Cell invasion with proliferation mechanisms motivated by time-lapse data

Matthew J. Simpson [a,*], Kerry A. Landman [b], Barry D. Hughes [b]

[a] Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Queensland 4001, Australia
[b] Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia

ARTICLE INFO

ABSTRACT

Cell invasion involves a population of cells which are motile and proliferative. Traditional discrete models of proliferation involve agents depositing daughter agents on nearest-neighbor lattice sites. Motivated by time-lapse images of cell invasion, we propose and analyze two new discrete proliferation models in the context of an exclusion process with an undirected motility mechanism. These discrete models are related to a family of reaction–diffusion equations and can be used to make predictions over a range of scales appropriate for interpreting experimental data. The new proliferation mechanisms are biologically relevant and mathematically convenient as the continuum–discrete relationship is more robust for the new proliferation mechanisms relative to traditional approaches.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Cell invasion involves combined cell proliferation and cell motility, and is essential to development [1–4], wound healing [5,6] and malignant progression [7,8]. Several types of experimental observation can be made to describe cell invasion [9]. Global properties, such as the speed of invasion fronts [3,4], as well as individual properties like the details of individual motility and proliferation events [1,2,4,10], are both described and quantified experimentally.

Continuum models of cell invasion typically combine a random motility mechanism and carrying capacity limited proliferation represented by logistic growth, giving reaction–diffusion equations, many of which are related to Fisher's model [11]. These models have been successfully used to obtain global properties in a variety of applications such as wound healing [5,6], malignant invasion [8], developmental morphogenesis [3] and gene propagation [12].

Modern microscopy techniques provide high quality imaging data giving us additional information about invasive systems. These observations unearth new opportunities to develop models based on observations at the level of an individual cell rather than a collective population-level description. Averaging these discrete models can lead to a continuum description of the system giving us a tool that is capable of representing both global and individual-level properties [13,14], compatible with experimental data [9].

Traditional lattice-based discrete models of cell proliferation involve a proliferative agent depositing a daughter agent on a nearest-neighbor site [15–17]. It is well-known that this discrete model is related to logistic proliferation in the appropriate continuum limit [5]. Time-lapse data shown here motivate us to consider two models—one with a new proliferation rule and the other with a generalization of the traditional discrete proliferation rule. Averaging each of the new discrete invasion models gives a partial differential equation (PDE), that is a generalization of Fisher's equation. Additional tools are developed allowing us to predict the average trajectory of a tagged cell within the invasive population [2,14]. We show that the

new proliferation mechanisms are both biologically relevant and mathematically convenient as the continuum–discrete relationship is more robust for the new proliferation mechanisms relative to the traditional proliferation mechanism.

2. Individual-level cell invasion model

A lattice-based simple exclusion process [18], with at most one agent per site, is used to model cell invasion. We use a two-dimensional square lattice with spacing Δ. Each site is indexed (i, j) where $i, j \in \mathbb{Z}$, and each site has position $(x, y) = (i\Delta, j\Delta)$. The lattice spacing may be thought of as being the size of a cell diameter. Our model and analysis can be easily implemented for a range of other lattices.

In any one realization of the model the occupancy of site (i, j) is $C_{i,j}$, with $C_{i,j} = 1$ for an occupied site, and $C_{i,j} = 0$ for a vacant site. If there are N agents on the lattice, during the next time step of duration τ, N agents are selected independently at random, one at a time. When chosen, an agent attempts to move with probability $P_m \in [0, 1]$ [19]. We interpret P_m as the probability that an agent will attempt to move a distance Δ in the time interval τ. Once the N motility events are attempted, another N agents are selected independently at random, one at a time, and these agents attempt to proliferate with probability $P_p \in [0, 1]$. In general N increases during each time step for $P_p > 0$ and we interpret P_p as the probability that an agent will attempt to proliferate in the time interval τ. This approach is appropriate for small values of P_p where the increase in N per time step is small. We always work with dimensionless simulations by setting $\Delta = \tau = 1$. The results can be rescaled using appropriate length and time scales for any particular application.

Time-lapse images showing cell invasion associated with the development of the enteric nervous system (ENS) motivate our discrete model. ENS development involves neural crest cells (NCCs) invading the developing gut tissues. NCCs are motile and proliferate to a maximum density [3]. The population of NCCs invades the gut tissue as a constant speed invasion wave and differentiates into neurons and glia to form the ENS [1,2,4].

Time-lapse data examining the movement of a few isolated NCCs in the absence of proliferation showed that the cells followed unpredictable random trajectories [4]. To mimic this, the motility mechanism in our discrete model is an unbiased simple exclusion process where a motile agent at (x, y) attempts to move to either $(x \pm \Delta, y)$ or $(x, y \pm \Delta)$, each with equal probability $1/4$ [19]. Since biological cells cannot occupy the same position in space [13,14], our discrete model is an exclusion process and motility events that would place an agent on an occupied site are aborted. An illustration of the potential outcomes of a motility event is given in Fig. 1(b).

The focus of this work is to investigate several new biologically plausible proliferation mechanisms. The key difference between the mechanisms is how the daughter agents are arranged spatially relative to the location of the original agent. Previous studies have considered the model where a proliferative agent at (x, y) deposits a daughter agent in one of $(x \pm \Delta, y)$ or $(x, y \pm \Delta)$ with equal probability $1/4$ [5,15–17]. These models have a separation distance $n = 1$ between the original cell and the daughter cell.

Time-lapse images of ENS development [1], illustrated in Fig. 2, show the physical details of a NCC proliferation event. The NCC division is composed of three distinct phases: (i) the cell ceases random motion and rounds up (Fig. 2(a)); (ii) the cell divides into two daughter cells (Fig. 2(b)); (iii) the daughter cells separate ballistically in opposing directions and are placed approximately six cell diameters equidistant apart from the location of the original cell (Fig. 2(c)). It is only on completion of the third phase that the daughter cells commence random motility. The proliferation event shown in Fig. 2 is typical (H. Young, personal communication, 2009) and has been observed in other time-lapse images of NCC invasion [4]. To represent these details in our discrete model we will concatenate the three phases of proliferation into a single event.

The time-lapse images in Fig. 2 demonstrate that the proliferation event is not to nearest-neighbor sites for this biological system and the separation distance is much larger than $n = 1$. Therefore, an alternative to the traditional discrete proliferation mechanism is needed to model NCC invasion. We propose and analyze two proliferation models that are shown schematically in Fig. 1(c)–(d).

Model 1 is a generalization of the traditional proliferation mechanism. A proliferative agent at (x, y) deposits a daughter agent at one of $(x \pm n\Delta, y)$ or $(x, y \pm n\Delta)$ with equal probability $1/4$. Here n is a positive integer representing the number of cell diameters that the daughter agent is placed relative to the original agent. We call n the separation distance and note that the traditional proliferation model corresponds to Model 1 with $n = 1$.

Model 2 is a new model where a proliferative agent at (x, y) produces two daughter agents which either reside in $(x + n\Delta/2, y)$ and $(x - n\Delta/2, y)$ or $(x, y + n\Delta/2)$ and $(x, y - n\Delta/2)$, each with equal probability $1/2$. Here n is a positive even integer representing the separation distance between the daughter agent and the original agent. Model 2 describes the proliferation event in Fig. 2 with $n = 6$.

Given we are dealing with an exclusion process, proliferation events that would place a daughter agent on an occupied site are aborted. Now that we have introduced the biologically motivated mechanism of placing a daughter agent at a site which is not necessarily a nearest neighbor of site (x, y), we must also check that each site between the original site and the target site is unoccupied, otherwise the proliferation event will be aborted. For example, when implementing Model 1 with $n = 2$, a proliferative agent at (x, y) attempting to place a daughter agent at $(x + 2\Delta, y)$ would only succeed in doing so if sites $(x + \Delta, y)$ and $(x + 2\Delta, y)$ were both vacant.

