAP space, revealing three clusters. Color-coding simulations by the ECM responsible for their presence in the top simulations suggests enrichment of most ECMs to only a single cluster (G). (H) This was confirmed when comparing cluster membership across the 12 represented ECMs. (I) All five model parameters (x-axis) show significant differences between the three clusters (****: $p <= 0.0001$). ( **J-K**) **Parameter sensitivity analysis**. (J) The % variance explained per parameter shows significant contribution of all five model parameters to at least one of the first three principal components (PCs). (K) The Sobol index (x-axis) tells us which parameter best explains which aspect of the cells' spatial distribution (color-code): its skewness, confluence or gradient near the edge.

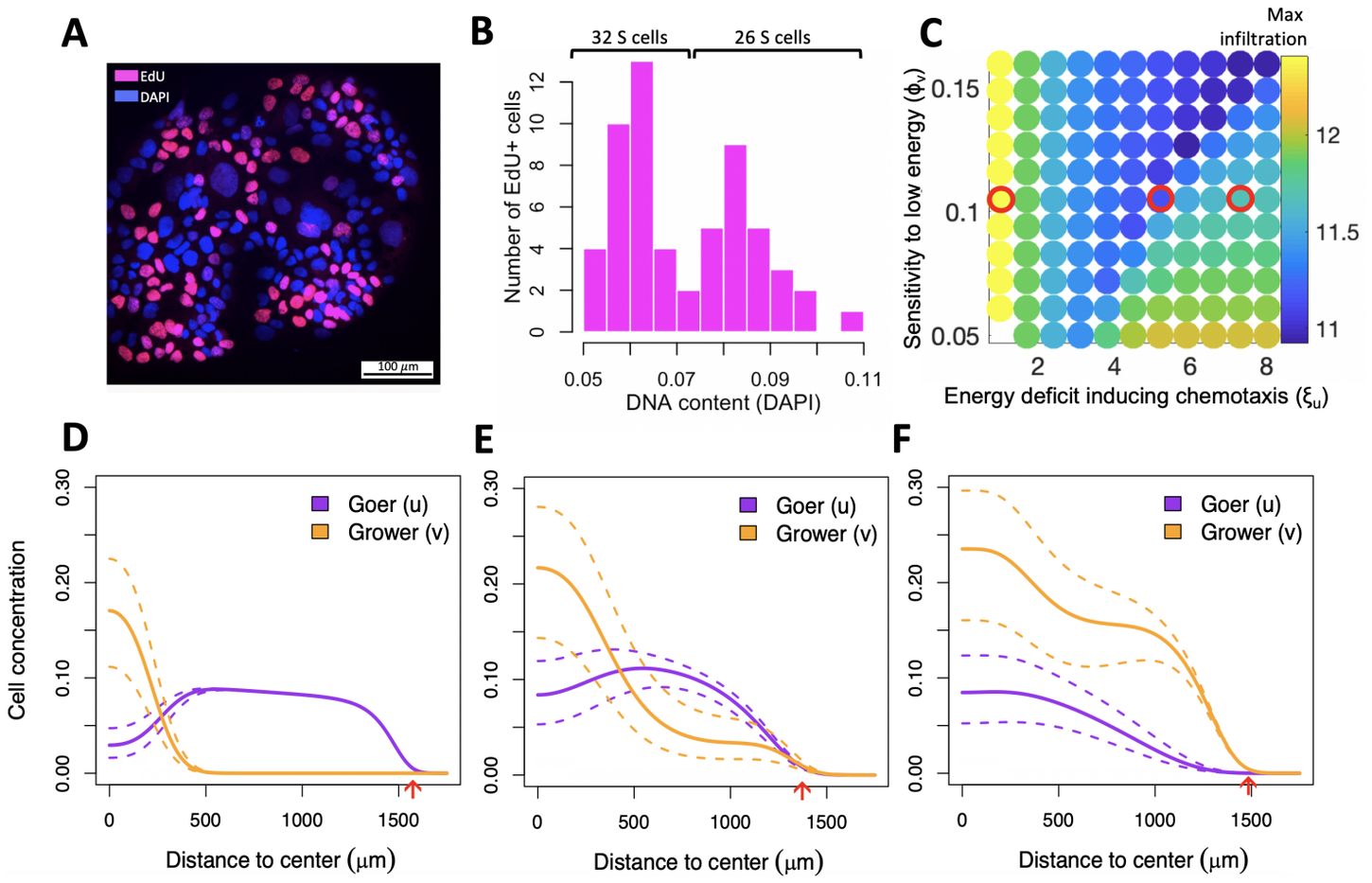



**Figure 4. Internal competition of co-existing subpopulations for same space slows down invasion of the metapopulation**. (A) DNA content and cell cycle state of 162 cells growing on HGF-exposed ICAM1. (B) DAPI intensity of 58 replicating (EdU+) cells shows a bimodal distribution, indicating the presence of two subpopulations – a low-ploidy population (grower) comprising 55% cells and a high-ploidy population (goer) comprising 45% cells. (C) Arms race between the grower's energetic sensitivity (y-axis) and the goer's chemotactic ability (x-axis) reduces infiltration distance (color bar). Red circles outline parameter combinations of interest explored in (D-F). (D-F) Spatial distribution of goer and grower for parameter values outlined in (C). Dotted lines outline extreme trajectories of expected cell concentrations due to incertitude in initial goer/grower proportions, as estimated from the silhouette coefficient of cells in panel B (see also Supplementary Fig. 6B). (D) High chemotactic motility will cause the goer to leave the center of the dish too soon, leaving room for the grower to expand there. (E) With an intermediate motility the goer succeeds maintaining high representation both at the center and edge of the dish. (F) Low motility will prevent the goer from gaining a sufficient spatial lead from the grower while energy is still abundant, and it will lose dominance at the edge of the dish once energy becomes sparse. Red arrows indicate maximum infiltration distance achieved by either of the two populations.

