Accepted Manuscript

Title: Single-cell approaches to cell competition:
high-throughput imaging, machine learning and simulations

Authors: Daniel Gradeci, Anna Bove, Guillaume Charras,
Alan R. Lowe, Shiladitya Banerjee

Please cite this article as: Gradeci D, Bove A, Charras G, Lowe AR, Banerjee S, Single-cell approaches to cell competition: high-throughput imaging, machine learning and simulations, *Seminars in Cancer Biology* (2019), https://doi.org/10.1016/j.semcancer.2019.05.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Single-cell approaches to cell competition: high-throughput imaging, machine learning and simulations

Daniel Gradeci[†1,2], Anna Bove[†2,3], Guillaume Charras[*2,3,4], Alan R. Lowe[*2,4,5], Shiladitya Banerjee[*1,4]

[1]Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[2]London Centre for Nanotechnology, University College London, 17-19 Gordon Street, London, WC1H 0AH, UK
[3]Department of Cell and Developmental Biology, University College London, Gower Street, London, WC1E 6BT, UK
[4]Institute for the Physics of Living Systems, University College London, Gower Street, London, WC1E 6BT, UK
[5]Institute for Structural and Molecular Biology, University College London, Gower Street, London, WC1E 6BT, UK

[†]These authors contributed equally

[*]correspondence: g.charras@ucl.ac.uk, a.lowe@ucl.ac.uk, shiladitya.banerjee@ucl.ac.uk

Abstract
Cell competition is a quality control mechanism in tissues that results in the elimination of less fit cells. Over the past decade, the phenomenon of cell competition has been identified in many physiological and pathological contexts, driven either by biochemical signaling or by mechanical forces within the tissue. In both cases, competition has generally been characterized based on the elimination of loser cells at the population level, but significantly less attention has been focused on determining how single-cell dynamics and interactions regulate population-wide changes. In this review, we describe quantitative strategies and outline the outstanding challenges in understanding the single cell rules governing tissue-scale competition dynamics. We propose quantitative metrics to characterize single cell behaviors in competition and use them to distinguish the types and outcomes of competition. We describe how such metrics can be measured experimentally using a novel combination of high-throughput imaging and machine

learning algorithms. We outline the experimental challenges to quantify cell fate dynamics with high-statistical precision, and describe the utility of computational modeling in testing hypotheses not easily accessible in experiments. In particular, cell-based modeling approaches that combine mechanical interaction of cells with decision-making rules for cell fate choices provide a powerful framework to understand and reverse-engineer the diverse rules of cell competition.

Keywords
Cell competition; single-cell approach; machine learning; imaging; computational modeling.

1. Introduction

1.1. What is cell competition?
Competition between cells is a phenomenon that results in the elimination of less fit cells (losers) from a tissue, that is then taken over by cells with greater fitness (winners) [1-3]. For the loser cells, the outcome depends strongly on the context: when they are cultured in a homotypic environment, they thrive but, when they are cultured in a mixed population, they are eliminated by the winner cells (**Fig 1A**).

Competition was originally identified in Drosophila, where it was thought to operate as an intrinsic homeostatic control mechanism, maximizing tissue fitness in the growing embryo by destroying suboptimal cells[4, 5]. Initially, most mutations leading to cell competition appeared to confer a loser status to the affected cells (e.g. proteins involved metabolic activity, cell growth, polarity and stress responses). However, mutations that conferred a competitive advantage were soon identified, allowing the affected cells to eliminate normal cells even when vastly outnumbered. Cell competition has now been identified in many species, at different stages of animal development, as well as in physiological and pathological conditions [6].

Over the years, a number of mechanisms of cell competition have been identified, involving either biochemical processes or mechanical interactions (**Fig. 1B**). In all cases, competition has generally been evaluated based on complete elimination of loser cells but significantly less attention has been placed on the dynamics of elimination or determining how cell-scale events lead to population wide changes. In this review, we first provide a short introduction to competition (leaving more detailed descriptions of signaling pathways to other reviews in this issue), then outline the outstanding questions in cell competition at the single-cell scale, and propose quantitative strategies to address these questions.

1.2. Competition through biochemical processes
In the broadest sense, biochemical competition can be classified into two major categories: the ligand capture model, which involves a passive competition for a diffusible survival factor, and the comparative fitness model, where cells in contact with one another compare their fitness through signaling[1, 2].

The ligand capture model suggested the existence of a Darwinian-like competition among cells for limiting amounts of pro-survival factors, resulting in the removal of cells with less efficient uptake [7-9]. In this model, loser cells in the vicinity of winners are starved of pro-survival factors but those further away can survive. In the ligand capture model, cell-type specific differences in physico-chemical parameters such as ligand diffusion, ligand-receptor reaction kinetics, and endocytic rate are all likely to play an important role in regulating the outcome of competition.

2

Subsequent studies, however, have challenged the ligand capture model[10, 11], suggesting an alternative mechanism of competition based on cells sensing their relative fitness. The comparative fitness model postulates that cell competition is governed by short-range cellular interactions. Cells are able to: (i) recognize differences with neighboring cells, (ii) acquire either a winner or loser status in response to the outcome of fitness comparison, (iii) eliminate cells less fit than themselves. This model is based on the definition of fitness as a parameter that cells can measure and compare with one another. A key cellular parameter that has been proposed to control the outcome of comparative fitness competition at the single cell level is the extent of heterotypic cell contact [12, 13], which in turn depends on physical parameters such as intercellular adhesion energy, junctional tension and cortical tension [14].

1.3. Mechanical competition

In recent years, experimental observations have suggested the existence of an entirely different mode of cell competition determined by mechanical forces within the tissue [15]. In this model, competition acts independently of cell recognition and fitness comparison, rather it results from the combination of faster growth rates in the winner cells and hypersensitivity to crowding in the loser cells [1, 2, 16]. The concept of mechanical competition was first introduced in a seminal paper by Boris Shraiman [17], where it was theorized that mechanical stresses generated by differential growth could underlie competitive interactions between distinct cell types. Due to differential sensitivities to mechanical compression, slow growing cells will be eliminated at a lower pressure when surrounded by a faster growing population. This theory subsequently led to the concept of homeostatic pressure [18, 19], which defines the tissue pressure at which the rate of cell death is precisely balanced by the rates of cell growth and proliferation. In this framework, rates of cell proliferation and death are modulated by mechanical pressure, such that the cell population with a higher homeostatic pressure will eliminate the other. The existence of mechanical competition was independently confirmed in experiments by Levayer and colleagues, while studying delamination occurring in the midline of *Drosophila Melanogaster* notum epithelia [20], and by Wagstaff and colleagues while studying interactions between *Scribble* depleted and wildtype MDCK cells[21]. Indeed, these experiments have revealed that each cell type possesses a preferred homeostatic density or pressure, at which cell deaths exactly compensate cell births. When crowding is suddenly increased from the homeostatic density, apoptoses and live delaminations decrease cell density [22, 23], while when crowding is decreased suddenly, cell divisions are upregulated [24, 25]. Therefore, if loser cells with low homeostatic density are placed in competition with winner cells with high homeostatic density, apoptoses will be triggered in losers but not in winners once density becomes sufficiently high [21, 26].

The physical or geometrical parameters measured by cells to gain information about their density remain unclear. Recent experiments examining extrusion in homotypic environments suggest that, at high density, live cell extrusion is preceded by increased compression on the extruded cell [27] and that extrusion occurs primarily at locations of defects in the nematic ordering of epithelial cells [28]. In competitive settings, compressive strains also lead to increased probability of extrusion [20]. Both nematic order and cellular strain are affected by physical variables such as cell density, intercellular adhesion, junctional/cortical tension, and cell motility.