One of the differences between the two proliferation models proposed here is that the originally occupied site (x, y) remains occupied after the proliferation event in Model 1, whereas site (x, y) becomes vacant after the proliferation event
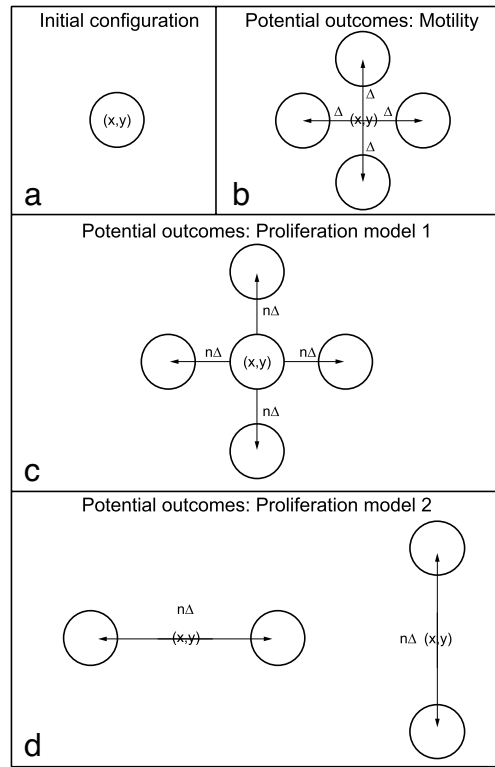
Fig. 1. Discrete motility and proliferation outcomes. (a) Consider an initially occupied site at location (x, y). (b) Motility events occur with probability P_m and a motile agent at (x, y) steps to either $(x \pm \Delta, y)$ or $(x, y \pm \Delta)$ with equal probability $1/4$. Proliferation events occur with probability P_p. (c) In Model 1 a proliferative agent at (x, y) deposits a daughter agent at either $(x \pm n\Delta, y)$ or $(x, y \pm n\Delta)$ with equal probability $1/4$ provided that all sites between (x, y) and the target site are vacant. For Model 1, n is a positive integer. (d) In Model 2 a proliferative agent at (x, y) divides into two daughter agents that are placed at either $(x + n\Delta/2, y)$ and $(x - n\Delta/2, y)$, or $(x, y + n\Delta/2)$ and $(x, y - n\Delta/2)$, with equal probability $1/2$, provided that all sites between (x, y) and the target sites are vacant. For Model 2, n is a positive and even integer.
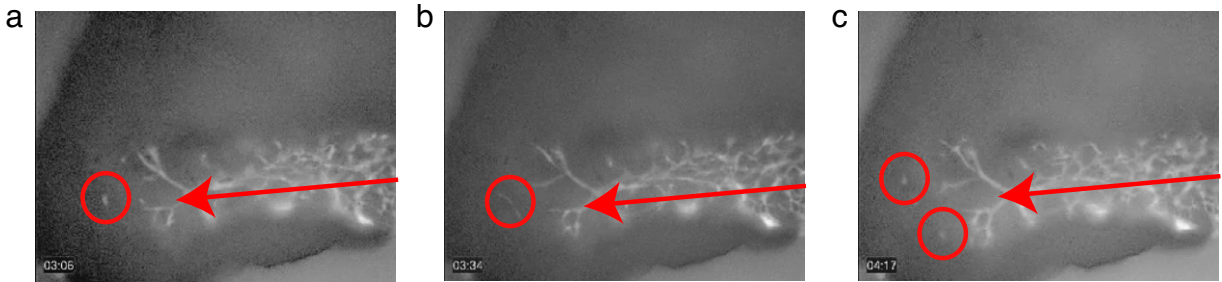

Fig. 2. Time-lapse images of a cell invasion assay showing NCCs that are both motile and proliferative [1]. The population of NCCs invades the developing intestine in a mouse model. The invasive cells are labeled white. The spatial organization of a single proliferation event is highlighted. In each snapshot the direction of invasion is shown with the red arrow. (a) A cell (red circle) prior to division. (b) The same cell (red circle) as it divides into two daughter cells. (c) The two resulting daughter cells are located approximately six cell diameters apart.
Source: Results reproduced with permission from Elsevier.

in Model 2. Neither of these proliferation models allow for agent removal from the system. This is appropriate for ENS development since NCC death is not observed experimentally [4].

To demonstrate visually the influence of implementing these different proliferation models a suite of simulation results is shown in Fig. 3. In each problem we consider a lattice of size 400×20. All sites where $180 \le x \le 220$ are initially occupied. Periodic boundary conditions were imposed on the horizontal boundaries and reflecting boundary conditions were imposed on the vertical boundaries. Six simulations over 3000 time steps with $P_m = 1$ and $P_p = 0.001$ were performed. This choice of parameter values is discussed in Section 3. Results for (i) Model 1 with $n = 1, 2, 3$, and (ii) Model 2 with $n = 2, 4, 6$ are given. In all cases the number of agents increased significantly and the population invaded the unoccupied regions of the lattice in both directions away from the initial location of the agents.
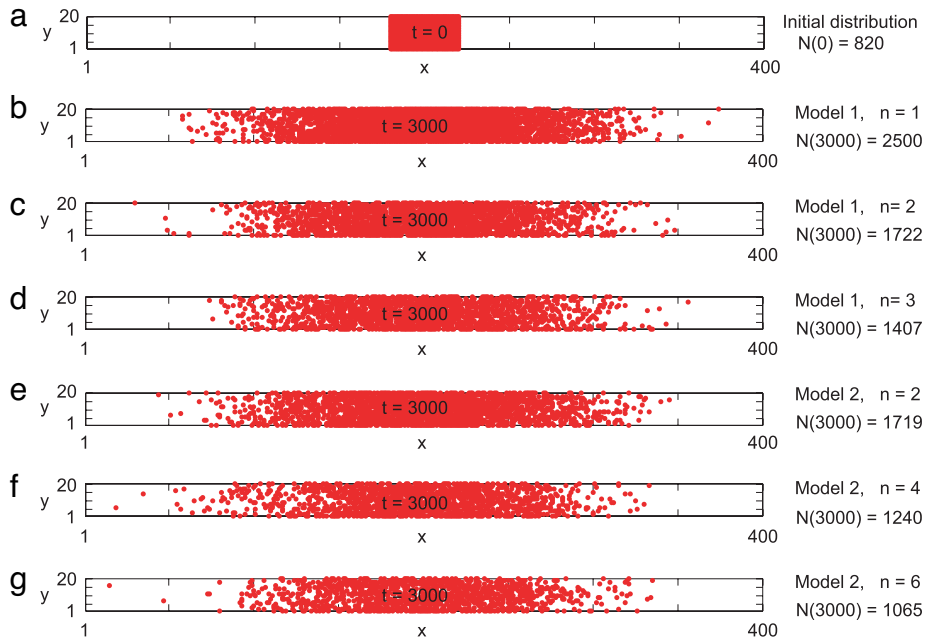
Fig. 3. (a) Simulations start with all sites occupied where $180 \le x \le 220$. (b)–(g) Simulations with $P_m = 1$ and $P_p = 0.001$ are performed for 3000 time steps for proliferation Model 1 with $n = 1, 2, 3$ and proliferation Model 2 with $n = 2, 4, 6$. All simulations start with $N(0) = 820$ agents; the number of agents after 3000 time steps $N(3000)$ is shown.

