Applied Network Science

Critical behavior of spatial networks as a model of paracrine signaling in tumorigenesis

Philip Tee[1,2,3]* and Allan Balmain[4]

*Correspondence:
phil@moogsoft.com;
ptee2@asu.edu;
p.tee@sussex.ac.uk
[1]Moogsoft Inc,1265 Battery Street,
94111 San Francisco, CA, USA
[2]Beyond Center, Arizona State
University, 85287 Tempe, AZ, USA
Full list of author information is
available at the end of the article

Abstract

Recent work on the study of cell populations in mouse tumors has revealed much about the clonal evolution of cancers from the initiated cell to metastasis. Although most cancers are clonal in origin, genetic instability leads to the emergence of new cell clones, some of which show cooperative behavior during progression to metastasis. The nature of these cell-cell interactions is unclear, and in particular it is possible that their spatial distribution could influence the emergence of fully malignant behavior. The spatial distribution would indicate a subtle dependence of the distribution of these cells and the emergence of malignancy. In this paper we model tumor evolution using dynamically evolving spatially embedded random graphs. The dynamic evolution of Spatio-Temporal graphs is not widely studied analytically, particularly when distance based preferences are included. In this paper we present analysis and simulations of such graphs, and demonstrate that the distance function relative to the mixing of the nodes can combine to create phase transitions in connectivity. This result supports the hypothesis that cell to cell interaction is a critical feature of malignancy in tumors.

Introduction

The development of cancer in mammals occurs in multiple stages leading from single initiated cells to benign or pre-malignant lesions, some of which eventually undergo a transition to malignancy, disseminating to distant sites in a process known as metastasis. These genetic and phenotypic transitions are accompanied (or caused) by acquisition of a series of hallmark traits that distinguish tumor cells from their normal counterparts (Weinberg 2011). Cancer evolution through these stages has been extensively studied both at the genetic and cellular levels, with some models positing a neo-Darwinian process leading to selection of the fittest cell at each stage, resulting in a series of clonal sweeps (Greaves and Maley 2012). This form of linear evolution has been seen in various types of human leukemias and some solid tumors (Venkatesan and Swanton 2016), and also applies to the early stages of initiation and benign to malignant progression of tumors in a mouse model of squamous carcinoma development (Reeves et al. 2018). In the latter model, cellular lineage tracing using a mouse strain that allows individual clones to be identified by their colors, showed that tumors arise from single initiated cells (i.e. they are clonal in origin) but can then generate new cell clones with different mutations, only one of which appears to gain a selective advantage and give rise to the progressing

lesions. These data therefore are compatible with the Darwinian model of survival of the fittest cell at each stage.

Subsequent stages of progression to metastasis are however more complex: continued branched evolution generates additional clones, some of which can interact positively or negatively with each other, resulting in an optimal configuration of synergistically interacting cells that promote metastatic dissemination. This view of cancer as a society of interacting cells has parallels with evolutionary models based in population biology. Sewell-Wright drew attention to this alternative view of evolution in his "shifting balance" theory, based upon natural selection among interactive systems (Wright 1982). The same concepts can be relevant to interactions among cell populations within cancers that could promote progression to full malignancy. Indeed, studies even of bacterial systems growing over thousands of generations in vitro suggest that both Darwinian-based clonal sweeps and population-based cooperation models can be correct, depending on the selective processes operative (Good et al. 2017).

The existence of extensive cellular heterogeneity within cancers has been appreciated for decades, and in particular the ability of specific sub-clones of cells to cooperate functionally to promote growth, invasion, or immune interactions has been experimentally demonstrated many times (Heppner 1984). More recent studies have confirmed many of these concepts (Calbo et al. 2011) and have begun to identify potential mechanisms by which this cooperation can take place (Cleary et al. 2014; Lim et al. 2017; Tammela et al. 2017; Marusyk et al. 2014). An emerging feature of these mechanistic studies is that cells may cooperate through direct cell-cell interactions, or through paracrine interactions by which growth factors or cytokines produced by one cell may promote survival of alternative cells that express the appropriate receptor systems. Clonal interference, by which single clones can inhibit the growth of neighboring cells, can also take place, suggesting that an optimal "sweet spot", determined by the ratio of interacting clones and their spatial relationships, may be necessary to ensure progressive tumor growth (Marusyk et al. 2014). These concepts of paracrine signaling at a distance and clonal competition have their roots in normal developmental biology, and are well known to biologists familiar with control of tissue patterning during embryogenesis (Tabata 2004; Moreno et al. 2014). Some growth factor-receptor interactions require direct interactions, while others can operate within a concentration gradient over multiple cell diameters. However we presently do not know if, and to what extent, similar mechanisms are operative in cancer evolution, and to our knowledge no attempts have been made to mathematically model such interactions using network theory.

The use of mathematical models in the study of cancer evolution is not unique. In particular the use of competitive network models has been the subject of much study, motivated from social, economic and game theory driven models of population evolution. An excellent review of these approaches can be found for example in the work of Perc and Szolnoki (2010). In addition, applications of chaos theory (Khajanchi et al. 2018), seek to model the fine balance between the immune system and cancer evolution by analyzing instabilities in the systems of coupled differential equations that model this fine balance. As such these techniques directly introduce the concept of an unstable equilibrium that can tip between a stable benign tumor and an unstably growing malignant one. In this work though we seek to explore mathematical models for such interactions using the tools of network science.

We start with a simple network model of two populations of cells, where the presence of an interaction between two cells is modeled by the presence of an edge in a graph where the nodes are the cells. In such a network model we can calculate and simulate various conditions on the presence of an edge, importantly capturing the relative populations and geometries of the two cellular populations. From there we can ask under what conditions are there widespread interactions between cells, by observing when a random cell in one population is likely to have a connection with a cell in the other as we vary average distance between the two cells. Given it is known that cell population admixture is characteristic of advanced malignancy, and may be essential for metastatic spread to distant sites (Venkatesan and Swanton 2016; Reeves et al. 2018), we would expect to see critical behavior whereby the connectivity of the graph fundamentally changes according to changes in population parameters of the graph evolution.

We discuss this in "Experimental evidence for symbiosis" section. In particular, we focus on the average distance between members of two cell populations, as experimentally that can be measured, providing we assume that members of a given cell population are at least approximately uniformly distributed in space. This assumption appears to be born out at least approximately, and can be seen in Fig. 3.

In models of evolving networks such phase transitions are well studied, in particular in the treatment by Barabási and Bianconi (2000; 2001) of Bose-Einstein condensation in networks. These networks are not spatially dependent, with the physical position of the nodes not considered. There are spatial random graph models, surveyed in (Barthélemy 2011), but these have mostly relied upon numerical simulations (Rozenfeld et al. 2002; Barthélemy 2003; Manna and Sen 2002). Models of Spatial networks that exhibit phase transitions have also been studied, (Sen et al. 2002; Ostilli and Bianconi 2015), but these models do not consider any deeper structure such as distinct populations of nodes. In the social sciences the role of geographical position on competition networks has been more extensively studied (Braha et al. 2011), but again the treatment has focused upon simulation rather than analytically solvable models. It is important to stress that the network model used is strictly a representation of the interaction of the cells in a dynamic environment, and the evolution of the network simply a method to investigate the role that the nature of the interaction has on the extent of the interaction. In short, we seek to demonstrate that if fitness for cooperation is governed by distance there exists a critical average distance between cells, below which, significant symbiosis cannot occur. This is represented by distinct changes in the distribution of links between two distinct populations of nodes.

In previous work (Tee and Balmain 2019), we presented some basic results from a model built to capture the characteristics of cellular symbiosis. In this work we provide more details on the network science underpinning the model, and the biology that provided the motivation. We also consider in "Conclusion and future work" section how such a model could be tested against experimentally prepared cultures of cancer tissue. Further we present some new results concerning the likely presence of constraints in the ability of a cell to 'service' multiple other cells of a different population and speculate at which point this would present a brake on the transition of the model from the benign disconnected state to the malignant smoothly connected one.

Experimental evidence for symbiosis

The motivation for this work derives from the results of experiments tracking the clonal evolution of cancers in humans and mouse models. In particular, the observation that although tumors often originate, and, are initially dominated by cells of a single clonal origin (Reeves et al. 2018), the malignant phase is characterized by the acquisition of new mutations in the original clone, resulting in emergence of multiple cell populations, each carrying distinct genetic alterations. Experimentally, these observations are based on Doexyribonucleic Acid (DNA) sequencing studies which have demonstrated the existence of "trunk" mutations that are present in every tumor cell and therefore existed in the original dominant clone, and sets of mutations that are only seen in subpopulations of tumor cells that evolved later as part of the branched evolutionary process (Venkatesan and Swanton 2016). Some of these sub-clones could be detected in primary tumors growing at their initial tissue site, but also in metastases from the same tumor, suggesting that multiple cell populations were involved in metastatic dissemination. These studies based on sequencing however do not provide any information on spatial relationships between clones that co-exist in the same tumors. Such information could however be obtained from mouse models in which individual cells can be labelled with one of four colors using a reporter cassette (the "confetti" mouse (Reeves et al. 2018)). In this mouse strain, individual cells can be permanently labelled at any time by activating an enzyme (Cre-recombinase) that randomly switches on these colors in adult tissues or tumors. Since the cells in which the recombination has taken place are now labelled with a specific color, and easily seen using a fluorescence microscope, their evolution and migration can be followed over time during tumor progression as the cells continue to mutate and proliferate. Upon treatment with Tamoxifen, these colors become visible under microscopy, and can then subsequently be used to track the evolution of cellular populations. Mice, thus treated, are referred to as 'confetti' mice. The mouse model used to trace the evolution of tumors from the initiated state to full metastasis are depicted graphically in Fig.1a and b (reprinted from Reeves et al. (2018)).