1.4 Quantitative approaches to deciphering the rules of competition

To date, cell competition has been mostly characterized at the population scale by measuring the elimination of loser cells in movies or still images. Indeed, in its most basic definition, the fitness of a population of cells can be approximated by its net growth rate $\mu = N^{-1} dN/dt$, where $N(t)$ is the number of cells at time t (**Fig. 1C**). A cell type is fitter if it grows more than its competitor (**Fig 1B**). Any factor that increases growth rate will increase fitness, whereas any factor that increases cell

death will decrease fitness. Although population scale statistics have been instrumental in describing competition, they do not provide insight into how population shifts arise from the multitude of interactions at the single cell level. Indeed, as with any game, knowing only the result is of little use for determining the rules. Rather, we need to follow the game in detail. In scientific terms, we need to observe each step of the competition, decide on the relevant parameters, characterize each step, and, once we have formulated a reasonable hypothesis, we need to evaluate our understanding by simulations.

In the following sections, we will focus on high-throughput data acquisition, parameterization, characterization, and simulation of cell competition. We will draw on our experience of mechanical competition [26] but similar approaches can be applied to any type of competition.

2. Towards a single cell understanding of competition
2.1 Single cell-level behaviors in competition
Cell competition is a balance problem that necessitates an understanding of both apoptosis and proliferation. Apoptosis is often quantified in studies of cell competition but faster growing cells may be able to better take advantage of any space freed as a result of cell death [29]. In subconfluent conditions, cell types with higher motility may be able to rapidly disperse to occupy space, whereas, in confluent conditions, mixing allows for more efficient elimination of loser cells [12]. Therefore, at a minimum, we need to characterize cell proliferation, death, and migration.

At a population level, these behaviors can be quantified temporally by computing the number of losers and winners, the fraction of the total population they occupy, their birth/death rate (**Fig 1C**), and their migration velocities or mean square displacement. When competition results in complete elimination of losers, the kinetics of competition can be compared for example by measuring the timescale for cell elimination. However, competition may also yield stable coexistence of the two cell types, with losers largely outnumbered but not eliminated (as observed in [12], **Fig 1A**). In this case, a quantification of the survival fraction may allow to compare across conditions (**Fig. 1C**).

To gain insight into competition at the single cell level, we must understand how each of these descriptors of population change is affected by the local environment and, as a consequence, we must also characterize local features of the tissue environment. For example, cell crowding is an important variable in mechanical competition. Therefore, characterizing the local cell density and its temporal evolution is of interest. In comparative fitness competition, contact between winners and losers is a key parameter which should be characterized over time for each cell. In the discussion that follows, we use our in vitro model system of Madin-Darby Canine Kidney (MDCK) cells to describe the methodologies for high throughput imaging of cell competition dynamics at the single-cell level.

2.2 High throughput imaging of cell competition
As competition takes place over durations of tens of hours, long duration movies acquired by time-lapse live cell microscopy are essential to decipher it (**Fig 2**). Considerations when designing imaging assays are setting the spatial and temporal resolutions, as well as a large enough field of view that enables imaging of larger amount of cells. The spatial resolution must be sufficient to easily distinguish the single cell events that underlie competition such as proliferation, death, and migration. The temporal resolution must be such that no event (death or division) goes undetected. If cells are highly migratory, it is imperative to choose a short enough time interval for cell tracking to allow for meaningful measurements of cell velocity or mean-squared displacements. If they are not, the duration of anaphase or cytokinesis will inform on the suitable

interval. Having chosen a temporal resolution, one can decide on how many fields of view to acquire based on frame acquisition time.

Analysis of the imaging data is a critical consideration. Experiments typically yield 10^2-10^3 cells per field of view, over hundreds of time-points and multiple fields of view, resulting in millions of cells worth of single-cell data [26]. Fluorescent markers are often used to distinguish winners from losers [21, 26, 30, 31]. Histone markers with different fluorophores are practical because they can be robustly segmented (**Fig 2A**). Although histone markers allow for identification of proliferation and death, they are late indicators of a commitment that has occurred several hours before. Hence, it may be useful to combine histone markers with additional reporter constructs, such as geminin [32] for entry into the cell cycle or active caspase reporters for cell death [33], to gain more precise information on when commitment occurred.

From the image data, each cell must be robustly identified, localized, labelled and tracked over the duration of its existence. Several software packages have been developed for general single-cell tracking. These include *tTt* [34], *CellProfiler* [35], *CellCognition* [36] and *LEVER* [37] amongst others (reviewed in [34]). For quantifying epithelial dynamics in vivo, several other software packages have also been developed [38-40]. However, single-cell tracking remains a challenging problem. For example, a method with a perceived high tracking accuracy of 99% actually means that up to 10 cells at every time point are mis-identified or mis-categorized. This is particularly problematic in the study of rare events such as cell death or division (see **section 2.5** for more discussion). Approaches utilizing advanced computer vision and machine learning to perform automated segmentation and event detection (**Fig 2B**) are improving the accuracy of such analysis methods. Our own software, *Sequitr*[1] [26], has been developed for cell competition, combining state-of-the-art deep learning for robust cell segmentation and classification together with a Bayesian approach to tracking (**Fig 2C**).

2.3 Population and single-cell metrics to characterize competition

As many different mutations can drive competition, defining general metrics to characterize competition allows us to categorize competitions based on quantitative signatures and determine the effect of perturbations.

Using segmented and classified image data, population fate metrics, such as the probability of cell death/division as a function of time or the cumulative number of deaths/divisions, can be computed for each cell type. Further insight into competition at the single cell level necessitates correlation of fate with features of the environment likely to influence competition. For example, in comparative fitness competition, losers in contact with winners are eliminated and increased contact leads to increased probability of apoptosis. Therefore, quantifying how proliferation and death depend on the composition of the local neighborhood will likely be informative. Conversely, in mechanical competition, crowding has been shown to be a key variable [21]. Hence, we need to determine the dependency of death and division on local cell density.

The most obvious definition for cell density is the inverse of the area of the cell of interest. However, this definition can be misleading because some epithelial cells have a similar morphology whether they are alone or surrounded by neighbors. For a more robust definition, the area of the cell and its direct neighbors should be taken into account, for example by defining the density as: $\rho(cell) = \frac{1}{A(cell)} + \sum_{i=1}^{n} \frac{1}{A(i)}$ with *n* the number of neighbors, $A(i)$ the area of neighbor

---

[1] https://github.com/quantumjot/CellTracking

i, and $A(cell)$ the area of the cell of interest (**Fig 3B**). Using this definition, a density can be defined for each cell at each time point and events (division, apoptosis, etc) can be binned as a function of density to yield curves relating the probability of the event to the local density. This approach reveals that, in mechanical competition, cell death in loser cells increases sharply with density. Cell death in winners appears to follow a curve similar to the losers but shifted towards higher densities [26]. These observations suggest that some aspects of mechanical competition may be understood through differences in homeostatic density between winners and losers.