Some simple but instructive observations can be made from Fig. 3. First, regardless of the proliferation mechanism, the horizontal length over which the population spreads is comparable for each proliferation mechanism. This is surprising. Given that we have introduced proliferation mechanisms that deposit daughter agents at variable distances away from the site of proliferation, we might have expected the spatial spread of the population would increase with n. Second, differences in the density of agents are observed. The agents are densely packed in the central region of the lattice for Model 1 with $n = 1$ whereas agents are less densely packed in the central region of the lattice for Model 2 with $n = 6$. Third, the number of agents after 3000 steps decreases with n for both models. This is reasonable as successful proliferation events require an increasing number of vacant sites as n increases. During the early part of the simulation, vacant sites near the central part of the lattice are unavailable and proliferation events are more likely to be aborted for larger values of n. In Section 3 we will analyze the discrete models and gain further insight into these observations.

3. Averaging the discrete mechanism

To connect the discrete mechanism with a continuum model we average the occupancy of site (i, j) over many statistically identical realizations to obtain $\langle C_{i,j} \rangle \in [0, 1]$ [13,14]. After averaging, we form a discrete conservation statement describing $\delta \langle C_{i,j} \rangle$, which is the change in average occupancy of site (i, j) during the time interval from t to $t + \tau$. The details of the conservation equations depend on the proliferation model. For Model 1 we obtain:

$$\delta \langle C_{i,j} \rangle = \frac{P_m}{4}(1 - \langle C_{i,j} \rangle) \sum \langle C_{i,j} \rangle - \frac{P_m}{4} \langle C_{i,j} \rangle \left(4 - \sum \langle C_{i,j} \rangle\right) + \frac{P_p}{4} \langle C_{i-n,j} \rangle \prod_{s=0}^{n-1}(1 - \langle C_{i-s,j} \rangle)$$

$$+ \frac{P_p}{4} \langle C_{i+n,j} \rangle \prod_{s=0}^{n-1}(1 - \langle C_{i+s,j} \rangle) + \frac{P_p}{4} \langle C_{i,j-n} \rangle \prod_{s=0}^{n-1}(1 - \langle C_{i,j-s} \rangle) + \frac{P_p}{4} \langle C_{i,j+n} \rangle \prod_{s=0}^{n-1}(1 - \langle C_{i,j+s} \rangle), \quad (1)$$

where, for brevity we define

$$\sum \langle C_{i,j} \rangle = \langle C_{i-1,j} \rangle + \langle C_{i+1,j} \rangle + \langle C_{i,j+1} \rangle + \langle C_{i,j-1} \rangle. \quad (2)$$

The positive terms on the right of Eq. (1) represent events that place an agent at site (i, j) while the negative terms represent events that remove agents from site (i, j). For Model 1, all proliferation events increase the occupancy of site (i, j) since Model 1 does not involve removing any agents from any site. All terms in the discrete conservation statement involve factors like $\langle C_{i,j} \rangle$ and $(1 - \langle C_{i,j} \rangle)$, which are interpreted as probabilities of occupancy and vacancy respectively. Furthermore, products of these factors are interpreted as transition probabilities. Therefore we make the standard assumption that the occupancy

of lattice sites is independent. This assumption is inappropriate for any single realization of the discrete model, but proves to be an extremely good approximation when considering averaged simulation data [13,14].

The discrete conservation statement for Model 2 is

$$
\begin{aligned}
\delta \langle C_{i,j} \rangle ={}& \frac{P_m}{4}(1 - \langle C_{i,j} \rangle) \sum \langle C_{i,j} \rangle - \frac{P_m}{4} \langle C_{i,j} \rangle \left(4 - \sum \langle C_{i,j} \rangle\right) \\
&+ \frac{P_p}{2} \langle C_{i-n/2,j} \rangle \prod_{\substack{s=0 \\ s \neq n/2}}^{n} (1 - \langle C_{i-s,j} \rangle) + \frac{P_p}{2} \langle C_{i+n/2,j} \rangle \prod_{\substack{s=0 \\ s \neq n/2}}^{n} (1 - \langle C_{i+s,j} \rangle) \\
&+ \frac{P_p}{2} \langle C_{i,j-n/2} \rangle \prod_{\substack{s=0 \\ s \neq n/2}}^{n} (1 - \langle C_{i,j-s} \rangle) + \frac{P_p}{2} \langle C_{i,j+n/2} \rangle \prod_{\substack{s=0 \\ s \neq n/2}}^{n} (1 - \langle C_{i,j+s} \rangle) \\
&- \frac{P_p}{2} \langle C_{i,j} \rangle \prod_{\substack{s=-n/2 \\ s \neq 0}}^{n/2} (1 - \langle C_{i+s,j} \rangle) - \frac{P_p}{2} \langle C_{i,j} \rangle \prod_{\substack{s=-n/2 \\ s \neq 0}}^{n/2} (1 - \langle C_{i,j+s} \rangle).
\end{aligned}
\tag{3}
$$

Unlike proliferation Model 1, we now have both positive and negative terms in the discrete conservation statement associated with proliferation events, since Model 2 involves both the removal and deposition of agents at certain sites. The product terms in Eqs. (1) and (3) arise because successful proliferation events require that all sites between the original agent and the target site be vacant.

For both proliferation models, the discrete conservation statements are related to a PDE in the appropriate limit as $\Delta \to 0$ and $\tau \to 0$ and the discrete values of $\langle C_{i,j} \rangle$ are written in terms of a continuous variable C. To see this relationship, all terms in Eqs. (1) and (3) are expanded in a Taylor series about site (i, j), keeping terms up to $\mathcal{O}(\Delta^2)$. Dividing the resulting expression by τ, we then take limits as $\Delta \to 0$ and $\tau \to 0$ jointly, with the ratio Δ^2/τ held constant [20,21]. In the continuum limit, for both proliferation models, we obtain a PDE which can be written as

$$
\frac{\partial C}{\partial t} = D \nabla^2 C + \lambda C (1 - C)^n.
\tag{4}
$$

This is a key result implying that the continuum description of the discrete models is independent of the details of the proliferation mechanism and depends only on the separation distance. Therefore we expect that agent density profiles obtained from Model 1 will be identical to agent density profiles from Model 2 provided that the same separation distance is used for each model. This is surprising given that the proliferation mechanisms in Model 1 and Model 2 are very different. It is reassuring that the main differences in the continuum models are governed by the separation distance n. The value of n can be easily identified from time-lapse images.

The diffusivity and proliferation rate are given by

$$
D = \frac{P_m}{4} \lim_{\Delta, \tau \to 0} \left(\frac{\Delta^2}{\tau}\right), \qquad \lambda = \lim_{\tau \to 0} \left(\frac{P_p}{\tau}\right).
\tag{5}
$$

These conditions imply that $P_p = \mathcal{O}(\tau)$. We emphasize that the continuum model is valid as $\Delta \to 0$ and $\tau \to 0$ jointly with the ratio Δ^2/τ held constant and $P_p = \mathcal{O}(\tau)$. Since discrete simulations must be performed for finite Δ and τ, we expect that the continuum model will match the discrete model only for small values of P_p [5].