Figure 1a shows the confetti cassette inserted into the ROSA-26 locus in the mouse genome, with the 4 available fluorescent protein cDNAs transcribed using a strong CAGG promoter. Treatment of the mice with a hormone (tamoxifen) activates expression of the proteins in such a way that each cell only expresses one single color. Figure 1b shows the results of activating these colors in a benign tumor on the skin of a confetti mouse, showing the multiple colors activated within tumor cells. The image was taken one week after activation, and shows that some of the initially activated cells have divided, giving rise to small groups of similarly colored cells.

As discussed in the introduction, using this technique it is possible to track the clones that participate in development of a tumor from an initiated stem cell, and their progression to malignancy. Figure 2 provides a graphic illustration of the conclusions reached using this model as applied to the initial clonal origin of tumors, and their progression to metastasis. Figure 2a shows the development of a benign tumor (papilloma) from a single initiated cell. The initiated clone is shown by the blue color at the clonal origin, and the yellow, red and green clones represent clones that have arisen from this original clone by generation of new mutations. Some minor clones (called "streaks" in this graphic) are not in fact tumor clones as they do not carry the necessary genetic mutation that initiates tumors in this model. These streak cells arise from normal stem cells, and do not
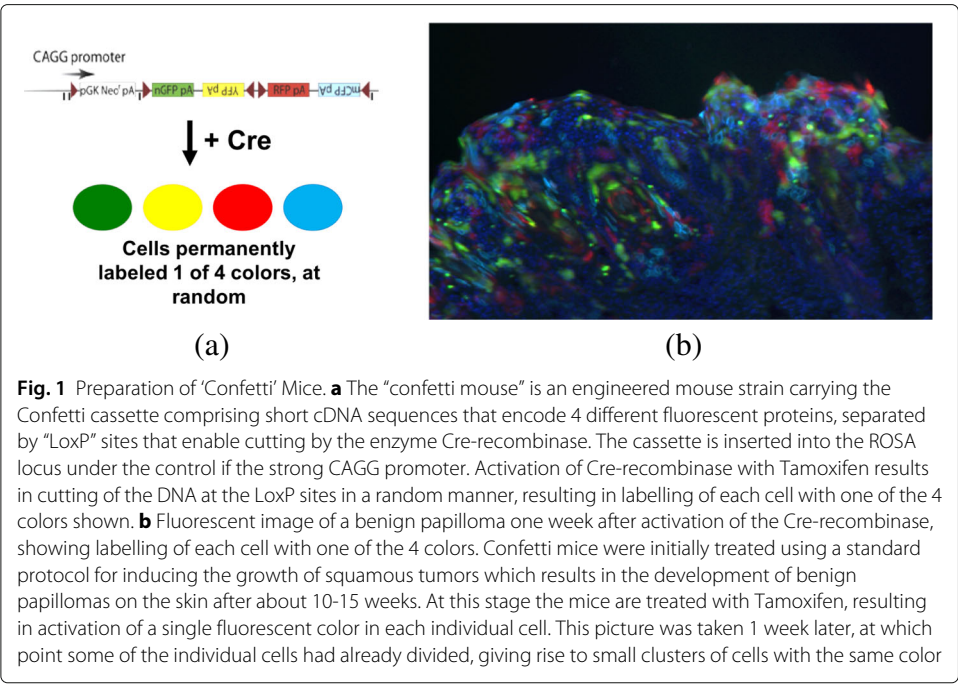
Fig. 1 Preparation of 'Confetti' Mice. **a** The "confetti mouse" is an engineered mouse strain carrying the Confetti cassette comprising short cDNA sequences that encode 4 different fluorescent proteins, separated by "LoxP" sites that enable cutting by the enzyme Cre-recombinase. The cassette is inserted into the ROSA locus under the control if the strong CAGG promoter. Activation of Cre-recombinase with Tamoxifen results in cutting of the DNA at the LoxP sites in a random manner, resulting in labelling of each cell with one of the 4 colors shown. **b** Fluorescent image of a benign papilloma one week after activation of the Cre-recombinase, showing labelling of each cell with one of the 4 colors. Confetti mice were initially treated using a standard protocol for inducing the growth of squamous tumors which results in the development of benign papillomas on the skin after about 10-15 weeks. At this stage the mice are treated with Tamoxifen, resulting in activation of a single fluorescent color in each individual cell. This picture was taken 1 week later, at which point some of the individual cells had already divided, giving rise to small clusters of cells with the same color

participate in progression of the benign tumor to malignancy. Figure 2b shows that of the three new clones in the original papilloma, the red clone has undergone a clonal sweep, and has taken over the entire carcinoma at the stage of progression (dotted line). However, as seen in the papilloma, this dominant new clone also undergoes additional evolution, giving rise to new clones (shown in green, yellow and blue). However in contrast to what is seen in the benign tumors, the carcinoma clones begin to intermix (shown by the dotted colors), showing signs of inter-clonal cooperation. These intermixing cells can then
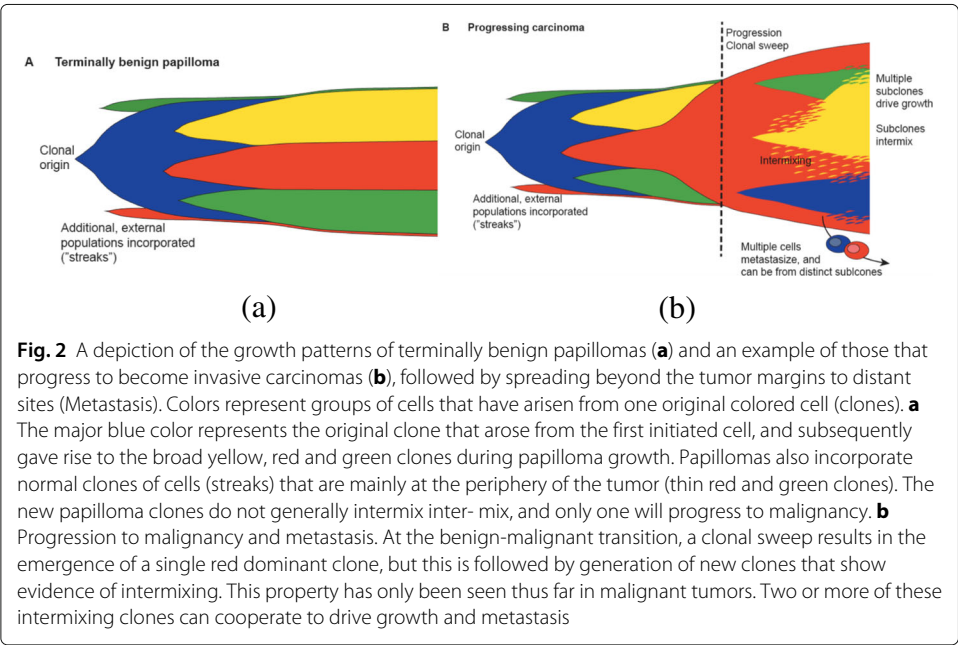

Fig. 2 A depiction of the growth patterns of terminally benign papillomas (**a**) and an example of those that progress to become invasive carcinomas (**b**), followed by spreading beyond the tumor margins to distant sites (Metastasis). Colors represent groups of cells that have arisen from one original colored cell (clones). **a** The major blue color represents the original clone that arose from the first initiated cell, and subsequently gave rise to the broad yellow, red and green clones during papilloma growth. Papillomas also incorporate normal clones of cells (streaks) that are mainly at the periphery of the tumor (thin red and green clones). The new papilloma clones do not generally intermix inter- mix, and only one will progress to malignancy. **b** Progression to malignancy and metastasis. At the benign-malignant transition, a clonal sweep results in the emergence of a single red dominant clone, but this is followed by generation of new clones that show evidence of intermixing. This property has only been seen thus far in malignant tumors. Two or more of these intermixing clones can cooperate to drive growth and metastasis

metastasize to distant sites, where they seed to form growing metastatic lesions that will eventually kill the host.

Figure 3 shows examples of the data upon which these conclusions are based. Figure 3a shows an image of a papilloma that was take a few weeks after the initial labelling. Of the many different colors seen one week after the labeling process (as shown in Fig. 1a), the papillomas now has only a small number of clones, each having arisen from a single cell in the papilloma. These distinct lobes of color persist throughout the lifetime of papillomas that do not progress to malignancy. Figure 3b shows a carcinoma from a similar mouse in which only one color is seen. Comparison of tumors from many animals showed that carcinomas arising from pre-existing benign multicolor papillomas invariably have only one major clone representing the clonal sweep. As shown in Fig. 2b, new clones emerge in carcinomas that can be seen if the Cre-recombinase is activated in mice with existing carcinomas. Figure 3c shows an image of a carcinoma with two distinct red and green cell populations seen both in the primary carcinoma on the skin, and in a metastasis to an adjacent lymph node, showing that at least two distinct cell populations participated in metastasis.