In comparative fitness competition, heterotypic contacts govern cell fate and thus a metric that quantifies this is necessary. One simple idea is to categorize cells by their number of winner and loser neighbors. After doing this for each cell and at each time point, we can classify each event and each observation as a function of their winner and loser neighbors to generate a neighborhood plot (**Fig 3A-C**) [26]. This plot has the number of winner neighbors on the *y* axis and the number of loser neighbors on the *x* axis. In each grid position, the probability of the event is color coded and the number of events observed is indicated. This presentation allows to rapidly grasp whether the event of interest depends on the number or identity of neighbors. For example, if there is no competition, the neighborhood plot for apoptosis will have a uniform color: apoptosis does not depend on the number or identity of neighbors (**Fig 3A**). If we are examining mechanical competition, the probability of apoptosis will be higher for cells with more neighbors but their identity is not important. In this case, the neighborhood plot is symmetric about its diagonal (**Fig 3B**). However, the number of neighbors is not directly proportional to the local cell density (**Fig. 3B** – top). Therefore, to assess density-based competition in heterogeneous tissues it is important to generate neighborhood plots of local cell density. In comparative fitness competition, loser cells may be more likely to die when most of their neighbors are winners. In this case, the neighborhood plot is asymmetric with high probabilities of death above the diagonal and low probabilities of death below the diagonal (**Fig 3C**).

Other metrics may of course be useful in characterizing competition. For example, in comparative fitness competition, the percentage of the perimeter in heterotypic contact and the duration of this contact may be important parameters in determining cell fate. In mechanical competition, stress in the epithelium may play a role. Therefore, computing strain rate from cell velocity maps may be informative.

2.4 Statistics of rare events
A major challenge with such metrics becomes evident when we want to compare experiments. The dominating observables in cell competition are rare events, such as cell division or death. How can we determine if a perturbation experiment is significantly different from its control? How can we determine if grid positions are significantly different from one another?

Generally, to compare probability estimators, we determine if their values differ by more than their absolute precision. This absolute precision is often estimated by the standard error: $SE = \sqrt{\frac{p(1-p)}{N}}$ with *p* the estimated probability: *p=n/N* with *n* the number of events and *N* the number of observations. One issue arises from the calculations of statistics of rare events [41]. To illustrate this, consider an event of cell elimination that occurs as an independent Poisson process over the whole duration of a typical competition assay with a constant rate of 1% (0.01 deaths per cell per duration of experiment). Any given cell has a probability 1-e^-0.01 = 0.00995 of dying over the whole experiment. For a population of 1000 cells observed over the whole experiment, the standard error associated with cell elimination is 0.0031. However, if cells are only observed for

10% of the experiment length, the probability of elimination for one cell becomes $1-e^{-0.01*0.1}$ = 0.0009995 and the standard error for a population of 1000 cells becomes 0.000992, far less than for the whole experiment. Yet, neither the elimination process nor the cell population have changed. This improvement of precision is an artifact of a rare event. To overcome this issue, one must use measures of relative precision (Coefficient of variation: $cv = \sqrt{\frac{(1-p)}{pn}}$), where the standard errors are calculated relative to the actual probability of the rare event [41]. A $cv < 10\%$ enables robust comparison of grid positions and across experiments. If the cv does not reach such low values, more data needs to be acquired or the existing data needs to be binned into coarser categories.

2.5 Measuring relevant physical and biochemical variables
Ultimately, we want to link our observations from competition to underlying differences in physical and biochemical variables. Computer simulations provide a powerful tool to predict cell competition outcomes and test experimental hypotheses. As input, they require to measure the variables that may play a role in competition for each cell type separately as well as in competitive contexts. For example, for mechanical competition, the strength of intercellular adhesion, the stiffness of cells, and the homeostatic density may all play a role. In comparative fitness or ligand capture competition, the cascade of different biochemical reactions and protein interactions underlying fitness comparison or growth factor uptake would need to be characterized. Methods for measuring each variable are specific to each type of competition and each signaling pathway.

3. Computational modelling approaches to reverse-engineer cell competition
3.1 Local vs global models of cell competition
The availability of multi-scale quantitative data opens up new possibilities to validate computational models that can provide a bottom-up understanding of the fundamental physical principles governing cell competition.

Models of cell competition can be categorized as global or local depending on whether they are population-based or cell-based, respectively. Global population-based models are usually formulated as ordinary differential rate equations (ODEs) mapping the evolution of the number of winners and losers to continuous variables (**Fig. 4A**). Such models implement phenomenological rules for cell colony growth and decay, without considering local interactions. This simplification enables the use of mathematical analysis to infer the general characteristics of the cell competition phenomena, which can be conveniently tested in population-scale experiments. As attractive as reducing a complex multivariate system to simple differential equations may seem, these models can only predict the emergent bulk properties such as population ratios, their relative growth rates or tissue size. The influence of local topography, mechanics and cell-cell signaling cannot be taken into account, despite experimental and computational evidence showing their importance in metastasis [42] and in cell competition [12, 21, 26, 29]. In particular, ODE models are unsuited to describe morphologies of co-existing colonies observed in comparative fitness competition [12].

ODE based models tend to reduce the multi-variate parameters of cell competition to a single asymmetry parameter for the fitness of competition. Fitness asymmetry is sometimes modelled as the difference in growth rates in an ODE model [6, 43]. Yet, many studies of cell competition have concluded that differential growth is not the only source of competition. Another metric for asymmetry in fitness is the ratio of survival fractions of the two cell populations, formulated using a predator-prey model much like the well-known *Lotka-Volterra* model [44]. In other coupled rate

7

equation models [26], an asymmetry in fitness is incorporated via differential sensitivity of proliferation rate to cell density. Indeed, experiments and modelling showed that differences in apoptosis sensitivity to density alone cannot explain the outcome of competition. One limitation common to all ODE models is that they cannot account for spatial heterogeneities in the population nor single-cell level configurations, topology, and mechanics.

To circumvent these limitations, a common approach is to implement partial differential equation based models as a framework to incorporate active cellular mechanics, motility, as well as couplings between mechanical and chemical fields [45, 46]. Here, the two cell types are modelled as viscoelastic media with pre-defined mechanical and kinetic properties [47-50] where the interaction mechanisms are implemented phenomenologically. A drawback, however, is that such an approach leads to many undetermined physical parameters whose molecular origins remain speculative. Furthermore, the choice of fixed mechanical properties limits the emergent states of self-organization. While continuum models have been used to infer the mechanistic principles of tissue homeostasis [51], they remain ill-suited to describe contact-dependent biochemical competition, as well as topology-based cell-cell interactions that are essential regulators of global population fate.

3.2 Cell-based models
To understand how the behavior of single cells and local cell-cell interactions lead to population-scale dynamics, one needs to explicitly model individual cells as physical entities with interaction rules. These are implemented in the form of computational algorithms using tools from soft matter physics, statistical physics, and dynamical systems.

The vertex model has been extensively used in modelling epithelial tissues due to its conceptual simplicity and computational parsimony. The vertex model treats each cell as a two-dimensional polygon, with vertices and rectilinear edges shared between neighboring cells [52]. Vertices move in response to forces that arise from cell growth, interfacial tension, and intracellular contractile stress (**Fig. 4B**) [53], resulting in changes in cell shape. The choice of polygonal shapes and simple mechanical rules reduce the computational complexity of simulations, while the burden lies in implementing the rules for cell rearrangements and active cellular processes. Vertex models have been widely used in epithelial monolayer modelling including cell sorting [54], germband extension [55], ventral furrow formation [56], cell division and tissue elongation [57], wound healing [58-60], as well as tumor growth [61]. Vertex models have been used in studies of mechanical cell competition [43], as well as to infer the role of local epithelial topology on mutant clone expansion [29]. Vertex models, however, have some key disadvantages, including the difficulty to model arbitrary cell shapes, isolated cells, or cell clusters. Therefore, this framework is unsuited to capture the competition for free space which is an important aspect of competition prior to confluency for in vitro 2D cell cultures. For well packed epithelial in vivo, vertex-based models are well suited for describing competition for space created by dying cells [29].