## References


1. Spiteri I, Caravagna G, Cresswell GD, Vatsiou A, Nichol D, Acar A, et al. Evolutionary dynamics of residual disease in human glioblastoma. Annals of Oncology: Official Journal of the European Society for Medical Oncology. 2019;30(3):456–463.

2. Brastianos PK, Carter SL, Santagata S, Cahill DP, Taylor-Weiner A, Jones RT, et al. Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets. Cancer Discovery. 2015 Nov;5(11):1164–1177. 00040.

3. Angelova M, Mlecnik B, Vasaturo A, Bindea G, Fredriksen T, Lafontaine L, et al. Evolution of Metastases in Space and Time under Immune Selection. Cell. 2018 Oct;175(3):751–765.e16. Available from: http://www.sciencedirect.com/science/article/pii/S0092867418312303.

4. Hannibal RL, Chuong EB, Rivera-Mulia JC, Gilbert DM, Valouev A, Baker JC. Copy number variation is a fundamental aspect of the placental genome. PLoS genetics. 2014 May;10(5):e1004290.

5. Dewhurst SM, McGranahan N, Burrell RA, Rowan AJ, Grönroos E, Endesfelder D, et al. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. Cancer Discovery. 2014 Feb;4(2):175–185. 00068.

6. Bielski CM, Zehir A, Penson AV, Donoghue MTA, Chatila W, Armenia J, et al. Genome doubling shapes the evolution and prognosis of advanced cancers. Nature genetics. 2018 Aug;50(8):1189–1195. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6072608/.

7. Amend SR, Torga G, Lin KC, Kostecka LG, Marzo Ad, Austin RH, et al. Polyploid giant cancer cells: Unrecognized actuators of tumorigenesis, metastasis, and resistance. The Prostate. 2019;79(13):1489–1497. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pros.23877. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/pros.23877.

8. Pienta KJ, Hammarlund EU, Axelrod R, Brown JS, Amend SR. Poly-aneuploid cancer cells promote evolvability, generating lethal cancer. Evolutionary Applications;n/a(n/a). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eva.12929. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/eva.12929.

9. Zhang S, Mercado-Uribe I, Xing Z, Sun B, Kuang J, Liu J. Generation of cancer stem-like cells through the formation of polyploid giant cancer cells. Oncogene. 2014 Jan;33(1):116–128.