We speculate that these different colored clones are cooperating either using paracrine signaling or an alternative mechanism that is essential for the tumor to enter the metastatic phase, and that the degree of mixing and separation between cellular populations could be a critical feature of the progression of the disease. It is this idea that suggests the well understood behavior of evolving complex networks that involve some notion of 'fitness' as a factor in the propensity for a node in the graph to attract connections. The fitness model introduced by Barabási and Bianconi (2001) has been long understood to possess critical behavior, and exhibit phase transitions. This model builds upon standard scale free models, where nodes accumulate connections according to the relative number of links they have already accumulated. Already, in a biological model of cooperation, this preferential attachment behavior captures an important feature in that cells involved in cooperation could be more attractive for further additional cooperation. However in the absence of a fitness parameter this would lead to a pure scale free model where there is no fundamental change to the network topology, that is the network will always exhibit a small number of nodes dominating the connectivity of the graph. The fitness model however, introduces a factor that works
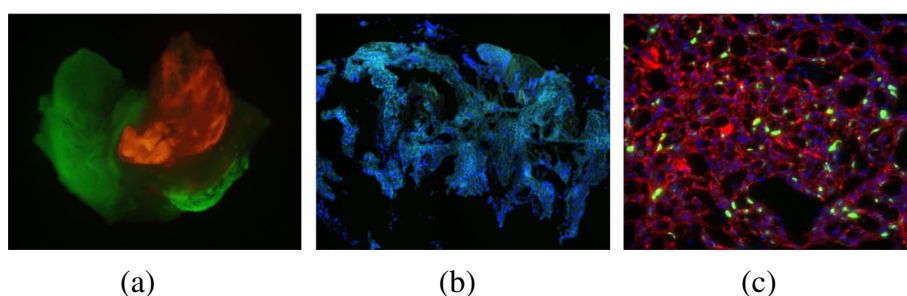

(a)   (b)   (c)

Fig. 3 Images of tumors from confetti mice at different stages of tumor development. **a** A benign papilloma that has retained distinct clones with no evidence of intermixing. **b** In the malignant carcinoma, one clone emerges to dominate in a clonal sweep. **c** After malignant conversion, new cell clones are generated (red and green cells) which show evidence of clone intermixing. This may facilitate the cell-cell interactions that eventually promote metastasis

alongside the number of links a node has to indicate a node's suitability for new links. This can lead to newer nodes dominating the connectivity of the network, something that can never happen in standard scale free networks. Crucially, these models also have distinct phases, as governed by an order parameter, whereby the network transitions from the so-called 'winner takes all' phase where one node captures substantially most of the links to a more normal random, or scale free distribution. In our treatment this is a transition from most nodes not being in effective contact with a partner in the other population to a phase where that is common. Critically the scale free nature, and the critical behavior of the network go hand in hand, and our model seeks to capture both.

In essence we model the cell to cell interaction as a graph, with the nodes representing an individual cell, and a link between nodes indicating that the cells are interacting to promote tumorous behavior. In particular we model the tumor as two distinct types of spatially embedded nodes, A and B, and the situation where the resultant graph has a direct connection between each cell type A to at least one member of the cell type B collection would indicate that symbiosis is occurring and that the tumor is in the phase where symbiosis leads to tumor dissemination.. The opposite is also true, if a large proportion of the two populations do not have interacting partners then we would equate that to the more benign phase. This model then allows us to investigate if there are parameters of the evolution of the graph with distinct phases that mirror the observed phenomenology. Importantly, the phase transition which brings most nodes into functional contact with each other would 'activate' the symbiosis is a widespread manner and we believe promote the tumor's ability to grow in an uncontrolled manner in a way associated with malignancy. It is interesting to note that this cell to cell contact, which may be direct or indirect, is also seen in metastasis, which would seem to indicate its causal role in uncontrolled tumor dissemination. In particular, if spatial arrangement and interaction range cause very different outcomes in the connectivity of the two populations it could support the hypothesis that the mixing of populations and their interactions are causal in the emergence of metastatic tumors.
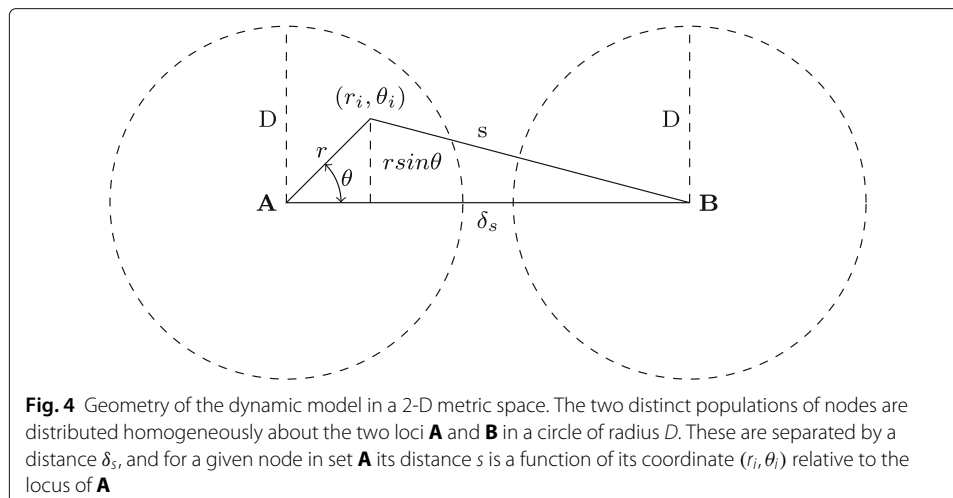
The Spatio-Temporal model

Design of the Spatio-Temporal network model

We begin by identifying the key characteristics of the biological system that we wish to model. We seek to identify when two populations of cells are connected in the sense that they are in effective contact via some as yet unknown chemical or enzymatic mechanism. This maps nicely to a bi-partite graph model between two sets of vertices with the presence of an edge between two nodes representing the fact that the two nodes are in causal contact with each other. The population of cells is dynamic and increasing, which we can model by adding a proportion of the two types of the nodes at a fixed time interval or time step. For simplicity we fix the ratio at 1:1, though we could easily extend our model to have any arbitrary ratio.

A key feature that we seek to capture is the role of physical distance between any two cells, which requires our evolving graph to be spatially embedded. Although tissue is manifestly three dimensional, we can simplify matters by restricting ourselves to a two dimensional space, and not sacrifice the critical features of the model. Generalization to

three dimensions is an area of future work, along the lines of Dall and Christensen (2002). Modeling the inter-cellular distance of each added node as a probability distribution is a difficult analytical task, and one that would add considerably to the complexity of the resultant model. We seek to produce an analytically soluble model, so in the interest of achieving this, we simplify the system by representing the position of one of the populations as at a fixed point, the center of its random distribution. The symmetry of the set up encourages us that the resultant analysis should capture the essential features of the network simulation. Further analysis of the model will investigate the possibility of a solution that includes the exact cell to cell distance. To introduce the concept of average separation of the two cellular populations, we use two independent random point processes, creating distributions of nodes about two loci, separated by a distance δ_s (see Fig. 4). We can control the overlap and therefore average separation of two randomly chosen nodes by shrinking or growing δ_s. As we shall demonstrate later, this parameter, along with an effective interaction range, acts as the critical order parameter for the model. We seek also to capture the concept of the activity of a cell, which we model as the number of other cells it is in causal contact with. In the familiar style of preferential attachment we will hypothesize that cells more actively interact with highly active cells, that is with nodes in the graph of higher degree. If the physical mechanism for symbiosis is some form of chemical or enzymatic interaction, there is of course some limit to the ability of one cell to 'service' other cells. In other work (Tee et al. 2017) we have modeled the effect of such a constraint on a dynamically evolving network, and although our base model does not include such a top constraint, we consider the effect of that in "Constraints in the dynamic model" section.

With these considerations. our network model is built upon the standard foundations of Spatial networks (Barthélemy 2011), and is constructed on a 2-Dimensional metric space equipped with a Euclidean metric, but with no boundary. The removal of the boundary simplifies the analysis, but in principle could be reintroduced following standard periodic proscriptions. The reason for its absence is to prevent boundary conditions from interfering with the distribution of the nodes, and in later work it would be instructive to investigate how boundary effects could affect the critical behavior of the network.

Fig. 4 Geometry of the dynamic model in a 2-D metric space. The two distinct populations of nodes are distributed homogeneously about the two loci **A** and **B** in a circle of radius D. These are separated by a distance δ_s, and for a given node in set **A** its distance s is a function of its coordinate (r_i, θ_i) relative to the locus of **A**

We establish on this metric space two loci A and B, at coordinates (x_A, y_A) and (x_B, y_B). Further we numerically equate the two coordinates, so that $x_A = y_A = A$, and $x_B = y_B = B$. This simplification recognizes that the fundamental property is the distance between the two loci, that is effectively unidimensional. Around these two loci we operate a homogenous point process to distribute nodes around the loci A, B in a disk of radius D. These two loci represent the populations of cell types A and B, and at any point in the evolution of the network we call the collection of nodes surrounding A as the set **A**, and those around B as **B**. In Fig. 4 we describe the essential features of the geometry of the model. In particular the loci are separated by a distance δ_s, which, although depicted as such in the diagram, does not have to separate the two sets of nodes into distinct and spatially separate communities. That is $\delta_s < 2D$ is permissible, although we do not model values of $\delta_s < 0$. Taking advantage of the circular symmetry of the model, we can label each point in the set by polar coordinates (r, θ) centered at A, B. In these coordinates the probability density function function for the addition of nodes is denoted $\rho(r_i, \theta_i)$, and in most of our treatment we assume that this is a uniform distribution, which has full angular symmetry. In polar coordinates a uniform distribution has Probability Density Function (PDF) $\rho(r, \theta) = \rho(r) = 2r/D^2$. This naturally leads to the concept of the distance s between a vertex $a \in \mathbf{A}$, and the locus of the set **B**.

