Bull Math Biol (2013) 75:871–889
DOI 10.1007/s11538-013-9839-0

Society for Mathematical Biology

ORIGINAL ARTICLE

Experimental and Modelling Investigation of Monolayer Development with Clustering

Matthew J. Simpson · Benjamin J. Binder · Parvathi Haridas · Benjamin K. Wood · Katrina K. Treloar · D.L. Sean McElwain · Ruth E. Baker

Received: 13 December 2012 / Accepted: 28 March 2013 / Published online: 13 April 2013

Abstract Standard differential equation-based models of collective cell behaviour, such as the logistic growth model, invoke a mean-field assumption which is equivalent to assuming that individuals within the population interact with each other in proportion to the average population density. Implementing such assumptions implies that the dynamics of the system are unaffected by spatial structure, such as the formation of patches or clusters within the population. Recent theoretical developments have introduced a class of models, known as moment dynamics models, which aim to account for the dynamics of individuals, pairs of individuals, triplets of individuals, and so on. Such models enable us to describe the dynamics of populations with clustering, however, little progress has been made with regard to applying moment dynamics models to experimental data. Here, we report new experimental results describing the formation of a monolayer of cells using two different cell types: 3T3 fibroblast cells and MDA MB 231 breast cancer cells. Our analysis indicates that the 3T3 fibroblast cells are relatively motile and we observe that the 3T3 fibroblast monolayer forms without clustering. Alternatively, the MDA MB 231 cells are less motile and we observe that the MDA MB 231 monolayer formation is associated with significant clustering. We calibrate a moment dynamics model and a standard mean-field model to both data sets. Our results indicate that the mean-field and moment dynamics models provide similar descriptions of the 3T3 fibroblast monolayer formation whereas these two models give very different predictions for the MDA MD 231 monolayer formation. These outcomes indicate that standard mean-field models of collective cell behaviour are not always appropriate and that care ought to be exercised when implementing such a model.

M.J. Simpson (✉) · B.J. Binder · P. Haridas · B.K. Wood · K.K. Treloar · D.L.S. McElwain · R.E. Baker
Mathematical Sciences, Queensland University of Technology, Brisbane Queensland, Australia
e-mail: matthew.simpson@qut.edu.au

Keywords Cell migration · Cell proliferation · Monolayer development · Clusters · Patchiness

1 Introduction

Continuum models of collective population dynamics typically invoke a mean-field assumption implying that individuals within the population interact with each other in proportion to their average density (Law and Dieckmann 2000; Law et al. 2003; Law et al. 2009). The mean-field assumption can be invoked explicitly, such as in the case of considering a lattice-based discrete random walk model where an approximate partial differential equation (pde) or ordinary differential equation (ode) description is derived by assuming that the occupancy status of lattice sites are independent (e.g. Deroulers et al. 2009; Penington et al. 2011; Penington et al. 2012; Plank and Simpson 2012; Simpson et al. 2011). Alternatively, the mean-field assumption can be invoked implicitly, such as in the case of applying standard ode or pde descriptions of collective cell behaviour without necessarily considering the underlying discrete process (e.g. Maini et al. 2004a, 2004b; Sengers et al. 2007; Sherratt and Murray 1990; Swanson et al. 2003; Tremel et al. 2009; Murray 2002). Invoking the mean-field assumption implies that the dynamics of the system is unaffected by spatial structure, such as the formation of patches or clusters of individuals (Bolker and Pacala 1997; Law and Dieckmann 2000; Law et al. 2003; Simpson et al. 2010a).

Great progress has been made in developing continuum models that relax the mean-field assumption for various applications including surface chemistry reactions (Mai et al. 1993, 1994), interacting plant communities (Bolker and Pacala 1997; Law and Dieckmann 2000; Law et al. 2003), infectious disease progression (Dangerfield et al. 2008; Keeling et al. 1997; Sharkey 2008, 2011) and point particle birth–death processes (Young et al. 2001). Our recent work has focused on adapting these techniques to discrete models of cell migration, proliferation and death processes which include finite size effects by allowing, at most, only one agent to occupy a particular location in space. This theoretical work implies that rapid cell proliferation (Baker and Simpson 2010; Simpson and Baker 2011) or strong cell-to-cell adhesion (Johnston et al. 2012) can lead to clustering (Simpson et al. 2010a). Our theoretical work has shown that alternative models, based on studying the dynamics of individuals, pairs of individuals, triplets of individuals, and so on (Bolker and Pacala 1997; Illian et al. 2008; Law and Dieckmann 2000; Law et al. 2003; Murrell et al. 2004), together with an appropriate moment closure approximation (Hansen and McDonald 2006; Singer 2004; Raghib et al. 2011), gives an improved prediction of the system behaviour (Baker and Simpson 2010; Simpson and Baker 2011; Johnston et al. 2012). We refer to such models as *moment dynamics models*.