A common alternative approach is to model cell competition using cellular automata [62], in which dynamic cell behavior is not governed by mechanical force balance but through probabilistic rules for cell apoptosis, mitosis, and mutations. Using this model, Basanta et al [62] have suggested that micro environmental factors such as the local concentration of oxygen or nutrients, and cell overcrowding may determine the expansion rate of the tumor colony. The results also show that tumor cells evolve to adapt their phenotypes to the microenvironment, such that environmental stresses determine the dominance of particular phenotypical traits. A similar approach was used in [42] to model tumor evolution, which showed that short-range cell dispersal and turnover can account for rapid intermixing of cells inside the tumor. How a small selective advantage in a

8

single cell within a large tumor allows colonization of the precursor mass in a clinically relevant time frame is a relevant mechanism for the rapid onset of resistance to chemotherapy. Cell automata models, however, are limited by the fact that they cannot account for the physical interactions between cells known to be important in cell competition and tumorigenesis. These limitations could be overcome using particle-based models [19, 63, 64] that allow for mechanical interactions as well as diverse multicellular configurations.

3.3 Cellular Potts Model for cell competition: a multi-layer model.

A comprehensive bottom-up insight into the mechanism of cell competition could be obtained using a combination of rule-based probabilistic models and mechanistic models to implement both signaling and realistic interactions via physical forces. This would result in a multi-layered model where, in its most fundamental layer, individual cells would follow robust mechanical principles governing their physical interactions. The physical layer is then coupled to a layer describing active cellular behavior as rules or (in game theoretical terms) strategies that cells follow to determine their fate. Such multi-layered framework for collective cell dynamics could be implemented following the commonly used Cellular Potts Model (CPM) [65].

The traditional CPM forms the physical layer of the model. At each time point, the CPM ensures that the cell shapes are such that free energy is minimized. The CPM is a mechanical lattice-based model (**Fig. 4C**), which represents each cell as a subset of lattice sites sharing the same cell identity. These lattice sites, referred to as cell-sites, constitute the building blocks of the cells in the computational model, and are only assigned the properties of the cell they belong to. For example, each cell type has some value of adhesion energy with other cells/objects as well as a preferred area and a compressibility modulus. The model evolves following a Metropolis algorithm, where a lattice site is chosen at random and assigned a new cell identity. The move is accepted (or rejected) if it lowers (or increases) the value of the Hamiltonian (free energy) function, which is a sum of adhesion energies and volume (or area) constraints. The Hamiltonian can be appropriately modified to implement active cell behaviors such as motility [66].

The CPM enables the addition of an exterior layer comprising the cellular automata rules. These rules are algorithmically executed for each cell separately at each time point to model active cellular behaviors such as growth, mitosis, apoptosis, etc. Such approaches have been applied to the growth of tumor cell clusters by implementing the effects of cell compressibility, active motility and contact inhibition [67]. In the context of cell competition, we have implemented a multi-layered CPM combining cell mechanics with probabilistic rules for cell growth and apoptosis for the two cell types, benchmarked from experimental data. Using this framework, we are beginning to simulate the outcomes of cell competition with precise control over physical, biochemical and decision-making parameters (**Fig. 4D**).

In summary, modeling approaches that combine local cellular mechanics with active decision-making rules would allow us to test the realism, redundancies and efficiencies of different single cell level strategies governing tissue-scale competition dynamics, and therefore to understand the single cell rules governing competition. With precise control over model parameters, multi-layered simulations would enable us to test the relative roles of each of the physical, biochemical and decision-making parameters on the outcome of cell competition. These outcomes could then be compared to high-throughput experimental data using metrics both at the single-cell level and at the population scale (**Section 2.3**). Such a predictive modeling framework would not only account for bulk physical properties but also predict local properties such as topology, organization, and the fate of single cells.

4. Future directions – artificial intelligence and numerical modelling

9

The bridging of scales from population level and bulk properties to single cell level and mechanism is crucial to understanding the underlying biology and physics of cell competition. Machine learning is now being routinely used to accurately identify and label objects in many forms of image data. We have used such methods to identify and label cells in time-lapse microscopy data of competition [26]. However, the more fascinating question is whether advances in Artificial intelligence (AI) can be leveraged to learn a basic set of deterministic rules of cell competition, as they have already done with chess and go [68] and clinical diagnostics [69]. Could one predict the outcome of cell competition if we knew the rules? For example, which physical, topological and biochemical parameters are important in determining the fate of a single cell? Under what circumstances does fate commitment occur? AI approaches could yield significant insight into these questions by taking advantage of the wealth of image data available to learn general models of behavior, and to uncover parameters which may not be immediately obvious to us. Combined with numerical models that enable us to test the foundations of our understanding by simulating outcomes and testing parameters, these quantitative approaches are likely to yield significant new insights into the mechanisms of cell competition.

Acknowledgements

References

1.   Levayer, R. and E. Moreno, *Mechanisms of cell competition: themes and variations.* J Cell Biol, 2013. **200**(6): p. 689-98.
2.   Vincent, J.P., A.G. Fletcher, and L.A. Baena-Lopez, *Mechanisms and mechanics of cell competition in epithelia.* Nat Rev Mol Cell Biol, 2013. **14**(9): p. 581-91.
3.   Merino, M.M., R. Levayer, and E. Moreno, *Survival of the Fittest: Essential Roles of Cell Competition in Development, Aging, and Cancer.* Trends Cell Biol, 2016. **26**(10): p. 776-788.
4.   Morata, G. and P. Ripoll, *Minutes: mutants of Drosophila autonomously affecting cell division rate.* Developmental biology, 1975. **42**(2): p. 211-221.
5.   Garcia-Bellido, A., P. Ripoll, and G. Morata, *Developmental compartmentalisation of the wing disk of Drosophila.* Nat New Biol, 1973. **245**(147): p. 251-3.
6.   Amoyel, M. and E.A. Bach, *Cell competition: how to eliminate your neighbours.* Development, 2014. **141**(5): p. 988-1000.
7.   Moreno, E., K. Basler, and G. Morata, *Cells compete for decapentaplegic survival factor to prevent apoptosis in Drosophila wing development.* Nature, 2002. **416**(6882): p. 755-9.
8.   Moreno, E. and K. Basler, *dMyc transforms cells into super-competitors.* Cell, 2004. **117**(1): p. 117-29.
9.   Tyler, D.M., et al., *Genes affecting cell competition in Drosophila.* Genetics, 2007. **175**(2): p. 643-57.
10. Rhiner, C., et al., *Flower forms an extracellular code that reveals the fitness of a cell to its neighbors in Drosophila.* Developmental cell, 2010. **18**(6): p. 985-998.
11. Yamamoto, M., et al., *The ligand Sas and its receptor PTP10D drive tumour-suppressive cell competition.* Nature, 2017. **542**(7640): p. 246.
12. Levayer, R., B. Hauert, and E. Moreno, *Cell mixing induced by myc is required for competitive tissue invasion and destruction.* Nature, 2015. **524**(7566): p. 476-80.