We begin the model by initializing the network with m_0 nodes, added to the set **A**, distributed according to the uniform PDF $\rho(r, \theta)$, that is an *homogenous point process.* In contrast to the normal setup of preferential attachment simulations, these nodes are initially disconnected. As each point a_i is added to the set **A**, however, its spatial position is fixed and so its distance to the locus of the collection of nodes in set **B**. This distance captures the average distance between a randomly chosen node of type A to a node of type B. As we vary the inter-loci distance this average separation varies too, and at a fixed node density it is the only parameter that governs that distance. It is essentially equivalent to fixing the loci and varying the ratio of cell types, which is ultimately the route to experimentally verifying the model, which we discuss in the "Conclusion and future work" section. This fact permits us in the later analysis to introduce the spatial dependence in a novel manner, and makes the model amenable to phase transition analysis similar to that outlined in Barabási and Bianconi (2001). In our model, we do not immediately connect nodes in accordance to a link proscription as is traditional in Spatial networks, but rather consider a dynamic process where nodes are added at to the sets **A**, **B** and as nodes are added to set **B** links are established to set **A**, somewhat similar to the model described in Barthelemy's dynamic model of spatial networks (Barthélemy 2003). Links are only permitted between nodes in different sets, so the end result of the model is a randomly generated, dynamic bi-partite graph, again in contrast to most dynamically generated network models, where nodes are all considered equivalent when assessing the probability of acquiring links. From the initial seed of m_0 nodes in **A**, the model is generated according to the following steps:

1. At each unit time step a new node b_j is added to the set **B** at coordinate $\left(r_i^b, \theta_i^b \right)$. As the node is added to the set **B**, its spatial location is 'quenched' from the probability distribution $\rho(r, \theta)$ in the disk radius D centered at the locus B.

2. As each node b_j is added node is added, it allowed to make m connections to nodes in the set **A**.

3.   The probability that a randomly chosen node $a_i \in \mathbf{A}$ acquires a link is defined as:

$$\Pi_i = \frac{f(s_i)g(k_i)}{\sum\limits_j f(s_j)g(k_j)} \tag{1}$$

where s_i is the euclidean distance between the node a_i and the locus of set $\mathbf{B}$, and f, g are potentially general functions of this distance and the nodes degree k_i respectively.

4.   After the node b_j is connected to its valence m, a new node of degree $k = 0$, is added to the set $\mathbf{A}$, according to the same spatial probability density function $\rho(r, \theta)$ as the nodes in set $\mathbf{B}$. The 'quenched' distance of this node to the locus of the set $\mathbf{B}$ is used via the function $f(s_i)$ to set the 'fitness' of the node to incoming connections to new nodes in the set $\mathbf{B}$. At this point the network has $m_0 + 2t$ nodes, and $2mt$ edges in total across both sets.

This process then repeats, theoretically indefinitely, although for the purposes of simulation we arrest the growth at specific times.

Scale free nature of the model

Before investigating the behavior of this model for non-trivial distance functions $f(s_i, r_c)$, it is worth verifying the perhaps unsurprising result that this model generates a scale free graph for constant distance functions, and the most simple form of linear dependence upon degree. As nodes in the set $\mathbf{A}$ are initially disconnected, to allow for any growth, the model must have a constant added to the degree. This is similar to the model originally proposed by Dorogovtsev et al. (2000), and is necessary to prevent nodes being isolated eternally with zero attachment probability. For simplicity we chose this constant to be 1, and the attachment probability collapses to:

$$\Pi_i = \frac{k_i + 1}{\sum\limits_j (k_j + 1)}. \tag{2}$$

Following the normal program of continuum analysis, and only considering nodes in the set $\mathbf{A}$, we can write a differential equation for the time evolution of the degree of these nodes. Firstly we note that $\sum_j (k_j + 1) = m_0 + (m+1)t \approx (m+1)t$, as $\mathbf{A}$ only captures half-edges. In the more normal model $\sum k_i = 2mt$, but half of the degree sum is excluded as we are only considering the set $\mathbf{A}$. This yields the degree evolution differential equation:

$$\frac{\partial k_i}{\partial t} = m\Pi_i = \frac{\alpha(k_i + 1)}{t}, \text{ where} \tag{3}$$

$$\alpha = \frac{m}{m + 1} \tag{4}$$

Applying the initial condition that at $t = t_0$, the time of introduction of a node into set $\mathbf{A}$, $k_i = 0$, this is easily soluble to give for the degree time evolution equation:

$$k_i(t) = \left(\frac{t}{t_0}\right)^{\alpha} - 1. \tag{5}$$

From here, we can derive the degree distribution by arguing that the probability of a node having degree less than k, is dictated solely by the choice of t_0. This argument yields a cumulative distribution as any time t_0, up to time t is equally probably, and so:

$$P(k(t) < k) = P\left(t_0 > t\left(k+1\right)^{-1/\alpha}\right) \tag{6}$$

$$= 1 - P\left(t_0 \leq t(k+1)^{-1/\alpha}\right) \tag{7}$$

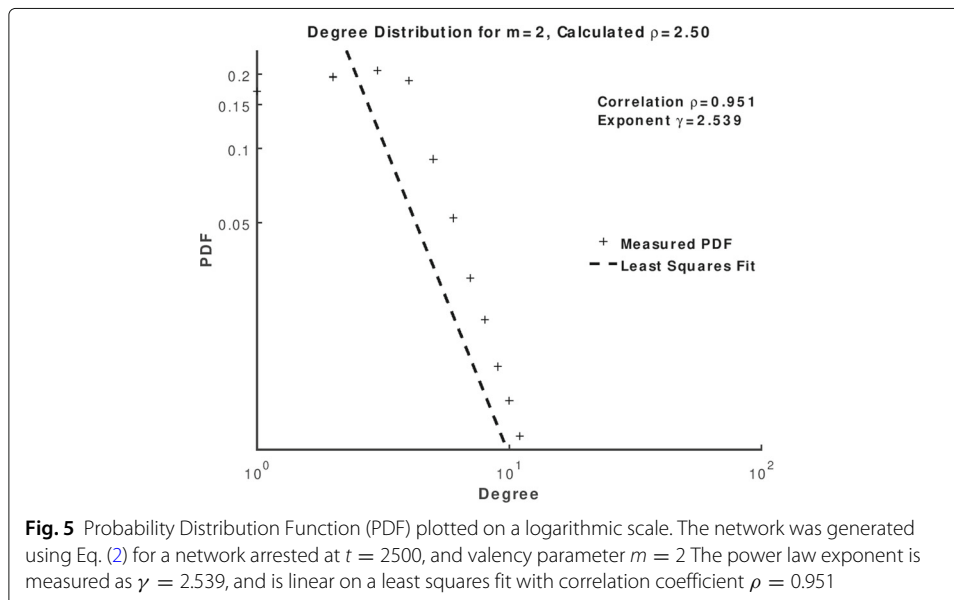$$= 1 - \frac{t}{(m_0 + t)(k+1)^{1/\alpha}} \tag{8}$$

When we differentiate this cumulative probability function, and noting that strictly $\alpha < 1$, we obtain the familiar form of a scale free degree distribution:

$$P(k) = \frac{t}{\alpha(m_0 + t)(k+1)^{(1/\alpha+1)}} \tag{9}$$

It is interesting to note that as $\frac{1}{2} \leq \alpha < 1$, the power law exponent of this model falls between 2 and 3. Recent detailed studies of real world networks (Broido and Clauset 2018; Tee et al. 2017) indicate that when scale freedom is observed in real networks, the power law exponent is generally in this range and lower than the classical result of 3. In Fig. 5 the simulation is arrested at $t = 2500$ for $m = 2$, and we plot the degree distribution on a log/log scale. Against the experimental data we also present a linear least squares fit for the PDF. As we are less interested in the degree distribution behavior of the model, than its critical point behavior we did not deploy the rigorous tests of scale freedom described in the standard approach by Clauset et al. (2009). Nevertheless the simulation is in broad agreement with the analysis that a scale free distribution is obtained, with a measured exponent of $\gamma = 2.54$. Our analysis predicts a value of $\alpha = 2.5$ for $m = 2$, and the PDF is broadly compatible with this value. It is not the intention of this work to propose a different model of preferential attachment, but we can be satisfied that the simulated model is at least consistent with the standard result before we introduce a more complex distance function.