10. López S, Lim EL, Horswell S, Haase K, Huebner A, Dietzen M, et al. Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. Nature Genetics. 2020 Mar;52(3):283–293.

11. Watson SS, Dane M, Chin K, Tatarova Z, Liu M, Liby T, et al. Microenvironment-Mediated Mechanisms of Resistance to HER2 Inhibitors Differ between HER2+ Breast Cancer Subtypes. Cell systems. 2018 Mar;6(3):329–342.e6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5927625/.

12. Keenan AB, Jenkins SL, Jagodnik KM, Koplev S, He E, Torre D, et al. The Library of Integrated Network-Based Cellular Signatures NIH Program: System-Level Cataloging of Human Cells Response to Perturbations. Cell Systems. 2018;6(1):13–24.

13. Pirkmajer S, Chibalin AV. Serum starvation: caveat emptor. American Journal of Physiology-Cell Physiology. 2011 May;301(2):C272–C279. Available from: https://www.physiology.org/doi/full/10.1152/ajpcell.00091.2011.

14. Growth of human diploid cells (strain MRC-5) in defined medium; replacement of serum by a fraction of serum ultrafiltrate. - PubMed - NCBI;. Available from: https://www.ncbi.nlm.nih.gov/pubmed/422676.

15. Bartholomew JC, Yokota H, Ross P. Effect of serum on the growth of balb 3T3 A31 mouse fibroblasts and an SV40-transformed derivative. Journal of Cellular Physiology. 1976 Jul;88(3):277–286. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jcp.1040880303.

16. Cai W, Rook SL, Jiang ZY, Takahara N, Aiello LP. Mechanisms of hepatocyte growth factor-induced retinal endothelial cell migration and growth. Investigative Ophthalmology & Visual Science. 2000 Jun;41(7):1885–1893.

17. Hatzikirou H, Basanta D, Simon M, Schaller K, Deutsch A. 'Go or grow': the key to the emergence of invasion in tumour progression? Mathematical medicine and biology: a journal of the IMA. 2012 Mar;29(1):49–65.

18. Kathagen-Buhmann A, Schulte A, Weller J, Holz M, Herold-Mende C, Glass R, et al. Glycolysis and the pentose phosphate pathway are differentially associated with the dichotomous regulation of glioblastoma cell migration versus proliferation. Neuro-Oncology. 2016 Sep;18(9):1219–1229. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4998991/.

19. Dhruv HD, McDonough Winslow WS, Armstrong B, Tuncali S, Eschbacher J, Kislin K, et al. Reciprocal Activation of Transcription Factors Underlies the Dichotomy between Proliferation and Invasion of Glioma Cells. PLoS ONE. 2013 Aug;8(8):e72134. 00009. Available from: http://dx.doi.org/10.1371/journal.pone.0072134.

20. Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Examining Go-or-Grow Using Fluorescent Cell-Cycle Indicators and Cell-Cycle-Inhibiting Drugs. Biophysical Journal. 2020 Mar;118(6):1243–1247. Available from: http://www.sciencedirect.com/science/article/pii/S0006349520301119.

21. Wang CH, Matin S, George AB, Korolev KS. Pinned, locked, pushed, and pulled traveling waves in structured environments. Theoretical Population Biology. 2019;127:102–119.

22. Bayliss A, Volpert VA. Complex predator invasion waves in a Holling–Tanner model with nonlocal prey interaction. Physica D: Nonlinear Phenomena. 2017 May;346:37–58. Available from: http://www.sciencedirect.com/science/article/pii/S016727891630481X.

23. Wasserstein L. Markov processes over denumerable products of spaces describing large systems of automata. 1969 Jan;Available from: https://scinapse.io/papers/51701124.

24. Sanderson C, Curtin R. Armadillo: a template-based C++ library for linear algebra. Journal of Open Source Software. 2016 Jun;1(2):26. Available from: https://joss.theoj.org/papers/10.21105/joss.00026.