13. Díaz-Díaz, C., et al., *Pluripotency surveillance by Myc-driven competitive elimination of differentiating cells.* Developmental cell, 2017. **42**(6): p. 585-599. e4.

14. Maitre, J.L., et al., *Adhesion functions in cell sorting by mechanically coupling the cortices of adhering cells.* Science, 2012. **338**(6104): p. 253-6.

15. Bras-Pereira, C. and E. Moreno, *Mechanical cell competition.* Curr Opin Cell Biol, 2018. **51**: p. 15-21.

16. Marinari, E., et al., *Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding.* Nature, 2012. **484**(7395): p. 542-5.

17. Shraiman, B.I., *Mechanical feedback as a possible regulator of tissue growth.* Proceedings of the National Academy of Sciences, 2005. **102**(9): p. 3318-3323.

18. Basan, M., et al., *Homeostatic competition drives tumor growth and metastasis nucleation.* HFSP journal, 2009. **3**(4): p. 265-272.

19. Basan, M., et al., *Dissipative particle dynamics simulations for biological tissues: rheology and competition.* Physical biology, 2011. **8**(2): p. 026014.

20. Levayer, R., C. Dupont, and E. Moreno, *Tissue Crowding Induces Caspase-Dependent Competition for Space.* Curr Biol, 2016. **26**(5): p. 670-7.

21. Wagstaff, L., et al., *Mechanical cell competition kills cells via induction of lethal p53 levels.* Nat Commun, 2016. **7**: p. 11373.

22. Eisenhoffer, G.T. and J. Rosenblatt, *Bringing balance by force: live cell extrusion controls epithelial cell numbers.* Trends Cell Biol, 2013. **23**(4): p. 185-92.

23. Eisenhoffer, G.T., et al., *Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia.* Nature, 2012. **484**(7395): p. 546-9.

24. Gudipaty, S.A. and J. Rosenblatt, *Epithelial cell extrusion: Pathways and pathologies.* Semin Cell Dev Biol, 2017. **67**: p. 132-140.

25. Gudipaty, S.A., et al., *Mechanical stretch triggers rapid epithelial cell division through Piezo1.* Nature, 2017. **543**(7643): p. 118.

26. Bove, A., et al., *Local cellular neighborhood controls proliferation in cell competition.* Mol Biol Cell, 2017. **28**(23): p. 3215-3228.

27. Kocgozlu, L., et al., *Epithelial Cell Packing Induces Distinct Modes of Cell Extrusions.* Current Biology, 2016. **26**(21): p. 2942-2950.

28. Saw, T.B., et al., *Topological defects in epithelia govern cell death and extrusion.* Nature, 2017. **544**(7649): p. 212-216.

29. Tsuboi, A., et al., *Competition for Space Is Controlled by Apoptosis-Induced Change of Local Epithelial Topology.* Curr Biol, 2018. **28**(13): p. 2115-2128 e5.

30. Hogan, C., et al., *Characterization of the interface between normal and transformed epithelial cells.* Nat Cell Biol, 2009. **11**(4): p. 460-7.

31. Norman, M., et al., *Loss of Scribble causes cell competition in mammalian cells.* J Cell Sci, 2012. **125**(Pt 1): p. 59-66.

32. Sakaue-Sawano, A., et al., *Visualizing spatiotemporal dynamics of multicellular cell-cycle progression.* Cell, 2008. **132**(3): p. 487-98.

33. Zhang, J., et al., *Visualization of caspase-3-like activity in cells using a genetically encoded fluorescent biosensor activated by protein cleavage.* Nat Commun, 2013. **4**: p. 2157.

34. Hilsenbeck, O., et al., *Software tools for single-cell tracking and quantification of cellular and molecular properties.* Nat Biotechnol, 2016. **34**(7): p. 703-6.

35. Carpenter, A.E., et al., *CellProfiler: image analysis software for identifying and quantifying cell phenotypes.* Genome Biol, 2006. **7**(10): p. R100.

36. Held, M., et al., *CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging.* Nat Methods, 2010. **7**(9): p. 747-54.

37. Winter, M., et al., *Vertebrate neural stem cell segmentation, tracking and lineaging with validation and editing.* Nat Protoc, 2011. **6**(12): p. 1942-52.

38. Heller, D., et al., *EpiTools: an open-source image analysis toolkit for quantifying epithelial growth dynamics.* Developmental cell, 2016. **36**(1): p. 103-116.
39. Etournay, R., et al., *TissueMiner: a multiscale analysis toolkit to quantify how cellular processes create tissue dynamics.* Elife, 2016. **5**: p. e14334.
40. de Reuille, P.B., et al., *MorphoGraphX: A platform for quantifying morphogenesis in 4D.* Elife, 2015. **4**: p. e05864.
41. Dixon, P.M., A.M. Ellison, and N.J. Gotelli, *Improving the precision of estimates of the frequency of rare events.* Ecology, 2005. **86**(5): p. 1114-1123.
42. Waclaw, B., et al., *A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity.* Nature, 2015. **525**(7568): p. 261-4.
43. Lee, S.W. and Y. Morishita, *Possible roles of mechanical cell elimination intrinsic to growing tissues from the perspective of tissue growth efficiency and homeostasis.* PLoS Comput Biol, 2017. **13**(7): p. e1005651.
44. Nishikawa, S., et al., *Mathematical model for cell competition: Predator-prey interactions at the interface between two groups of cells in monolayer tissue.* J Theor Biol, 2016. **404**: p. 40-50.
45. Banerjee, S. and M.C. Marchetti, *Continuum models of collective cell migration.* arXiv preprint arXiv:1805.06531, 2018.
46. Hakim, V. and P. Silberzan, *Collective cell migration: a physics perspective.* Reports on Progress in Physics, 2017. **80**(7): p. 076601.
47. Gatenby, R.A. and E.T. Gawlinski, *A reaction-diffusion model of cancer invasion.* Cancer Res, 1996. **56**(24): p. 5745-53.
48. Lorz, A., et al., *Modeling the effects of space structure and combination therapies on phenotypic heterogeneity and drug resistance in solid tumors.* Bull Math Biol, 2015. **77**(1): p. 1-22.
49. Williamson, J.J. and G. Salbreux, *Stability and Roughness of Interfaces in Mechanically Regulated Tissues.* Phys Rev Lett, 2018. **121**(23): p. 238102.
50. Cho, H. and D. Levy, *The Impact of Competition Between Cancer Cells and Healthy Cells on Optimal Drug Delivery.* arXiv preprint arXiv:1806.07477, 2018.
51. Podewitz, N., et al., *Interface dynamics of competing tissues.* New Journal of Physics, 2016. **18**.
52. Fletcher, A.G., et al., *Vertex Models of Epithelial Morphogenesis.* Biophysical Journal, 2014. **106**(11): p. 2291-2304.
53. Farhadifar, R., et al., *The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing.* Curr Biol, 2007. **17**(24): p. 2095-104.
54. Brodland, G.W., *New information from cell aggregate compression tests and its implications for theories of cell sorting.* Biorheology, 2003. **40**(1-3): p. 273-7.
55. Honda, H., M. Tanemura, and T. Nagai, *A three-dimensional vertex dynamics cell model of space-filling polyhedra simulating cell behavior in a cell aggregate.* J Theor Biol, 2004. **226**(4): p. 439-53.
56. Spahn, P. and R. Reuter, *A vertex model of Drosophila ventral furrow formation.* PLoS One, 2013. **8**(9): p. e75051.
57. Rauzi, M., et al., *Nature and anisotropy of cortical forces orienting Drosophila tissue morphogenesis.* Nat Cell Biol, 2008. **10**(12): p. 1401-10.
58. Staddon, M.F., et al., *Cooperation of dual modes of cell motility promotes epithelial stress relaxation to accelerate wound healing.* PLoS Comput Biol, 2018. **14**(10): p. e1006502.
59. Nagai, T. and H. Honda, *Wound healing mechanism in epithelial tissues cell adhesion to basal lamina.* WSEAS Transactions on Biology and Biomedicine, 2006. **3**(6): p. 389.
60. Kuipers, D., et al., *Epithelial repair is a two-stage process driven first by dying cells and then by their neighbours.* J Cell Sci, 2014. **127**(6): p. 1229-1241.