Introducing distance functions

The model as originally posed is crucially dependent upon the functions f, g, but is essentially generic. The specific behavior we are seeking to model are the circumstances where nodes in the set **A** have degree distributions that exhibit distinct phase transitions. We

Fig. 5 Probability Distribution Function (PDF) plotted on a logarithmic scale. The network was generated using Eq. (2) for a network arrested at $t = 2500$, and valency parameter $m = 2$ The power law exponent is measured as $\gamma = 2.539$, and is linear on a least squares fit with correlation coefficient $\rho = 0.951$

consider a phase transition to have occurred where the degree distribution $P(k)$ has distinct and different properties. These phase transitions are well studied and understood in scale free networks (Barabási and Bianconi 2001; Dorogovtsev et al. 2007; Bianconi 2017), and are characterized by percolation phenomena such as the emergence of a giant component which captures a significant proportion of the nodes, and sharp discontinuities in the degree distribution sometimes characterized by the transition between 'rich get richer' to 'winner takes all' distributions. In all of these critical phenomena there is an order parameter that when varied triggers the transition between these phases. In traditional statistical thermodynamic models this is typically performed by the temperature (in fact inverse temperature $\beta = \frac{1}{T}$), but of course in our model we do not have a direct analog of thermodynamic temperature. Further, it should be stressed that this model argues by analogy to statistical thermodynamics, and we do not propose that we have a classical equilibrium model. In fact the network is evolving in time, and is not a thermodynamic system in equilibrium. However, the fundamental stochastic nature of the model is similar and, as we shall show, exhibits a number of characteristics of such systems. Instead, it is the geometry which supplies a candidate order parameter. We consider a characteristic length of the distance function f, which we denote r_c, so that our distance function becomes $f(s, r_c)$ (in fact the order parameter r_c is also dependent upon the distance between the loci δ_s and it is the ratio of the two quantities that control the phase transition behavior of the system). This choice of candidate parameter has profound effect on the behavior of the model. In the first instance, it rules out the normal and most basic Random Geometric Graph (RGG), as a 'hard' connection model in which $f(s, r_c) = \Theta(D - s_i)$, essentially replicates the standard results of variants of the preferential attachment models. This is easily seen until $D > \langle s_i \rangle$ no links will be made, and once $D > sup(s_i)$, the attachment is a function solely of k_i. We have already demonstrated in the previous section that this form of connection function yields a scale free network.

For the purposes of this investigation we chose a very specific choice of f, and g, namely:

$$f(s_i, r_c) = exp\Big(-\beta s_i\Big) \tag{10}$$

$$g(k_i) = k_i + 1, \text{ where} \tag{11}$$

$$\beta = \frac{1}{r_c}, \text{ and it is assumed } \delta_s \text{ is fixed.} \tag{12}$$

Our choice of distance function $f(s_i, r_c)$, is similar to that used in wireless models (Dettmann and Georgiou 2016; Georgiou et al. 2013) which models the physical drop of signal strength with distance. Such an exponential drop off is a characteristic of diffusion processes which are likely to be present in the biological model if the signaling agent is a molecule exchanged between the cells. The distance between elements of **A** and **B** is controlled by the characteristic scale factor r_c, which in the case of chemical signaling would be governed by molecular weight, temperature and ionic content of the intercellular fluid. As we discuss in "Conclusion and future work" section this maybe critical in the practical utility of the model. It will also be shown in the theoretical treatment, that this exponential form of dependence in Eq. 10, naturally admits Bose-Einstein condensation behavior.

The choice of $g(k_i)$ is somewhat more straightforward, and follows the analysis in "Scale free nature of the model" section. As noted before, preferential attachment models are usually linear in k_i, but this would result in no link creation as the nodes $a_i \in \mathbf{A}$ are

'born' with degree zero. In this case the attachment probability is identically zero for all nodes (assuming a finite and non singular value for $f(r_i, r_c)$) and the network would not grow. Again it is possible to imagine more complex functions, and indeed in earlier work (Tee et al. 2017) this has motivated a more in depth analysis of real world networks. For the purposes of this study, we have constricted ourselves to Eq. 11.

Critical behavior and phase transitions

Using Eqs. (10), (11), we write for the total attachment probability the following expressions:

$$\Pi_i = \frac{e^{-\beta s_i}(k_i + 1)}{Z(t)}, \text{ where} \tag{13}$$

$$Z(t) = \sum_j e^{-\beta s_j}(k_j + 1). \tag{14}$$

The time varying quantity $Z(t)$ plays the role of the partition function at time t, and normalizes the probability of attachment. As each node is added the quantity s_i is a random variable, being related to the coordinate of the node i, (r_i, θ_i), which itself is chosen from the probability distribution $\rho(r, \theta) = \rho(r)$, which as stated before has no dependence on θ. From Fig. 4 we can obtain the following relationship between s_i and r_i,

$$s_i^2 = r_i^2 sin^2\theta + (\delta_s - r_i cos\theta)^2, \tag{15}$$

$$= r_i^2 + \delta_s^2 - 2r_i\delta_s cos\theta. \tag{16}$$

As r_i is a random variable selected from the distribution $\rho(r)$, because s_i is a function of r_i, it is also a random variable drawn from a distribution $\rho(s)$. This distribution can be derived from $\rho(r)$ by using Eq. 16, and standard change of variable techniques.. We can apply this technique to the uniform distribution PDF in polar coordinates to obtain the corresponding distribution for $\rho(s)$ for the model. We recall that for a uniform point process on the disk radius D the PDF is:

$$\rho(r, \theta) = \rho(r) = \frac{2r}{D^2}. \tag{17}$$

However, the precise form of the relationship for $\rho(s)$ is not strictly necessary for the argument. Following (Barabási and Bianconi 2001) we can investigate the critical behavior of our model. Using the mean field approximation, we can write for the time evolution of the degree of any node in our set **A**, the following differential equation:

$$\frac{\partial k_i}{\partial t} = m\Pi_i = m\frac{e^{-\beta s_i}(k_i + 1)}{Z(t)}. \tag{18}$$

Allowing $r_c \to \infty$ we recover Eq. 3, for which we previously solved to verify the scale free behavior of the model. Because of the subtly different form of the solution from the standard Barabási-Albert model, and the difference in initial condition that at t_0, the time of introduction of the node, $m = 0$, we cannot proceed in the identical manner to (Barabási and Bianconi 2001). Instead, we investigate solutions to Eq. 18 of the form, for reasons that will become apparent later on:

$$k_i(t, t_0, s_i) = m\left(\frac{t}{t_0}\right)^{h(s_i)} - 1, \tag{19}$$

where $h(s_i)$ is a general function of the initial position of the node.

We start by forming an expression for the partition function at time t, $Z(t)$. We can write, perfectly generically:

$$\langle Z(t) \rangle = \int_s \rho(s_i) ds_i \int_0^t e^{-\beta s_i} k(t, t_0, s_i) dt_0, \tag{20}$$

noting that the lower limit in the integral of introduction time t_0 now ranges from 0 due to the different initial conditions. We can substitute Eq. 19 into this integral to obtain:

$$\langle Z(t) \rangle = \int_s e^{-\beta s} \rho(s) ds \int_o^t \left(m t^{h(s)} t_0^{-h(s)} - 1 \right) dt_0, \tag{21}$$

$$= \int_s e^{-\beta s} \rho(s) ds \left[\frac{m t^{h(s)} t_0^{(1-h(s))}}{1 - h(s)} - t_0 \right]_0^t, \tag{22}$$

$$= m t \int_s \frac{h(s) e^{-\beta s} \rho(s)}{1 - h(s)} ds, \tag{23}$$

after discarding terms of $O\left(t^{-h(s)}\right)$, which as $h(s) < 1$ will be negligible as $t \to \infty$.

We now introduce the chemical potential μ of the model, by exploiting the infinite time average of this partition function as follows:

$$e^{-\beta \mu} = \lim_{t \to \infty} \frac{\langle Z(t) \rangle}{mt} \tag{24}$$

If this is introduced into the original differential equation, it is possible to place a constraint upon the form of $h(s_i)$, which is easy to verify must be $h(s_i) = e^{-\beta(s-\mu)}$. Further, and by manipulation of the form for $\langle Z(t) \rangle$, we can deduce that this is true providing that the following equality holds, essentially an integral equation for the chemical potential μ:

$$\int_s \frac{e^{-\beta(s-\mu)}}{e^{\beta(s-\mu)} - 1} \rho(s) ds = 1. \tag{25}$$

We can rearrange Eq. 24 to yield:

$$\mu = -\frac{1}{\beta} \log \left(\lim_{t \to \infty} \frac{\langle Z(t) \rangle}{mt} \right). \tag{26}$$

It is clear from the denominator of Eq. 23, that this possesses discontinuities, as $h(s) = 1$ at certain values of s, μ. Discontinuities in the chemical potential are a precondition for phase transitions.

This equation is typical of the types of statistical constraints that is seen in the phenomenon of Bose-Einstein condensation (see for example (Schrödinger 1989)), and is highly sensitive to the value of our parameter β. Typically in the physical phase transition analysis, this is inverse temperature, and as $T \to 0$, this parameter diverges. At this stage it is possible to see that Eq. 25 is not soluble in the limit $\beta \to \infty$, and we would expect discontinuities, similar to Bose-Einstein condensation. These are characterized by discontinuities in the value of the chemical potential, and its first derivative which is analogous to the free energy of this system (for a discussion see Park and Newman (2004)). In our model β is dependent upon the value of r_c the interaction distance scale factor (at fixed δ_s). The larger the value of r_c relative to δ_s, the less dependent upon the distance function our attraction probability becomes, as in the limit $\beta \to \infty$, or distance function $g(s_i, r_c) \to 1$ and we recover the spatially independent model, which we have already demonstrated has a scale free degree distribution. We can conclude that as β varies, we

are likely to get distinct phases in the connectivity of our model, and in the next section we will simulate our model to investigate further the critical behavior.