25. van der Meer D, Barthorpe S, Yang W, Lightfoot H, Hall C, Gilbert J, et al. Cell Model Passports—a hub for clinical, genetic and functional datasets of preclinical cancer models. Nucleic Acids Research. 2019 Jan;47(D1):D923–D929. Publisher: Oxford Academic. Available from: https://academic.oup.com/nar/article/47/D1/D923/5107576.

26. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012 Mar;483(7391):603–607. 00591. Available from: http://www.nature.com/nature/journal/v483/n7391/full/nature11003.html.

27. Hafner M, Niepel M, Chung M, Sorger PK. Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. Nature Methods. 2016;13(6):521–527.

28. Hafner M, Heiser LM, Williams EH, Niepel M, Wang NJ, Korkola JE, et al. Quantification of sensitivity and resistance of breast cancer cell lines to anti-cancer drugs using GR metrics. Scientific Data. 2017 Nov;4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5674849/.

29. Heiser LM, Sadanandam A, Kuo WL, Benz SC, Goldstein TC, Ng S, et al. Subtype and pathway specific responses to anticancer compounds in breast cancer. Proceedings of the National Academy of Sciences of the United States of America. 2012 Feb;109(8):2724–2729.

30. Daemen A, Griffith OL, Heiser LM, Wang NJ, Enache OM, Sanborn Z, et al. Modeling precision treatment of breast cancer. Genome Biology. 2013;14(10):R110.

31. Bräutigam K, Mitzlaff K, Uebel L, Köster F, Polack S, Pervan M, et al. Subtypes of Triple-negative Breast Cancer Cell Lines React Differently to Eribulin Mesylate. Anticancer Research. 2016 Jun;36(6):2759–2766.

32. Dai X, Cheng H, Bai Z, Li J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. Journal of Cancer. 2017 Sep;8(16):3131–3141. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5665029/.

33. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013 Jan;14(1):7. Available from: https://doi.org/10.1186/1471-2105-14-7.

34. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic Acids Research. 2014 Jan;42(Database issue):D472–D477. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3965010/.

35. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Research. 2013 Jan;41(Database issue):D955–D961. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531057/.

36. Chen WYJ, Abatangelo G. Functions of hyaluronan in wound repair. Wound Repair and Regeneration. 1999;7(2):79–89. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1524-475X.1999.00079.x.

37. Ellis IR, Schor SL. Differential effects of TGF-beta1 on hyaluronan synthesis by fetal and adult skin fibroblasts: implications for cell migration and wound healing. Experimental Cell Research. 1996 Nov;228(2):326–333.

38. Fisher RA. THE WAVE OF ADVANCE OF ADVANTAGEOUS GENES. Annals of Eugenics. 1937 Jun;7(4):355–369. Available from: http://doi.wiley.com/10.1111/j.1469-1809.1937.tb02153.x.

39. Liotta LA. Tumor invasion and metastases–role of the extracellular matrix: Rhoads Memorial Award lecture. Cancer Research. 1986 Jan;46(1):1–7.

40. Hynes RO. Extracellular matrix: not just pretty fibrils. Science (New York, NY). 2009 Nov;326(5957):1216–1219. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3536535/.

41. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nature Biotechnology. 2018 Dec;.

42. Xu C, Su Z. Identification of cell types from single-cell transcriptomes using a novel clustering method. Bioinformatics (Oxford, England). 2015 Jun;31(12):1974–1980.

43. Sobol IM. Global Sensitivity Indices for Nonlinear Mathematical Models; 2004. .

44. Lim S, Nam H, Jeon JS. Chemotaxis Model for Breast Cancer Cells Based on Signal/Noise Ratio. Biophysical Journal. 2018;115(10):2034–2043.

45. Keller EF, Segel LA. Model for chemotaxis. Journal of Theoretical Biology. 1971 Feb;30(2):225–234. Available from: http://www.sciencedirect.com/science/article/pii/0022519371900506.

46. Anderson ARA, Rejniak KA, Gerlee P, Quaranta V. Microenvironment driven invasion: a multiscale multimodel investigation. Journal of mathematical biology. 2009 Apr;58(4-5):579–624. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5563464/.