61.   Hardin, J. and T. Walston, *Models of morphogenesis: the mechanisms and mechanics of cell rearrangement.* Curr Opin Genet Dev, 2004. **14**(4): p. 399-406.
62.   Basanta, D., et al., *Computational analysis of the influence of the microenvironment on carcinogenesis.* Math Biosci, 2011. **229**(1): p. 22-9.
63.   Szabo, B., et al., *Phase transition in the collective migration of tissue cells: experiment and model.* Physical Review E, 2006. **74**(6): p. 061908.
64.   Basan, M., et al., *Alignment of cellular motility forces with tissue flow as a mechanism for efficient wound healing.* Proc Natl Acad Sci U S A, 2013. **110**(7): p. 2452-9.
65.   Graner, F. and J.A. Glazier, *Simulation of biological cell sorting using a two-dimensional extended Potts model.* Phys Rev Lett, 1992. **69**(13): p. 2013-2016.
66.   Szabo, A. and R.M. Merks, *Cellular potts modeling of tumor growth, tumor invasion, and tumor evolution.* Front Oncol, 2013. **3**: p. 87.
67.   Li, J.F. and J. Lowengrub, *The effects of cell compressibility, motility and contact inhibition on the growth of tumor cell clusters using the Cellular Potts Model.* J Theor Biol, 2014. **343**: p. 79-91.
68.   Silver, D., et al., *Mastering the game of Go with deep neural networks and tree search.* Nature, 2016. **529**(7587): p. 484-+.
69.   De Fauw, J., et al., *Clinically applicable deep learning for diagnosis and referral in retinal disease.* Nature medicine, 2018. **24**(9): p. 1342.
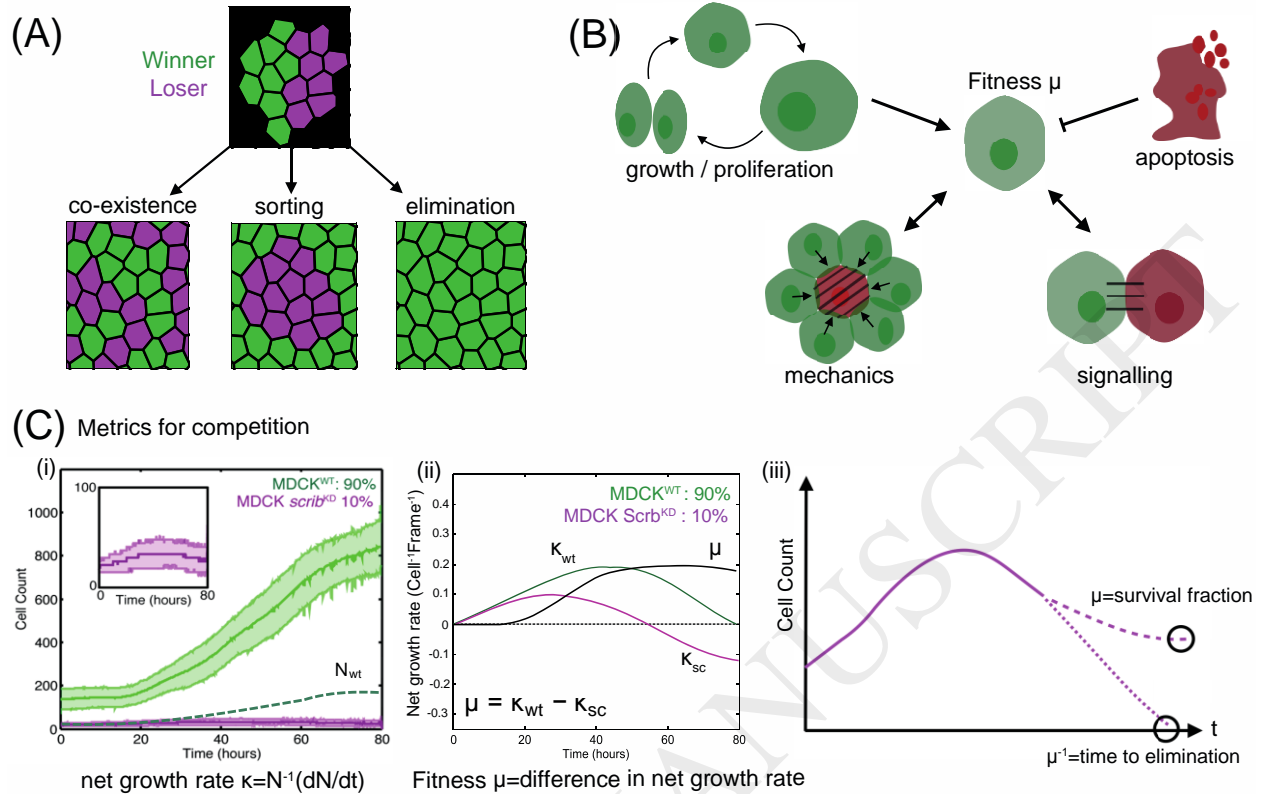
Figure 1: Cell competition outcomes and strategies. (A) Cell competition can result in diverse outcomes depending on cell-cell interaction parameters, including stable co-existence, sorting into colonies, or complete elimination of one cell type. (B) The fitness μ of a cell is increased by factors that increase cell growth/division rate, and is decreased by factors that enhance the rate of apoptosis. Both cell proliferation and apoptosis can be regulated by mechanical forces within the tissue or by short-range biochemical signaling. (C) (i) Fitness for a cell type can be approximated by its net growth rate $\mu=\left(\frac{1}{N}\right)dN/dt$, where $N(t)$ is the number of cells of that type at time t. Figure shows the cell count dynamics for MDCK[WT] (green) and scribble[kd] (purple) in a competitive setting. Dashed curve (green) shows the cell count profile for non-induced MDCK[WT] (N_{wt}) in pure population. Inset: Evolution of cell count for scribble[kd] cells on a smaller scale. Figure adapted from Ref. [26]. (ii) In a competition setting, fitness for a cell type can defined as the difference in its net growth rate from its competitor, i.e. how fast the cell type grows relative to its competitor. The figure shows the time evolution of the net growth rates for the winner (WT) and loser (scribble[kd]) cell types, and also the dynamics of μ, the difference in net growth rate. (iii) Fitness could also be quantified by the time to elimination of the loser cell type or by the survival fraction, as defined in the curve for loser cell count.
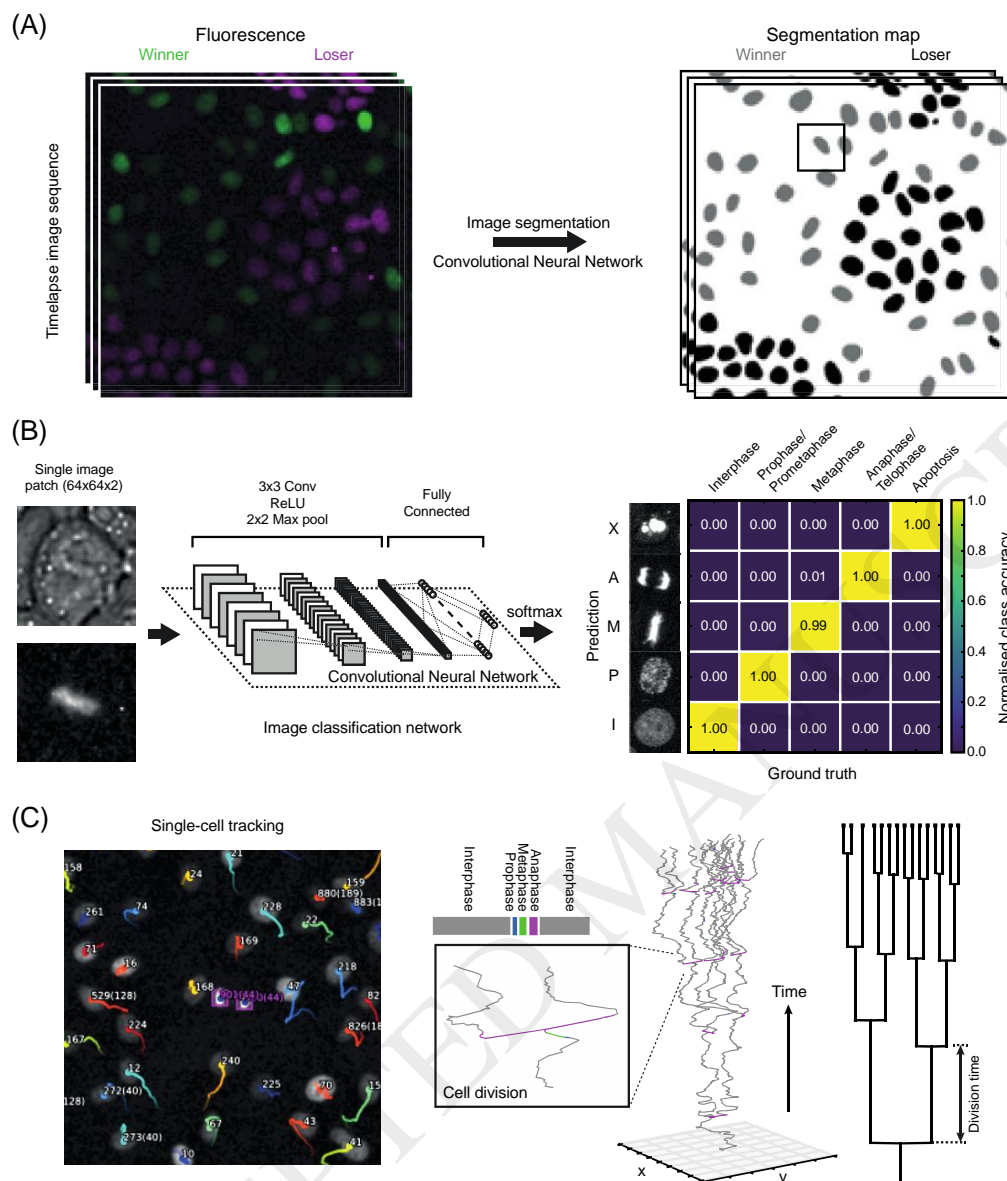
14