Constraints in the dynamic model

In previous work (Tee et al. 2017) we explored scale free networks with a top constraint. In our spatial model of the biology of cell symbiosis we cannot assume that a given cell has an endless supply of signaling agent, and in some sense this must place a constraint on the number of cells it can be in causal contact with. This translates to a maximum degree c, and amounts to a modification of the degree function $g(k_i)$. For this model our distance moderated preferential attachment equation can be written as (with the validity restriction that strictly $k_i < c$):

$$\Pi_i = \frac{e^{-\beta s_i}(k_i + 1)(c - k_i)}{Z(t)}, \text{ where} \tag{27}$$

$$Z(t) = \sum_j e^{-\beta s_j}(k_j + 1)(c - k_i). \tag{28}$$

Solving this model analytically is beyond the scope of this work, but in the limit $c \to \infty$ the reader should convince themselves that the $c - k_i$ factors cancel out as a constant and we retrieve our original model. However, it is interesting to perform the criticality analysis as in the previous section, using the form for the degree evolution for small values of k_i relative to c as the standard scale free evolution equation:

$$k_i(t, t_0, s_i) = m\left(\frac{t}{t_0}\right)^{h(s_i)}. \tag{29}$$

Our expression for the time average of the partition function now becomes, approximating $k_i + 1$ to k_i:

$$\langle Z(t) \rangle = \int_s e^{-\beta s} \rho(s) ds \int_o^t m t^{h(s)} t_0^{-h(s)} \left(c - m t^{h(s)} t_o^{h(s)}\right) dt_0, \tag{30}$$

$$= \int_s e^{-\beta s} \rho(s) ds\, mt \left[\frac{(c-1)t^{h(s)} t_0^{1-h(s)}}{1 - h(s)} + \frac{m t^{2h(s)} t_0^{1-2h(s)}}{1 - 2h(s)}\right]_0^t, \tag{31}$$

$$= mtc \int_s \frac{h(s) e^{-\beta s} \rho(s)}{1 - h(s)} ds, \tag{32}$$

where we have ignored terms small compared to c, and assumed $c \gg 1$, and $c \gg m$. It is interesting that under these conditions, we are able to follow exactly the logic of the previous section and we will maintain the same critical behavior until starvation of signaling factor (a low value of c relative to m for example) is in place, with the difference that we have to replace the definition of chemical potential with:

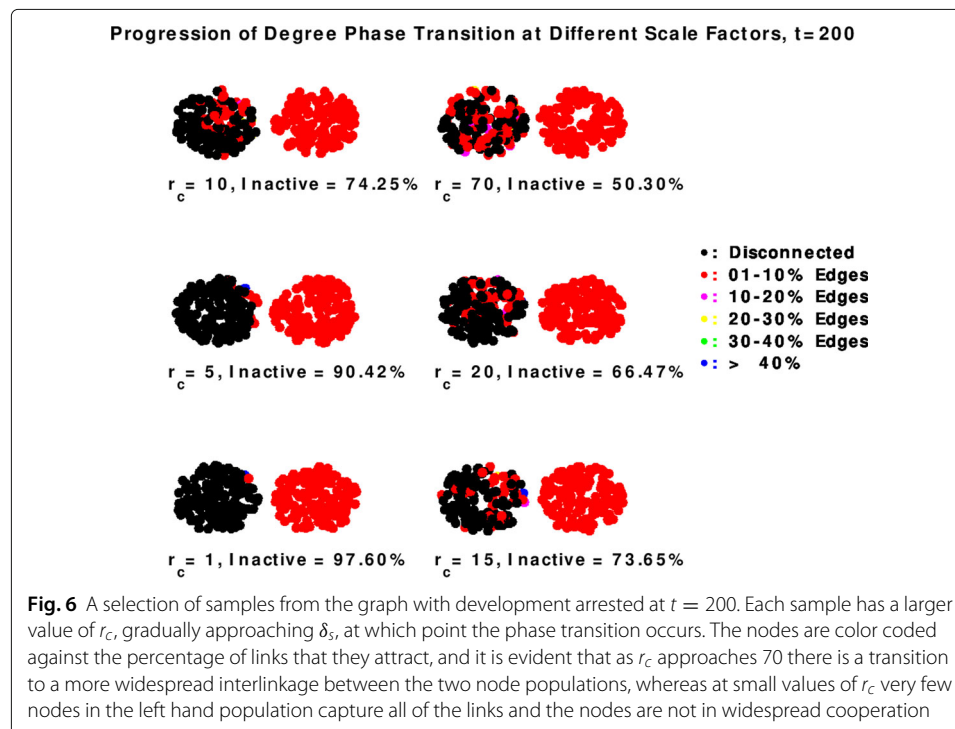$$e^{-\beta \mu} = \lim_{t \to \infty} \frac{\langle Z(t) \rangle}{mtc}, \tag{33}$$

and self consistency requires $h(s) = \frac{1}{c} e^{-\beta(s-\mu)}$. The constant factor does not change the critical behavior of $Z(t)$, and we should expect discontinuities in chemical potential and phase transitions.

When we cannot assert $c \gg 1$, and $c \gg m$, the denominator of Eq. 25 is replaced with a quadratic term in $h(s)$, $1 - 3h + 2h^2$. This time however the quadratic term in h, introduces

an additional factor of c into the denominator, which in turn removes the singularity for values of $c \geq 2$, at $s = \mu$. This is verified by noting that the roots of the quadratic in $h(s)$ are $(\frac{1}{2}, 1)$, and $h(s) < \frac{1}{2}$ for $c \geq 2$. The critical behavior is evidently more complex, but intuitively adding a constraint will prevent any one node consuming all of the links for times $t > c/m$, by definition. Deeper analytic treatment of this scenario is the subject of ongoing work, but the constraint c acts like a tuning parameter between Bose-Einstein statistics with critical behavior, and in the limit $c = 1$ a network that should possess Fermi-Dirac statistics.

Simulation and analysis

The model as described in "The Spatio-Temporal model" section, has been simulated to generate a series of random samples of graphs produced. It is apparent from the data the expected phase transitions in the connectivity of the graphs produced. Firstly, in Fig. 6, we graphically depict the spatial graph across a range of scale factors r_c. In the diagram the nodes are depicted in their actual position, but to aid the understanding of the connectivity edges are omitted. Instead we color code the nodes by the degree, as a fraction of the total number of edges in the graph, arrested at time $t = 200$. The parameter in all of our simulations $m = 2$ and we start with $m_0 = 20$ nodes. Different values of r_c have been modeled, and the results presented here are representative of the general results. The simulation itself was performed using code written in Matlab, and can be obtained from the author on request. The nodes in set **A** are depicted on the left, and the nodes in **B** on the right. For reference the distance between the loci of the two distributions is 70 units, and the nodes are constrained to a disk of radius 30 units. For very small r_c, the majority of the nodes in **A** attract any connections, and example of the 'winner takes all' phenomena well studied in Barabási and Bianconi (2001). Indeed all 800 edges generated

Fig. 6 A selection of samples from the graph with development arrested at $t = 200$. Each sample has a larger value of r_c, gradually approaching δ_s, at which point the phase transition occurs. The nodes are color coded against the percentage of links that they attract, and it is evident that as r_c approaches 70 there is a transition to a more widespread interlinkage between the two node populations, whereas at small values of r_c very few nodes in the left hand population capture all of the links and the nodes are not in widespread cooperation

for a value of $r_c = 1$ are consumed by just 2.4% or 19 nodes. As r_c increases in value a sudden increase in the number of nodes in **A** occurs, and the distribution of connectivity begins to resemble the familiar form of a scale free distribution, with some nodes acquiring many connections. It is an interesting feature of this evolving spatial network that some nodes in the set **A** remain disconnected.

The simulation was also used to generate data for the collapse of the super hub as r_c increases, and using Eq. 24 values for the chemical potential were also obtained. We present those results in Fig. 7a, b, c, d. As expected in the analysis, the chemical potential shows a sharp discontinuity at around $r_c = 70$. This is approximately the distance between the loci. In Fig. 7a, c, we also plot the maximum degree in the graph as a fraction of the total number of edges incident on the set **A**. Again a very clear drop off is witnessed around value of r_c around 20 where the distance function gives more nodes in the set comparable values and the evolution can settle down into a regular preferential attachment regime.

Finally, we wish to demonstrate that the degree distributions for the base model satisfy, at least approximately, a power law. In Fig. 8a and b we reproduce the cumulative degree distributions for both the base model and also the model with a top constraint described in "Constraints in the dynamic model" section. We chose Cumulative Distribution Functions (CDFs) to avoid the distortion in degree distributions for high degree, low frequency nodes, and the data is presented with the network growth arrested at $t = 2000$, implying a 4000 node network. In Fig. 8a, we see that that above approximately $k = 3$, the distributions are linear on a logarithmic scale, even as the network approaches criticality. We have not analyzed the data rigorously to determine the power law exponent, but the fundamental point of a power law operating over a range of degree values can be seen by inspection. In Fig. 8b we have applied a top constraint of $c = 50$, and the distribution is evidently not a power law, and terminates abruptly at $k = c$. As discussed in the previous section, the presence of a top constraint has the effect of preventing critical behavior, and this at least is supported by the degree distribution. Along with the analytical work and experimental verification of the model, the constrained growth regime is the subject of future research.

Conclusion and future work

We have been able to demonstrate in this work that a carefully constructed dynamic spatial model of network evolution exhibits a critical transition in connectivity of the graph from a state where two cell populations show minimal interactions to one in which cell communication is widespread. Further, the order parameter in this phase transition is the scale that governs the effective distance at which interactions between the nodes operate.