47. Kimmel GJ, Gerlee P, Altrock PM. Time scales and wave formation in non-linear spatial public goods games. PLOS Computational Biology. 2019 Sep;15(9):e1007361. Publisher: Public Library of Science. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1007361.

48. Einstein A. On the Motion of Small Particles Suspended in Liquids at Rest Required by the Molecular-Kinetic Theory of Heat. 1905;17:208.

49. Andor N, Simonds EF, Czerwinski DK, Chen J, Grimes SM, Wood-Bouwens C, et al. Single-cell RNA-Seq of lymphoma cancers reveals malignant B cell types and co-expression of T cell immune checkpoints. Blood. 2018 Jan;p. blood–2018–08–862292. Available from: http://www.bloodjournal.org/content/early/2018/12/26/blood-2018-08-862292.

50. Birkbak NJ, Eklund AC, Li Q, McClelland SE, Endesfelder D, Tan P, et al. Paradoxical relationship between chromosomal instability and survival outcome in cancer. Cancer research. 2011 May;71(10):3447–3452. 00068 PMID: 21270108.

51. Morrissy AS, Garzia L, Shih DJH, Zuyderduyn S, Huang X, Skowron P, et al. Divergent clonal selection dominates medulloblastoma at recurrence. Nature. 2016 Jan;529(7586):351–357. 00000. Available from: http://www.nature.com/nature/journal/v529/n7586/full/nature16478.html.

52. Laughney AM, Elizalde S, Genovese G, Bakhoum SF. Dynamics of Tumor Heterogeneity Derived from Clonal Karyotypic Evolution. Cell Reports. 2015 Aug;12(5):809–820.

53. Elizalde S, Laughney AM, Bakhoum SF. A Markov chain for numerical chromosomal instability in clonally expanding populations. PLoS computational biology. 2018;14(9):e1006447.

54. Donovan P, Cato K, Legaie R, Jayalath R, Olsson G, Hall B, et al. Hyperdiploid tumor cells increase phenotypic heterogeneity within Glioblastoma tumors. Molecular bioSystems. 2014 Jan;.

55. Bailly M, Wyckoff J, Bouzahzah B, Hammerman R, Sylvestre V, Cammer M, et al. Epidermal Growth Factor Receptor Distribution during Chemotactic Responses. Molecular Biology of the Cell. 2000 Nov;11(11):3873–3883. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC15043/.

56. Gulhati P, Bowen KA, Liu J, Stevens PD, Rychahou PG, Chen M, et al. mTORC1 and mTORC2 regulate EMT, motility, and metastasis of colorectal cancer via RhoA and Rac1 signaling pathways. Cancer Research. 2011 May;71(9):3246–3256.

57. James RG, Davidson KC, Bosch KA, Biechele TL, Robin NC, Taylor RJ, et al. WIKI4, a novel inhibitor of tankyrase and Wnt/ß-catenin signaling. PloS One. 2012;7(12):e50457.

58. Du L, Li X, Zhen L, Chen W, Mu L, Zhang Y, et al. Everolimus inhibits breast cancer cell growth through PI3K/AKT/mTOR signaling pathway. Molecular Medicine Reports. 2018 May;17(5):7163–7169.

59. Alalem M, Ray A, Ray BK. Metformin induces degradation of mTOR protein in breast cancer cells. Cancer Medicine. 2016 Oct;5(11):3194–3204. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5119975/.

60. Liu L, Li F, Cardelli JA, Martin KA, Blenis J, Huang S. Rapamycin inhibits cell motility by suppression of mTOR-mediated S6K1 and 4E-BP1 pathways. Oncogene. 2006 Nov;25(53):7029–7040.

61. Witelski TP, Ono K, Kaper TJ. ON AXISYMMETRIC TRAVELING WAVES AND RADIAL SOLUTIONS OF SEMILINEAR ELLIPTIC EQUATIONS. Natural Resource Modeling. 2008 Jun;13(3):339–388. Available from: http://doi.wiley.com/10.1111/j.1939-7445.2000.tb00039.x.

62. Lin CH, Lee JK, LaBarge MA. Fabrication and use of microenvironment microarrays (MEArrays). Journal of Visualized Experiments: JoVE. 2012 Oct;(68).