(A)

Fluorescence
Winner   Loser

Timelapse image sequence

Image segmentation
Convolutional Neural Network

Segmentation map
Winner   Loser

(B)

Single image patch (64x64x2)

3x3 Conv
ReLU
2x2 Max pool

Fully Connected

softmax

Convolutional Neural Network

Image classification network

Prediction

Interphase Prophase/Prometaphase Metaphase Anaphase/Telophase Apoptosis

	Interphase	Prophase/Prometaphase	Metaphase	Anaphase/Telophase	Apoptosis
X	0.00	0.00	0.00	0.00	1.00
A	0.00	0.00	0.01	1.00	0.00
M	0.00	0.00	0.99	0.00	0.00
P	0.00	1.00	0.00	0.00	0.00
I	1.00	0.00	0.00	0.00	0.00

Normalised class accuracy

Ground truth

(C)

Single-cell tracking

Interphase Anaphase Metaphase Prophase Interphase

Cell division

Time

x   y

Division time

Figure 2: Experimental workflow for high-throughput analysis of cell competition at the single cell scale. (A) High-throughput time-lapse microscopy of mixed cultures of winner and loser cells yields thousands of images per experiment. Each image shows the fluorescent histone markers demarcating cell type (winner or loser). These images can be robustly segmented using computer vision and convolutional neural networks (CNN), allowing the cell type, centroids and cell count to be determined at each time point. (B) A CNN classifier can be trained to accurately identify the mitotic or apoptotic state using image data alone. The table shows the confusion matrix that quantifies the matching of human annotations vs annotations of the CNN system. Correct predictions are located on the diagonal of the matrix. Following completion of training, the CNN possessed an accuracy of 99% for all cell states in the matrix. (C) Using the segmented and classified objects, automated tracking of the cells over time allows the reconstruction of lineage trees. Each lineage tree defines the spatial position, motion, lineage, instantaneous state and fate of every cell in the field of view. Figure adapted from Ref. [26].
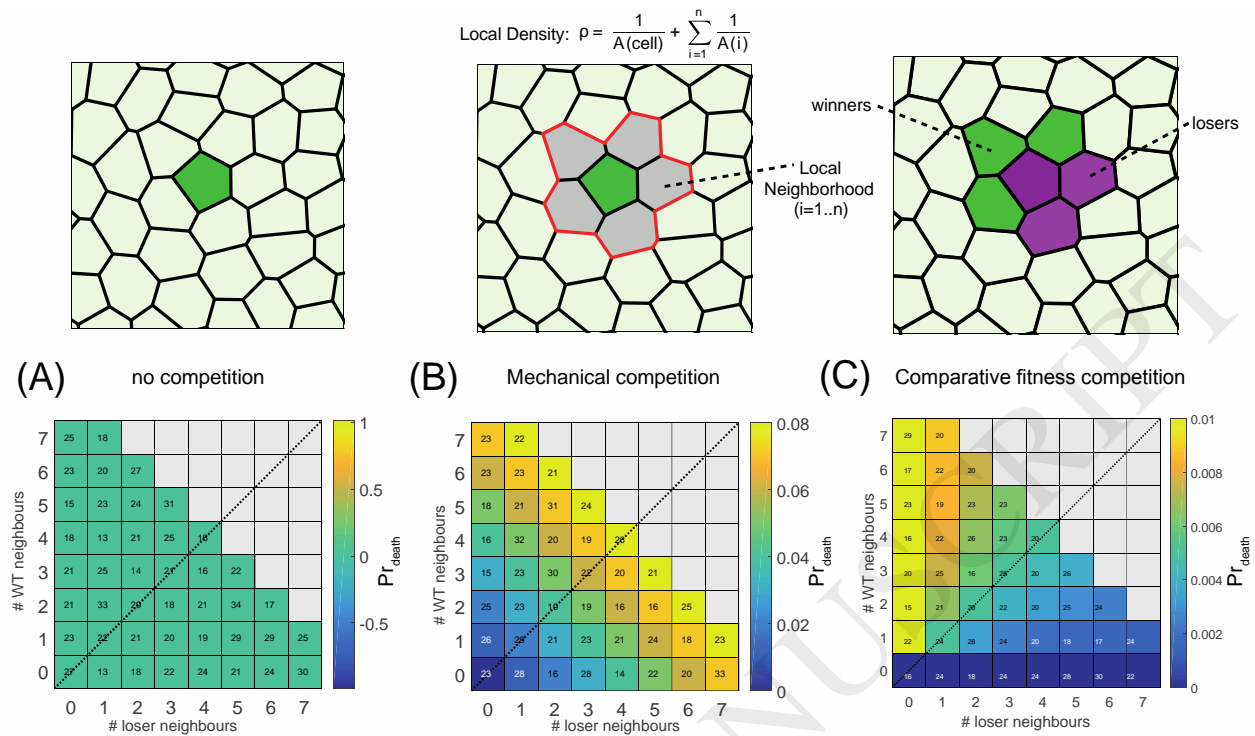
15