Eq. (5) connects the parameters in the continuum model (D, λ) to the parameters in the discrete model (P_m, P_p, Δ, τ). An important parameter for biological applications is P_p/P_m, which compares the relative frequency of proliferation and motility events for isolated agents. Biological observables from which this ratio may be estimated are the mitotic rate λ (related to the doubling time t_d, by $\lambda = \log_e 2/t_d$), the cell diffusivity D and the cell diameter Δ, used as our lattice spacing. Eq. (5) gives

$$
\frac{P_p}{P_m} = \frac{\Delta^2 \lambda}{4D}.
\tag{6}
$$

Note that the choice of τ affects the values of P_p and P_m individually, but not their ratio. A typical diffusivity is $D = 1 \times 10^{-6}$ mm^2/s [10,22]. For NCCs, a typical cell diameter and doubling time are $\Delta \approx 20$ μm and $t_d = 18$ h [3]. This gives $P_p/P_m \approx 0.001$, meaning that NCC proliferation events occur far less often than motility events, as seen in time-lapse movies of NCC invasion [2,4]. Using estimates of D and Δ here, with each non-dimensional time step representing 100 s, we can simulate NCC invasion by setting $P_m = 1$ and $P_p = 0.001$, corresponding to the parameter values in Fig. 3. Since $P_p \ll 1$ simulations must be performed for a sufficiently large number of time steps to see the number of agents increase significantly. For example, for an individual simulation shown in Fig. 3(a)–(b), it took 3000 non-dimensional time steps, or 3.47 days, for the population to increase from 820 to 2500 agents for Model 1 with $n = 1$.

In addition to developing a PDE to predict the distribution of agent density, we also develop a continuum model to describe the evolution of the average position of a tagged agent within the invasive population. For both models, if site (i, j)

is occupied, then the expected displacement of that agent during the next time step is

$$\delta p_x = \Delta \frac{P_m}{4}(1 - \langle C_{i+1,j}\rangle) - \Delta \frac{P_m}{4}(1 - \langle C_{i-1,j}\rangle) \tag{7}$$

$$\delta p_y = \Delta \frac{P_m}{4}(1 - \langle C_{i,j+1}\rangle) - \Delta \frac{P_m}{4}(1 - \langle C_{i,j-1}\rangle). \tag{8}$$

We note that δp_x and δp_y are independent of the proliferation mechanism and the value of P_p. This occurs because Model 1 does not involve any change in position of a proliferative agent during a proliferation event. For Model 2, a proliferative tagged agent at (x, y) produces two daughter agents placed at either $(x + n\Delta/2, y)$ and $(x - n\Delta/2, y)$ or $(x, y + n\Delta/2)$ and $(x, y - n\Delta/2)$ and we chose to track one of the daughter agents in Model 2 with equal probability $1/2$. This means that the changes in position of the two daughter agents are equal in length and opposite in direction, ensuring that their symmetric contributions to δp_x and δp_y in Eqs. (7) and (8) cancel.

Following the same procedure to obtain Eq. (4), we divide Eqs. (7) and (8) by τ and expand all terms about (i, j) in a Taylor series up to $\mathcal{O}(\Delta^2)$. Keeping Δ^2/τ constant, we let $\Delta \to 0$ and $\tau \to 0$ jointly giving

$$\frac{dp_x}{dt} = -2D\frac{\partial C}{\partial x}, \tag{9}$$

$$\frac{dp_y}{dt} = -2D\frac{\partial C}{\partial y}. \tag{10}$$

The solution of these differential equations gives $p_x(t)$ and $p_y(t)$, which are the coordinates of the average trajectory of a tagged agent initially at position $(p_x(0), p_y(0))$. These trajectories are called pathlines because of the analogy with potential flow [14].

Although the pathline models are independent of P_p, proliferation still influences the pathline models since proliferation affects the solution of Eq. (4) which, in turn, influences the pathlines through the appearance of $\partial C/\partial x$ and $\partial C/\partial y$ in Eqs. (9) and (10).

Where possible the continuum models developed here will be simplified and solved analytically. When this is not possible we solve the continuum models numerically. The solution of Eq. (4) is approximated with a finite difference method using a constant grid spacing δx and implicit Euler stepping with constant time steps δt. Picard iteration with convergence tolerance ϵ, is used to solve the resulting nonlinear equations [13]. Eqs. (9) and (10) are also solved numerically using a technique described by Simpson et al. [14].

## 4. Comparing continuum and discrete cell invasion models

We now compare the solution of Eq. (4) with simulations shown previously in Fig. 3. These simulations are equivalent to a one-dimensional problem since the initial condition and the boundary conditions impose no asymmetry in the vertical direction. We compare column averaged occupancy data from the discrete simulations with the numerical solution of a one-dimensional version of Eq. (4) [13,14]. Column averaged density data from the discrete simulations are further averaged over 40 identically prepared realizations. Results in Fig. 4 demonstrate an excellent correspondence between the discrete simulation data and the solution of Eq. (4) for Model 1 with $n = 1, 2, 3$ and Model 2 with $n = 2, 4, 6$. In all cases the initial local density near $x = 200$ decreases during the early part of the simulation as the agents spread making more room on the lattice near $x = 200$. At later times the local density around $x = 200$ increases as proliferation events take place. Comparing the simulation profiles for the different proliferation mechanisms we see different evolution behavior. For example, with Model 1 and $n = 1$, the density near $x = 200$ increases quickly after $t = 1000$, while for Model 2 with $n = 6$, we are yet to see the density near $x = 200$ increase during the time interval considered. This difference makes sense physically as we expect there to be more aborted proliferation events for Model 2 with larger n compared with Model 1 and small n for the initially close-packed group of agents.

Results in Fig. 4 confirm that the details of the proliferation mechanism do not affect the density profiles provided that the separation distance is the same. For example the discrete density profiles in Fig. 4(b) and (d) for $n = 2$ are indistinguishable even though the proliferation mechanism is different in each case. Further simulations (not shown) confirm that this is also true for other separation distances, such as $n = 4$. Given that the details of the proliferation mechanism are unimportant at the continuum-level for the same value of n and that ENS development corresponds with Model 2, we will focus on Model 2 from this point onward.

5. Divergence between the continuum and discrete models

The ability of the continuum models developed in Section 3 to replicate averaged discrete data depends on several assumptions, namely: (i) that the occupancies of lattice sites are independent, (ii) that $P_p = \mathcal{O}(\tau)$ as $\tau \to 0$, and (iii) the truncated Taylor series accurately relate the occupancies of sites on the lattice. The failure of any of these assumptions could mean that the continuum model is invalid. For example we expect that as P_p increases the continuum and discrete models
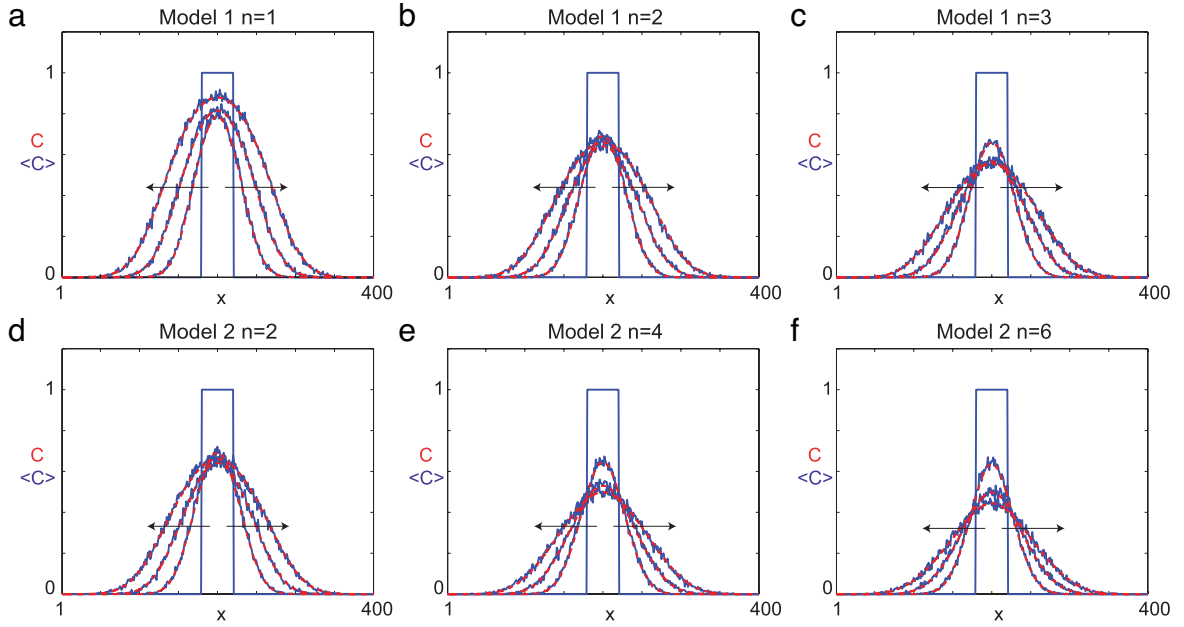
Fig. 4. Averaged agent density data for $P_m = 1$ and $P_p = 0.001$. All sites with $180 \leq x \leq 220$ are initially occupied. The column density of agents, averaged over 40 identically prepared simulations, are shown at $t = 0, 1000, 2000, 3000$ (solid blue) and compared with the solution of Eq. (4) (dotted red) with the arrows showing the direction of increasing time. Six sets of results are shown: results for proliferation Model 1 with $n = 1, 2, 3$ given in (a)–(c) and results for proliferation Model 2 with $n = 2, 4, 6$ are given in (d)–(f). The solution of Eq. (4) is obtained numerically with $\delta x = 0.25$, $\delta t = 0.1$ and $\epsilon = 1 \times 10^{-6}$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