It is of interest to consider why linear clonal evolution applies to the stage of cancer initiation in the model used here, and in early progression to carcinoma, but appears to require a transition to a more interactive state during development of the metastasis. It is possible that new interactions at the stage of metastasis allow the tumor to re-create a microenvironmental niche similar to that which existed at the primary tumor site. At the site of origin, the clonal evolution of early tumor cells is still taking place in a microenvironment that has all of the cell types. These include stem cells and fully differentiated cells, whose normal functions are to promote cell migration, growth and responses to the stress of tissue regeneration. At distant sites, many of these alternative niche signals are
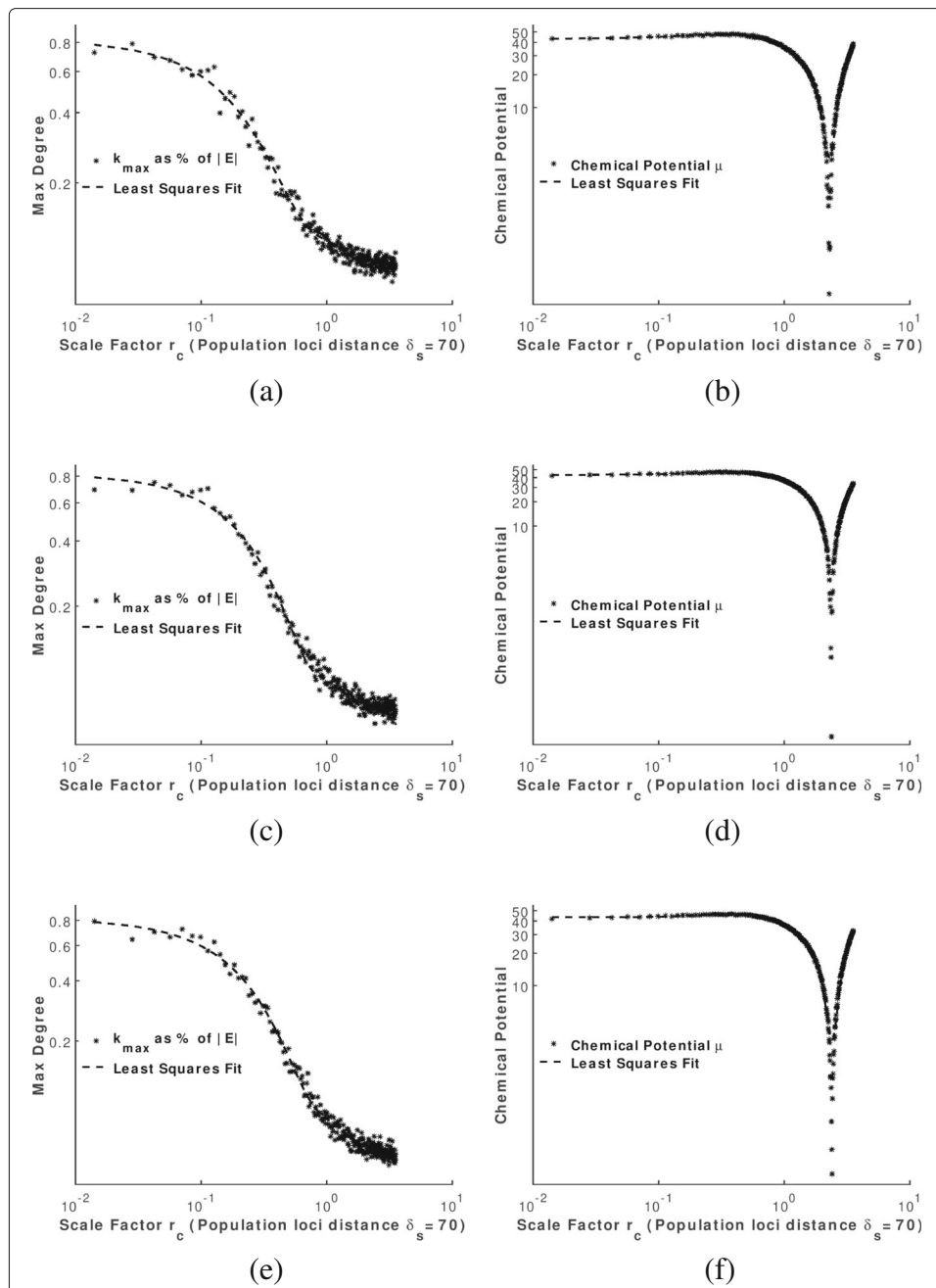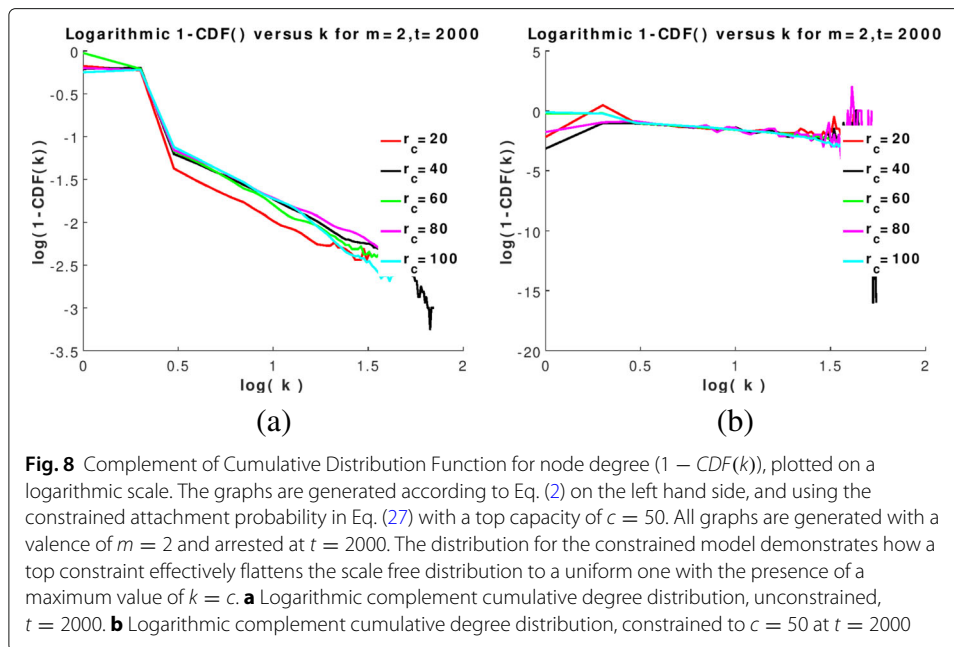
Fig. 7 A series of plots for increasing size networks, demonstrating the Critical Behavior of Networks as r_c varies. On the left hand side we present the degree of the highest connectivity node as a fraction of all links in the network on a logarithmic scale. As the distance factor increases after a period of stability this collapses from 0.8 indicating that on average nodes are interconnected between node populations. On the right hand side we plot absolute chemical potential μ, again on a logarithmic scale. As r_c approaches the separation between the population loci $\delta_s = 70$ there is a clear discontinuity in the chemical potential, an essential characteristic of a phase transition. **a** k_{max} as a fraction of all edges, arrested at $t = 200$. **b** Chemical Potential μ, arrested at $t = 200$. **c** k_{max} as a fraction of all edges, arrested at $t = 500$. **d** Chemical Potential μ, arrested at $t = 500$. **e** k_{max} as a fraction of all edges, arrested at $t = 1000$. **f** Chemical Potential μ, arrested at $t = 1000$

Fig. 8 Complement of Cumulative Distribution Function for node degree ($1 - CDF(k)$), plotted on a logarithmic scale. The graphs are generated according to Eq. (2) on the left hand side, and using the constrained attachment probability in Eq. (27) with a top capacity of $c = 50$. All graphs are generated with a valence of $m = 2$ and arrested at $t = 2000$. The distribution for the constrained model demonstrates how a top constraint effectively flattens the scale free distribution to a uniform one with the presence of a maximum value of $k = c$. **a** Logarithmic complement cumulative degree distribution, unconstrained, $t = 2000$. **b** Logarithmic complement cumulative degree distribution, constrained to $c = 50$ at $t = 2000$

not available, resulting in a bottleneck that the successful metastatic cell has to overcome in order to survive. This bottleneck may necessitate the generation of clones that can produce the correct combinations of growth and survival factors that mimic, as closely as possible, niche conditions available for tumor growth at the primary tumor site. In this sense, the phase transition observed in this mathematical model may represent an emergent property of the progressing tumor – the acquisition of collective cellular behavior to overcome the new challenges of growth in a foreign environment (Lyons et al. 2008).

In the model developed here, there should emerge ratios of cooperating cell populations at which the growth rate increases non-uniformly. Interestingly, experimental studies have suggested that the ratio of admixed populations is a critical determinant of malignant tumor growth. Elegant studies by Marusyk et al. (2014) showed that the ratio of tumor cells expressing the growth factor IL11 and the cell matrix enzyme Lysyl oxidase 4 (LOXL4) was a major determinant of tumor volume in xenograft models of human cancer cell lines. These novel studies were however limited to growth of human tumor cell lines in immunocompromised mice. In view of the importance of the immune system in controlling systemic spread of metastatic cells, it would be important to carry out similar studies to estimate the functional consequences of cell population mixing in immunocompetent models where the phenotypes of local tumor growth and metastasis can be separately evaluated. These studies could reveal the mechanistic basis and specific factors involved in cell-cell cooperation leading to tumor dissemination. As indicated in the earlier text, the authors intend to test the model by identifying whether in laboratory conditions this behavior can be replicated. The outline of the experiment is as follows. Two cell populations can be separately cultured and mixed according to fixed ratios from 0:1 to 1:1. Fixing the ration of cell populations, in a fixed area, will have the effect of fixing the average distance between two randomly selected cells from the two different populations. From there we can measure the growth rates of the two populations. If our symbiotic model is promoting tumorigenesis, there should emerge ratios at which the growth rate

increases non-uniformly. These ratios can then be tested across other different mixes of cell populations. Confidence in the model would grow if the ratio was repeatedly identical for a given population ratio.