Local Density: $\rho = \dfrac{1}{A(cell)} + \sum\limits_{i=1}^{n} \dfrac{1}{A(i)}$

winners

losers

Local Neighborhood (i=1..n)

(A) no competition

(B) Mechanical competition

(C) Comparative fitness competition

WT neighbours

loser neighbours

$\mathrm{Pr_{death}}$

Figure 3: Single cell metrics to characterize competition. The mechanisms of competition at the single cell level can be inferred from neighborhood plots (bottom row) which have the number of winner (wildtype – WT) neighbors on the y-axis and the number of loser neighbors on the x-axis. The probability of an event (cell fate) is color coded and the number of observations is indicated in each grid position. (A) In a no-competition scenario, the probability of a certain cell fate (death) is independent of the number of winner/loser neighbors, resulting in a uniform color map. (B) In pure mechanical competition, the local density of a cell's neighborhood is important, but not the identity of the cell type. In this case, the probability of cell apoptosis will be symmetric about the diagonal, and will increase with increasing number of cells in the neighborhood. Top: Local density of the cell is defined as: $\rho(cell) = \dfrac{1}{A(cell)} + \sum_{i=1}^{n} \dfrac{1}{A(i)}$, where n is the number of neighbors, $A(i)$ the area of neighbor cell i, and $A(cell)$ the area of the cell of interest. (C) In comparative fitness competition, loser cells may possess a higher likelihood of apoptosis when most of their neighbours are winners. As a result, the neighbourhood plot is asymmetric, with higher probabilities of death above the diagonal.
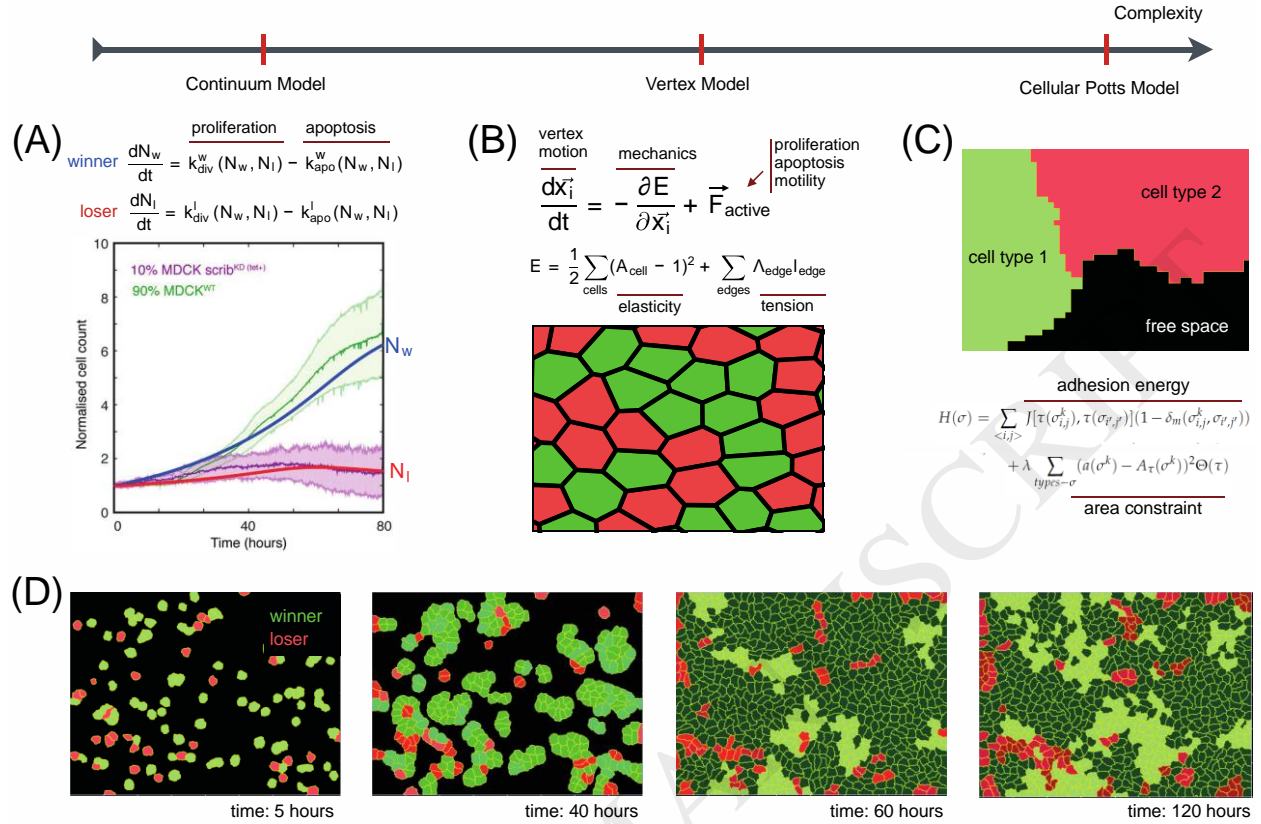
Figure 4: Computational modeling approaches to reverse-engineer cell competition. (A) Continuous model for cell competition based on ordinary differential equations for the rate of proliferation (k_{div}) and apoptosis (k_{apo}) for the winner and loser cells with cell counts N_{w} and N_{l}, respectively [26]. (B) Vertex model of cell competition, where individual cells are represented by polygons in a confluent tissue. Each vertex, x_i, evolves in time at a velocity proportional to the net force acting on it from mechanical and active processes. Mechanical forces arise from the gradient in tissue mechanical energy, E, which is a sum of cell areal elastic energy and line tension on cell edges arising from intercellular adhesions and cortical actomyosin contractility [52]. (C) Cellular Potts Model for cell competition, where each cell is a sub-lattice of pixels, that can take arbitrary shapes and undergo motile behavior based on mechanical forces or stochastic rules. Mechanical forces are derived from a Potts model Hamiltonian, H, which is a sum of intercellular adhesion energies and area constraint for each cell [65]. In the Hamiltonian (bottom), σ is the cell index, i and j are lattice sites, τ is the cell type, J is the adhesion energy at cell boundaries, δ is the Kronecker delta function, a is the current cell area, A is the preferred cell area, and λ is a elastic constant quantifying the strength of area constraint. (D) Snapshots of simulated cell competition dynamics using a multi-layered cellular Potts model, where winner cell types are indicated in shades of green and loser cell types are shown in shades of red. Darker shades indicate higher generation number.