will diverge [5]. Unfortunately, our limiting analysis does not give any insight into the details of this transition. To examine this transition in detail we performed a range of simplified simulations on a lattice with periodic boundary conditions on all boundaries. The lattice was initially occupied with a spatially uniform distribution of agents meaning that, on average, there were no spatial gradients in the system.

The advantage of working with the uniform density problem is that the governing PDE (Eq. (4)) collapses to an ordinary differential equation, namely

$$\frac{dC}{dt} = \lambda C (1 - C)^n. \tag{11}$$

To solve Eq. (11) we make use of the result

$$\frac{1}{C(1-C)^n} = \frac{1}{C} + \sum_{k=1}^{n} \frac{1}{(1-C)^k}. \tag{12}$$

Using Eq. (12) with $C(0) = C_0$ the solution of Eq. (11) can be written as

$$T = t\lambda = \begin{cases} \log_e \left[\dfrac{C(1 - C_0)}{C_0(1 - C)} \right], & n = 1, \\ \log_e \left[\dfrac{C(1 - C_0)}{C_0(1 - C)} \right] + \sum_{k=1}^{n-1} \left[\dfrac{1}{k(1 - C)^k} - \dfrac{1}{k(1 - C_0)^k} \right], & n \geq 2. \end{cases} \tag{13}$$

Since we are interested in comparing the solution of Eq. (11) with simulation data for different values of P_p, the solution (Eq. (13)) is written in terms of non-dimensional time $T = \lambda t$. This allows us to take simulation data for different values of P_p and collapse the solution profiles onto a universal curve independent of P_p.

For the uniform density problem we performed a range of simulations on a lattice of size I × J. Each site was initially occupied with probability $C_0 \in [0, 1]$ and simulations were performed with $P_m = 1$ and different values of P_p. Since there were no spatial gradients in the system, the average lattice site occupancy could be evaluated using

$$\langle C \rangle = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} C_{i,j}. \tag{14}$$

To compare continuum and discrete models, averaged values of $\langle C \rangle$ over several identically prepared realizations were compared with the solution to Eq. (11), given by Eq. (13).
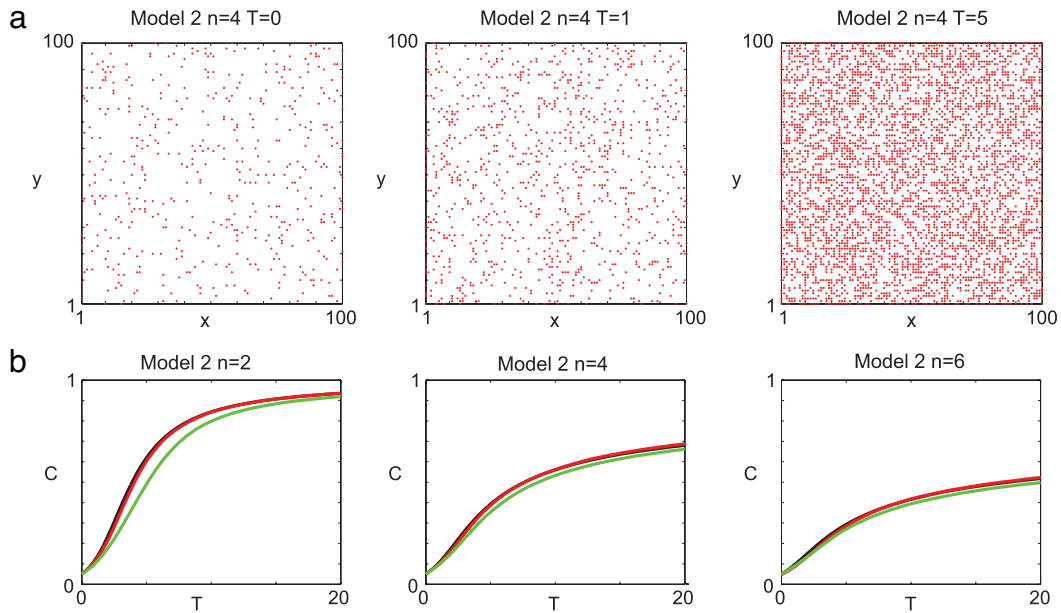
Fig. 5. Uniform density results for Model 2 are obtained by randomly populating a lattice of size 100×100 so that 5% of sites are initially occupied and performing simulations with $P_m = 1$ and different values of P_p. Snapshots of a simulation are shown in row (a) with $n = 4$, and the distribution of agents is shown at $T = 0, 1, 5$ with $P_p = 0.001$. Simulation data averaged over 40 identically prepared realizations are given in row (b). The analytical solution of the continuum model (Eq. (13)) (black) is superimposed on simulation data for $P_p = 0.001$ (red) and $P_p = 0.5$ (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Row (a) of Fig. 5 shows three snapshots of a single realization with proliferation Model 2 for $n = 4$ confirming the absence of spatial gradients in the system during the period of the simulation. Although the motility parameter does not appear in the simplified model (Eq. (11)), it is critical that motility is included and $P_m > 0$. Indeed, if we simulate this problem with $P_m = 0$, we observe clusters of agents that grow in size and eventually coalesce [17,23]. Agent clustering occurs when the ratio P_p/P_m is sufficiently large [17]. When clusters of agents are present, the independence assumption leading to Eqs. (1) and (3) is violated. Clustering of agents also means that averaging the occupancy across the lattice using Eq. (14) is inappropriate.

Comparing the discrete data and Eq. (13) in row (b) of Fig. 5 reveals certain trends that cannot be deduced from the limiting analysis. Simulation data with $P_p = 0.001$ matches the continuum model reasonably well for all values of n considered, whereas the simulation data with $P_p = 0.5$ does not always match the continuum result. The quality of the continuum–discrete comparison depends on the discrete mechanism as P_p increases and the comparison improves as n increases. This is best illustrated with the green curves for $P_p = 0.5$ where the continuum–discrete comparison is very poor when $n = 2$, and the comparison improves as n increases. Further simulation data at intermediate values of P_p (not shown) follow the same trends.