Author details
[1]Moogsoft Inc,1265 Battery Street, 94111 San Francisco, CA, USA. [2]Beyond Center, Arizona State University, 85287 Tempe, AZ, USA. [3]University of Sussex BN1 9RH Brighton, UK. [4]UCSF Helen Diller Family Comprehensive Cancer Center, 1450 3rd Street, 94143 San Francisco, California, USA.

References

Barabási G, Bianconi AL (2000) Competition and multiscaling in evolving networks. Europhys Lett 54(May):13. https://doi.org/10.1209/epl/i2001-00260-6. http://arxiv.org/abs/cond-mat/0011029%5Cnhttp://iopscience.iop.org/article/10.1209/epl/i2001-00260-6/meta;jsessionid=0D1D3E5DC321F5EFACF4772446228069.c2.iopscience.cld.iop.org

Barabási AL, Bianconi G (2001) Bose-Einstein Condensation in Complex Networks. Phys Rev Lett 86(24):5632–5635. https://doi.org/10.1103/PhysRevLett.86.5632. http://link.aps.org/doi/10.1103/PhysRevLett.86.5632%5Cnhttp://prl.aps.org/abstract/PRL/v86/i24/p5632_1%5Cnhttp://prl.aps.org/pdf/PRL/v86/i24/p5632_1

Barthélemy M (2011) Spatial networks. Phys Rep 499(1-3):1–101. https://doi.org/10.1016/j.physrep.2010.11.002

Barthélemy M (2003) Crossover from scale-free to spatial networks. Europhys Lett 63(6):915–921. https://doi.org/10.1209/epl/i2003-00600-6

Bianconi G (2017) Rare events and discontinuous percolation transitions 022314(February):1–11. https://doi.org/10.1103/PhysRevE.97.022314. http://arxiv.org/abs/1707.00348

Braha D, Stacey B, Bar-Yam Y (2011) Corporate competition: A self-organized network. Soc Networks 33(3):219–230. https://doi.org/10.1016/j.socnet.2011.05.004

Broido AD, Clauset A (2018) Scale-free networks are rare June:26–28. http://arxiv.org/abs/1801.03400

Calbo J, van Montfort E, Proost N, van Drunen E, Beverloo HB, Meuwissen R, Berns A (2011) A Functional Role for Tumor Cell Heterogeneity in a Mouse Model of Small Cell Lung Cancer. Cancer Cell 19(2):244–256. https://doi.org/10.1016/j.ccr.2010.12.021

Clauset A, Rohilla Shalizi C, Newman MEJ (2009) Power-Law Distributions in Empirical Data. SIAM Rev 51(4):661–703. https://doi.org/10.1214/13-AOAS710

Cleary AS, Leonard TL, Gestl SA, Gunther EJ (2014) Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. Nature 508(1):113–117. https://doi.org/10.1038/nature13187

Dall J, Christensen M (2002) Random Geometric Graphs March:1–9. https://doi.org/10.1103/PhysRevE.66.016121. http://arxiv.org/abs/cond-mat/0203026%0Ahttp://dx.doi.org/10.1103/PhysRevE.66.016121

Dettmann CP, Georgiou O (2016) Random geometric graphs with general connection functions. Phys Rev E 93(3):1–16. https://doi.org/10.1103/PhysRevE.93.032313

Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) Structure of growing networks with preferential linking. Phys Rev Lett 85(21):4633–4636. https://doi.org/10.1103/PhysRevLett.85.4633

Dorogovtsev SN, Goltsev AV, Mendes JFF (2007) Critical phenomena in complex networks 80(December). https://doi.org/10.1103/RevModPhys.80.1275. http://arxiv.org/abs/0705.0010%0Ahttp://dx.doi.org/10.1103/RevModPhys.80.1275

Georgiou O, Dettmann C, Coon J (2013) K-Connectivity for Confined Random Networks. arXiv preprint arXiv:1304.1329 103(July):1–6. https://doi.org/10.1209/0295-5075/103/28006. http://arxiv.org/abs/1304.1329

Greaves M, Maley CC (2012) Clonal evolution in cancer. Nature 481(7381):306. https://doi.org/10.1038/nature10762

Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM (2017) The dynamics of molecular evolution over 60,000 generations. Nature 551(7678):45–50. https://doi.org/10.1038/nature24287

Heppner GH (1984). Cancer Res 44(6):2259–2265

Khajanchi S, Perc M, Ghosh D (2018) The influence of time delay in a chaotic cancer model. Chaos 28(10). https://doi.org/10.1063/1.5052496

Lim JS, Ibaseta A, Fischer MM, Cancilla B, O'Young G, Cristea S, Luca VC, Yang D, Jahchan NS, Hamard C, Antoine M, Wislez M, Kong C, Cain J, Liu YW, Kapoun AM, Garcia KC, Hoey T, Murriel CL, Sage J (2017) Intratumoural heterogeneity

generated by Notch signalling promotes small-cell lung cancer. Nature 545(7654):360–364. https://doi.org/10.1038/nature22323

Lyons JG, Lobo E, Martorana AM, Myerscough MR (2008) Clonal diversity in carcinomas: Its implications for tumour progression and the contribution made to it by epithelial-mesenchymal transitions. Clin Exp Metastasis 25(6):665–677. https://doi.org/10.1007/s10585-007-9134-2

Manna SS, Sen P (2002) Modulated scale-free network in Euclidean space. Phys Rev E - Stat Phys Plasmas Fluids Relat Top 66(6):4. https://doi.org/10.1103/PhysRevE.66.066114

Marusyk A, Almendro V, Michor F, Altrock PM, Tabassum DP, Polyak K (2014) Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature 514(7520):54–58. https://doi.org/10.1038/nature13556

Moreno E, Champalimaud LF, Basler K, Morata G (2014) Cells compete for Decapentaplegic survival factor to prevent apoptosis in Drosophila wing development Growth regulation by morphogen gradients View project Drosophila TNF View project:755–759. https://doi.org/10.1038/416755a. www.nature.com

Ostilli M, Bianconi G (2015) Statistical mechanics of random geometric graphs: Geometry-induced first-order phase transition. Phys Rev E Stat Nonlinear Soft Matter Phys 91(4):1–14. https://doi.org/10.1103/PhysRevE.91.042136

Park J, Newman MEJ (2004) Statistical mechanics of networks. Phys Rev E Stat Nonlinear Soft Matter Phys 70(6 2):1–13. https://doi.org/10.1103/PhysRevE.70.066117

Perc M, Szolnoki A (2010) Coevolutionary games-A mini review. BioSystems 99(2):109–125. https://doi.org/10.1016/j.biosystems.2009.10.003

Reeves MQ, Kandyba E, Harris S, Del Rosario R, Balmain A (2018) Multicolour lineage tracing reveals clonal dynamics of squamous carcinoma evolution from initiation to metastasis. Nat Cell Biol 20(6):699–709. https://doi.org/10.1038/s41556-018-0109-0

Rozenfeld AF, Cohen R, Ben-Avraham D, Havlin S (2002) Scale-Free Networks on Lattices. Phys Rev Lett 89(21):2–5. https://doi.org/10.1103/PhysRevLett.89.218701

Schrödinger E (1989) Statistical Thermodynamics. Dover Books on Physics Series. Dover Publications

Sen P, Banerjee K, Biswas T (2002) Phase transitions in a network with a range-dependent connection probability. Phys Rev E Stat Phys Plasmas Fluids Relat Top 66(3):1–4. https://doi.org/10.1103/PhysRevE.66.037102

Tabata T (2004) Morphogens, their identification and regulation. Development 131(4):703–712. https://doi.org/10.1242/dev.01043. http://dev.biologists.org/cgi/doi/10.1242/dev.01043

Tammela T, Sanchez-Rivera FJ, Cetinbas NM, Wu K, Joshi NS, Helenius K, Park Y, Azimi R, Kerper NR, Wesselhoeft RA, Gu X, Schmidt L, Cornwall-Brady M, Yilmaz ÖH, Xue W, Katajisto P, Bhutkar A, Jacks T (2017) A Wnt-producing niche drives proliferative potential and progression in lung adenocarcinoma. Nature 545(7654):355–359. https://doi.org/10.1038/nature22334

Tee P, Balmain A (2019) Phase Transitions in Spatial Networks as a Model of Cellular Symbiosis 2:1–12. https://doi.org/10.1007/978-3-030-05414-4_47

Tee P, Wakeman I, Parisis G, Dawes J, Kiss I (2017) Constraints and entropy in a model of network evolution. Eur Phys J B 90(11). https://doi.org/10.1140/epjb/e2017-80185-5

Venkatesan S, Swanton C (2016) Tumor Evolutionary Principles: How Intratumor Heterogeneity Influences Cancer Treatment and Outcome. American Society of Clinical Oncology Educational Book 36:e141–e149. https://doi.org/10.14694/edbk_158930

Weinberg H (2011) Cell - Hallmarks of Cancer: The Next Generation. Cell Vol 144, Issue 5, 646-674, 4 March 2011 144(5):646–674. http://www.cell.com/fulltext/S0092-8674(11)00127-9

Wright S (1982) The Shifting Balance Theory and Macroevolution. Annu Rev Genet 16(1):1–20. https://doi.org/10.1146/annurev.ge.16.120182.000245. http://www.annualreviews.org/doi/10.1146/annurev.ge.16.120182.000245

Publisher's Note