The computational data in Fig. 5 gives more information than the limiting analysis which simply tells us that provided that we consider simulation data with $\Delta \rightarrow 0, \tau \rightarrow 0, P_p = \mathcal{O}(\tau)$ and $P_m > 0$ we expect to see good correspondence between the discrete and continuum models. The analysis does not give any insight into how or why the discrete and continuum models diverge as these limiting conditions are not satisfied as P_p increases sufficiently. Further research to analyze this transition is required. However, for practical purposes in terms of modelling ENS development, this limitation is of no consequence as the relevant parameters are $P_m = 1$ and $P_p = 0.001$ and the simulation data matches the continuum models very well.

The results in Fig. 5 illustrate the complicated relationship between the discrete and continuum models. Without comparing the discrete and continuum models in this way, we might have incorrectly anticipated that the continuum–discrete comparison would have been most favorable for low n since the truncation error in the truncated Taylor series increases with n. Alternatively, we may have anticipated that the continuum–discrete comparison would be most favorable for higher values of n as the process of depositing agents on remote lattice sites could reduce agent clustering. As we have shown in Fig. 5, the relationship between the continuum and discrete models is sufficiently complex that these intuitive arguments are unable to explain the observed trends. Instead we rely on comparing the continuum and discrete models using simulation data to obtain a deeper understanding of the relationship between these models.

Fig. 5 shows that care must be taken when analyzing cell proliferation experiments. A standard experiment to estimate the proliferation rate is to measure how the density of a uniformly distributed population of cells grows with time, as shown in row (a) of Fig. 5 [10]. In the absence of any information about the relevant proliferation mechanism, it is possible to take
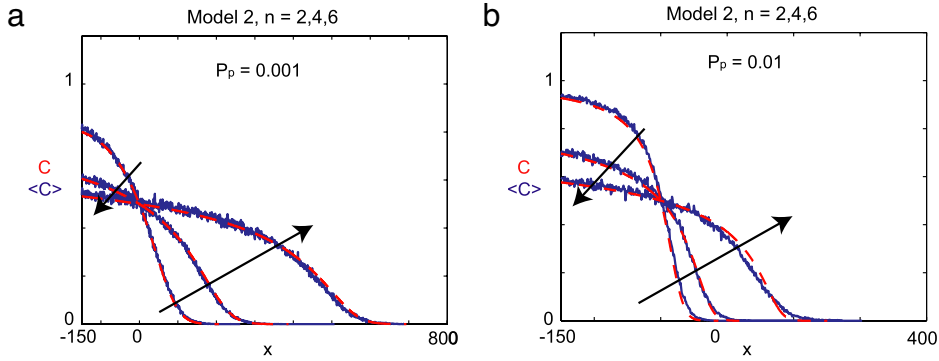
Fig. 6. Averaged agent density data for Model 2 with $P_m = 1$ and different values of P_p are compared with the solution of Eq. (4) for simulations performed over a large enough time for traveling waves to develop. Each site with $1 \leq x \leq 10$ is initially occupied. Column averaged simulation data, averaged over 40 identically prepared realizations are shifted so that $C(0, t) = 0.5$. Simulations were performed until the stopping time criteria described in the text was reached. Simulation profiles (solid blue) are compared with the solution of Eq. (4) (dotted red). Results in subfigures (a) and (b), are for $P_p = 0.001$ and $P_p = 0.01$ respectively. The arrows indicate the direction of increasing n. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

any of the growth curves in row (b) of Fig. 5 and fit the observed data to a logistic curve (Eq. (13) with $n = 1$) to estimate λ. Comparing the shape of the different growth curves in Fig. 5 shows that the population grows differently for different values of n. Therefore, simply fitting a logistic curve without any detailed knowledge of the particular proliferation mechanism can be meaningless.

6. Comparing continuum and discrete traveling waves

A key feature of cell invasion systems is the existence of constant speed invasion fronts, the speed of which can be characterized experimentally. For example, Maini et al. [22] measured the rate at which a population of cells closed a scrape wound *in vitro*, while Druckenbrod and Epstein [1] and Young et al. [4] measured the speed of advance of a population of NCCs.

A one-dimensional version of Eq. (4) is

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} + \lambda C(1 - C)^n, \tag{15}$$

which is a generalization of Fisher's equation ($n = 1$) that supports traveling wave solutions [11]. We expect that traveling wave solutions will exist for other values of n. A phase plane analysis of Eq. (15) shows that traveling wave solutions evolving from initial data with compact support move with speed $s = 2\sqrt{D\lambda}$ [24,25]. The wave speed is independent of the parameter n since it is determined by the stability about $C = 0$ and the $(1 - C)^n$ factor is not involved. Therefore the speed of the traveling wave is independent of the details of the proliferation mechanism. Although the simulations in Fig. 3 had not fully developed into traveling waves, the fact that the wave speed is independent of the proliferation mechanism is consistent with our observation that the distance the invasive fronts moved in Fig. 3 was indistinguishable regardless of the proliferation mechanism.

We now compare continuum and discrete traveling wave solutions for a range of separation distances. Previous comparisons of continuum and discrete traveling wave solutions have focused on measuring either the wave speed alone [26], or approximating the width of the wavefront for the traditional proliferation model (Model 1 with $n = 1$) [5]. Here we focus on comparing the exact details of the shape of the invasive front for a range of separation distances n.

To investigate the traveling wave profiles we performed a suite of simulations on a lattice of size 2000×20. Periodic boundary conditions were imposed on the horizontal boundaries and reflecting boundary conditions were imposed on the vertical boundaries. Each lattice site with $1 \leq x \leq 10$ was initially occupied and simulations were performed for a sufficiently long period of time that a constant speed traveling wave formed. All simulations correspond to $P_m = 1$ and different values of P_p. In each case the column-averaged density profile was obtained by averaging over many identically prepared realizations. To compare results for different proliferation models, the position of each traveling wave was shifted so the profile was centered at $x = 0$ where $C = 0.5$. Results in Fig. 6 compare discrete profiles with a numerical solution of Eq. (15).

The numerical solution of Eq. (15) was obtained using the same initial condition and domain as the discrete simulations. As the numerical solution was generated the speed of the invasion fronts was approximated [27] and the numerical solution was stopped when the speed of invasion was approximately $s = 2\sqrt{D\lambda}$. For $P_p = 0.001$ it took until $t = 30{,}000$ for the invasion waves to form. When $P_p = 0.01$ it took until $t = 5000$ for the invasion waves to form.

Although the speed of the invasion waves in Fig. 6 is independent of the proliferation mechanism, the width of the wavefront varies dramatically depending on the details of the proliferation mechanism. There are several different ways to
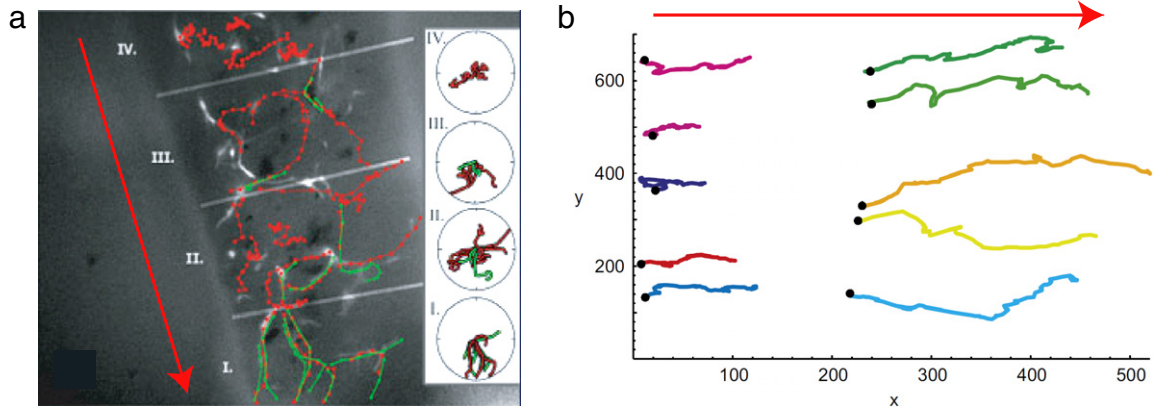
Fig. 7. (a) Experimental pathline results obtained by Druckenbrod and Epstein [2]. The direction of invasion of the NCC population is shown with the red arrow. Individual pathlines of cells within the population are give by the red and green trajectories. Polar diagrams are given in the inset, showing the length and direction of particular trajectories along the invasion wave that has been divided into four sections, I–IV. See Druckenbrod and Epstein [2] for a complete description of the experimental results. (b) Experimental pathlines from Cai et al. [10]. The direction of invasion for a scrape assay is shown with the red arrow. Pathlines of cells are shown, and in both cases the pathlines at the leading edge are biased to move in the same direction as the net invasion direction whereas tagged cells well behind the leading edge do not move as far. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: (a) Reproduced with permission from Wylie. (b) Reproduced with permission from Elsevier.

quantify the width of the invasion front [24]. Here we define the width as the distance between two contours at $C = C_{\max}$ and $C = C_{\min}$. Choosing $C_{\max} = 0.5$, $C_{\min} = 0.1$ and $P_p = 0.001$, the width is 90 for Model 2 with $n = 2$ while the width is 520 for Model 2 with $n = 6$.

7. Trajectory data

Time-lapse data provides cell-level details within an invading population [2]. Representative experimental results in Fig. 7(a) show that the movement of tagged NCCs at the leading edge of the invasion wave was biased to move in the same direction as the invasion front while the tagged cells behind the leading edge moved a smaller distance relative to their leading edge counterparts. These observations gave rise to the hypothesis that the behavior of NCC varies with respect to the position of the wavefront [2]. The same observations were made in an *in vitro* analysis of wound healing by Cai et al. [10] shown in Fig. 7(b).

To demonstrate how we can use the discrete invasion model and Eqs. (9) and (10) to recreate and interpret experimental trajectory data, we performed a range of simulations shown in Fig. 8. In each simulation we considered a lattice of size 400×20, with periodic boundary conditions along the horizontal boundaries and reflecting boundary conditions on the vertical boundaries. Initially all sites with $1 \le x \le 200$ were occupied, and a single tagged agent was placed at (199, 10). The trajectory of this tagged agent was recorded during each simulation. The tagged agent was identical to all other agents in the system. Simulation results in Fig. 8 show that the net movement of the tagged agent at $t = 1000$ has drifted in the positive x direction in all cases. The time scale of these simulations is not long enough for a traveling wave to form.

To generate pathline data from the discrete model we considered M identically prepared realizations of the simulations in Fig. 8 and averaged the horizontal coordinate of the mth tagged pathline denoted x_m, giving

$$\langle x(t) \rangle = \frac{1}{M} \sum_{m=1}^{M} x_m(t). \tag{16}$$

The averaged pathline data are compared with a numerical solution of Eq. (9) in Fig. 9 showing that the continuum and discrete pathline data match reasonably well for all separation distances considered. Consistent with experimental observations [2,10], all models predict that tagged agents at the leading edge moves in the positive x direction. Furthermore we see that differences in the pathline data between the different proliferation models are small for this problem.

Pathlines for tagged agents well behind the leading edge were very short regardless of the proliferation mechanism (not shown). This occurs because $\partial C / \partial x = 0$ in this region, and according to Eq. (9), the average position does not change with time. This is also consistent with experimental data [2,10].

In general, the difference between the pathlines of tagged agents at the leading edge compared to the pathlines of tagged agents well behind the leading edge is a universal result relevant to a range of initial distributions and densities of agents. Although results in Figs. 8 and 9 correspond to a Heaviside initial distribution, further simulations confirm that the same trends are observed for other initial distributions provided that $\partial C / \partial x < 0$ at the leading edge, and $\partial C / \partial x = 0$ behind the leading edge. This is true regardless of the density of agents behind the leading edge.
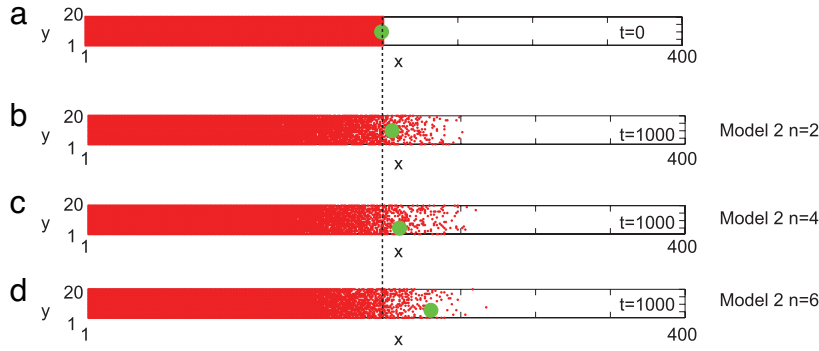
Fig. 8. (a) All simulations on a 400×20 lattice start with sites occupied (red) where $1 \leq x \leq 200$ and a tagged agent (enlarged green) is placed at $(199, 10)$. (b)–(d) Three simulations for Model 2 with $n = 2, 4, 6$. Each simulation corresponds to $P_m = 1$ and $P_p = 0.001$. In each case the net movement of each tagged agent is in the positive x direction during the 1000 time steps of the simulation. A vertical dashed line is placed at $x = 119$ to emphasize the movements of the tagged agents. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
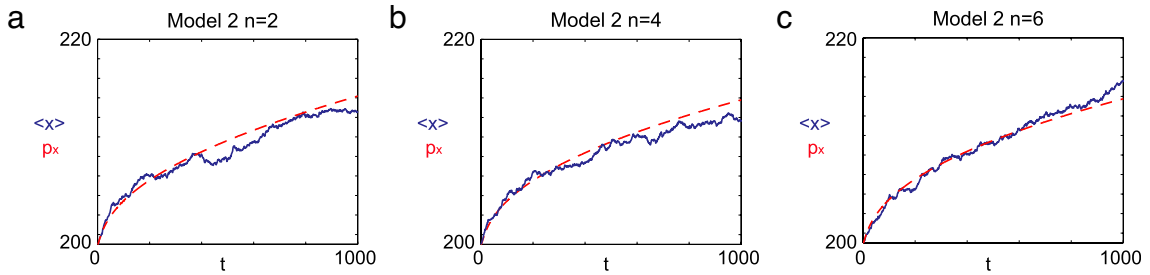
Fig. 9. Pathline data associated with the simulations in Fig. 8 are given for proliferation Model 2 with $n = 2, 4, 6$. Averaged simulation data, $\langle x(t) \rangle$ (solid blue), are compared with $p_x(t)$ (dotted red). All simulation data corresponds to the same domain, boundary conditions and initial conditions shown in Fig. 8 with $P_m = 1$ and $P_p = 0.001$. Averages are constructed using $M = 40$ identically prepared realizations. The numerical solutions of Eqs. (4) and (9) are obtained with $\delta x = 0.25$, $\delta t = 0.1$ and $\epsilon = 1 \times 10^{-6}$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In summary, we find that pathline data in the invasive populations shown in Fig. 9 are relatively insensitive to the details of the proliferation mechanism. These results are relevant to short-term simulations where the traveling wave profile is yet to form. This result is surprising given that proliferation is essential for the formation and maintenance of invasion waves [3,22]. Proliferation plays an indirect role as the key factor governing the pathline data is the geometry of the invasion wave as tagged agents at the leading edge move in the same direction as the invasion wave, while the net displacement of tagged agents behind the leading edge is far less than their leading edge counterparts [2,10]. All discrete pathline data shown here compare well with the solution of the corresponding continuum models since we only presented results for sufficiently small P_p. Of course, the continuum–discrete comparison deteriorates as P_p increases.

8. Discussion and conclusions

In this work we have presented and analyzed two models of cell invasion. The models can describe population-level information in terms of spatial cell density profiles that are compatible with experimental data [1,22]. The models also capture individual-level information allowing us to visualize individual motility and proliferation events as well as pathline data, which are also observed experimentally [2,10].

Using time-lapse data to develop a discrete model enables us to replicate realistic proliferation events. Traditional proliferation models in an exclusion process involve proliferative agents depositing daughter agents on nearest-neighbor lattice sites [5,15–17]. We have proposed and analyzed two alternative models motivated by particular experimental observations. The discrete models are related to a family of reaction–diffusion equations. The continuum models do not depend on the exact details of the proliferation mechanisms considered here provided that the separation distance between the placement of agents during a proliferation event is the same.

The invasion models give us new insight into the utility of different kinds of experimental data. Since proliferation is essential for the formation and maintenance of invasion waves [3], we might have anticipated that differences in the proliferation mechanism would have a significant impact on pathline data. Our modeling indicates that this is not always the case. Instead, pathline data is governed by the geometry of the invasion wave. From this we conclude that an experiment aimed at investigating the details of the proliferation mechanism in an invasive system ought not to focus on pathline data.

Instead the experiment ought to focus on measuring wave speed data, wavefront width data, or collecting time-lapse images showing the details of proliferation events.

The averaging arguments presented here show that the discrete mechanisms are related to a suite of continuum models for the agent density (Eq. (4)) and pathlines (Eqs. (9) and (10)). These models are valid in the limit as $\Delta \to 0$, $\tau \to 0$ and $P_p = \mathcal{O}(\tau)$ with the ratio Δ^2/τ held constant and $P_m > 0$. Since we must always perform discrete simulations with finite Δ and τ, we expect to see a transition where the continuum–discrete comparison is good for some parameter values and poor for others. We found that all continuum–discrete comparisons were good for sufficiently small P_p. However as P_p increased, the value of P_p at which the continuum–discrete comparison became poor increased as the separation distance n increased. This means that the new proliferation model developed here for ENS development with a relatively large separation distance $n = 6$ is both biologically relevant and mathematically convenient compared to traditional models (Model 1 with $n = 1$) as the continuum–discrete comparison is more robust over a wider range of parameters. We also emphasize that the continuum models developed here are only valid if we have a sufficient amount of motility in the system and $P_m > 0$. With little or no motility, agent proliferation leads to local clustering [17,23] which violates the independence assumptions underlying Eqs. (1) and (3).

All analysis and simulation data presented here correspond to undirected motility. We focused on undirected motility for two reasons. First, this is the simplest possible motility mechanism [13,14,18]. Second, our modeling has been inspired by NCC invasion and time-lapse data describing the movement of isolated NCCs in the absence of proliferation shows that they move along random and unbiased trajectories [4]. We repeated the analysis and simulations presented here using other motility mechanisms including biased motility [14] and adhesive motility [28]. These alternatives lead to different flux terms in the continuum models giving a different advection-diffusion-reaction PDE with the same requirement that $P_p = \mathcal{O}(\tau)$ as $\tau \to 0$. Equivalent continuum–discrete comparisons for these alternative motility mechanisms were made and the exact same trends were observed: the continuum–discrete comparison is good provided that P_p is sufficiently small and that the continuum–discrete comparison becomes poor as P_p increases.

Acknowledgements

References

[1] N.R. Druckenbrod, M.L. Epstein, Dev. Biol. 287 (2005) 125–133.
[2] N.R. Druckenbrod, M.L. Epstein, Dev. Dyn. 236 (2007) 84–92.
[3] M.J. Simpson, D.C. Zhang, M. Mariani, K.A. Landman, D.F. Newgreen, Dev. Biol. 302 (2007) 553–568.
[4] H.M. Young, A.J. Bergner, R.B. Anderson, H. Enomoto, J. Milbrandt, D.F. Newgreen, P.M. Whitington, Dev. Biol. 270 (2004) 455–473.
[5] T. Callaghan, E. Khain, L.M. Sander, R.M. Ziff, J. Stat. Phys. 122 (2006) 909–924.
[6] G.J. Pettet, H.M. Byrne, D.L.S. McElwain, J. Norbury, Math. Biosci. 136 (1996) 35–63.
[7] J.A. Sherratt, Proc. R. Soc. Lond. A 456 (2000) 2365–2386.
[8] K.R. Swanson, C. Bridge, J.D. Murray, E.C. Alvord Jr., J. Neurol. Sci. 216 (2003) 1–10.
[9] B.C. Thorne, A.M. Bailey, D.W. DeSimone, S.M. Peirce, Birth Defects Res. C 81 (2007) 344–353.
[10] A.Q. Cai, K.A. Landman, B.D. Hughes, J. Theoret. Biol. 245 (2007) 576–594.
[11] R.A. Fisher, Ann. Eugenics 7 (1937) 353–369.
[12] B.H. Bradshaw-Hajek, P. Broadbridge, Math. Comput. Model. 39 (2004) 1151–1163.
[13] M.J. Simpson, K.A. Landman, B.D. Hughes, Physica A 388 (2009) 399–406.
[14] M.J. Simpson, K.A. Landman, B.D. Hughes, Phys. Rev. E 79 (2009) 031920.
[15] M. Bramson, P. Calderoni, A. De Masi, P. Ferrari, J. Lebowitz, R.H. Schonmann, J. Stat. Phys. 45 (1986) 905–920.
[16] A.R. Kerstein, J. Stat. Phys. 45 (1986) 921–931.
[17] E. Khain, L.M. Sander, Phys. Rev. E 77 (2008) 051129.
[18] T.M. Liggett, Stochastic Interacting Systems: Contact, Voter and Exclusion Processes, Springer-Verlag, Berlin, 1999.
[19] D. Chowdhury, A. Schadschneider, K. Nishinari, Phys. Life Rev. 2 (2005) 318–352.
[20] E.A. Codling, M.J. Plank, S. Benhamou, J. R. Soc. Interface 5 (2008) 813–834.
[21] B.D. Hughes, Random Walks and Random Environments, vol. 1, Oxford University Press, 1995.
[22] P.K. Maini, D.L.S. McElwain, D.I. Leavesley, Tissue Eng. 10 (2004) 475–482.
[23] K. Zygourakis, R. Bizios, P. Markenscoff, Biotechnol. Bioeng. 38 (1991) 459–470.
[24] J.D. Murray, Mathematical Biology I: An Introduction, Springer, Heidelberg, 2002.
[25] J. Canosa, IBM J. Res. Dev. 17 (1973) 307–313.
[26] M.J. Simpson, A. Merrifield, K.A. Landman, B.D. Hughes, Phys. Rev. E 76 (2007) 021918.
[27] M.J. Simpson, K.A. Landman, Appl. Math. Lett. 19 (2006) 604–612.
[28] M.J. Simpson, K.A. Landman, B.D. Hughes, A.E. Fernando, Physica A 389 (2010) 1412–1424.