In this manuscript, we present a combined experimental and modelling study where our recently-developed moment dynamics models are applied to new experimental measurements. We consider a two-dimensional assay where an initially-uniform population of cells are placed on a tissue culture plate and monitored in

real time as the cells move and proliferate to form a monolayer. Two cell types are considered: 3T3 fibroblast cells (3T3 cells) and MDA MB 231 breast cancer cells (231 cells). Visual observations indicate that the 3T3 fibroblast cells remain approximately spatially uniform as the monolayer forms. Conversely, the 231 breast cancer cell monolayer displays significant clustering. Separate measurements of the random motility of the cells (cell diffusivity) and the population growth are made for both cell types allowing us to calibrate a standard mean-field model and a more sophisticated moment dynamics model to the growth data. Our results indicate that the mean-field model and the moment dynamics model give similar results for the 3T3 fibroblast cell monolayer formation. Alternatively, the mean-field model and the moment dynamics model lead to very different predictions for the 231 monolayer formation. The presence, or absence, of clustering in the experimental data is confirmed by separately calculating a spatial index which allows us to quantitatively assess whether the spatial distribution of individuals is uniform or clustered (Binder and Landman 2011; Phelps and Tucker 2006). We conclude by reiterating that continuum models based on the mean-field assumption should not be applied to population dynamics problems where cell clustering is present, and we also outline further extensions to our modelling and experimental work.

2 Methods

2.1 Experimental Methods

3T3 Culture Murine fibroblast 3T3 cells (Todaro and Green 1963) were cultured in Dulbecco's modified Eagle medium (Invitrogen, Australia) supplemented with 5 % fetal calf serum (FCS) (Hyclone, New Zealand), 2 mM L-glutamine (Invitrogen) and 1 % v/v Penicillin/Streptomycin (Invitrogen) in 5 % CO_2 at 37 °C. Monolayers of 3T3 cells were cultured in T175 cm^2 tissue culture flasks (Nunc, Thermo Scientific, Denmark). Cells were lifted just prior to confluence using 0.05 % trypsin (Invitrogen). Viable cells were counted using a Trypan blue exclusion test and a haemocytometer.

MDA MB 231 Culture This breast cancer cell line (Cailleau et al. 1974) was maintained in Dulbecco's modified Eagle medium (Invitrogen) supplemented with 10 % FCS (Hyclone) and 1 % v/v Penicillin/Streptomycin (Invitrogen). The cells were cultured in 5 % CO_2 at 37 °C. Monolayers of 231 cells were cultured in T175 cm^2 tissue culture flasks (Nunc, Thermo Scientific, Denmark). Cells were lifted just prior to confluence using 0.05 % trypsin (Invitrogen, Australia). Viable cells were counted using a Trypan blue exclusion test and a haemocytometer.

Imaging of Monolayer Formation Two different densities of cell suspension were used; 5,000, and 25,000 cells/100 μL. The cell suspension was carefully introduced into wells of a 24-well tissue culture plate so that some wells contained 5,000 cells initially ($N = 3$) and others contained 25,000 cells initially ($N = 3$). Cells were monitored in real time using a Leica AFLX 6000 widefield microscope. Image sequences were acquired at a magnification of 10x. A series of images, each covering an area of 640 μm × 480 μm, was recorded every hour over a period of 48 hours.

2.2 Modelling Methods

We consider a model of cell motility and proliferation which has been described, and analysed, previously (Baker and Simpson 2010). The model can be implemented using a discrete random-walk framework (Codling et al. 2008) and the continuum limit of the model can be obtained by techniques outlined in (Baker and Simpson 2010). In brief, the discrete model consists of a two-dimensional square lattice with lattice spacing Δ. Each site may be occupied by, at most, one agent. Each agent has a transition rate P_m per unit time of moving to another lattice site and a proliferation rate P_p per unit time of giving birth to another agent. We suppose that motility events take place so that a motile agent at location (x, y) will attempt to step to $(x, y \pm \Delta)$ or $(x \pm \Delta, y)$ such that each target site is chosen with equal probability of $1/4$. Similarly, a proliferative agent at site (x, y) will attempt to deposit a daughter agent at $(x, y \pm \Delta)$ or $(x \pm \Delta, y)$ such that each target site is chosen with equal probability $1/4$. The model is an exclusion process (Liggett 1999) in the sense that potential movement and proliferation events can only take place if the target site is vacant. We will consider the case where the initial distribution of agents is spatially uniform and we denote the number of agents on the lattice at time t by $Q(t)$. Discrete simulations of this process can be performed using the Gillespie algorithm (Baker and Simpson 2010; Gillespie 1977). Such simulations of this discrete model have been presented, at length, previously (Baker and Simpson 2010). Here, we focus on applying a moment dynamics description of this discrete process to a new set of experimental data so that we can explore the implications of invoking the mean–field assumption.

To model the discrete system we use k-point distribution functions (Mai et al. 1993, 1994), $\rho^{(k)}$ $(k = 1, 2, \ldots)$, to gain information about the correlations between occupancy of different lattice sites. Essentially, the $\rho^{(k)}$ functions are multivariate probability distribution functions describing the occupancy of k-tuplets of lattice sites. We use l, m and n to denote various lattice sites, and $\sigma_l \in \{0, A\}$ to be the lattice variable describing the state of site l. For $k = 1$ we have

$$\rho^{(1)}(A_l) = c_l, \qquad \rho^{(1)}(0_l) = 1 - c_l, \tag{1}$$

where c_l is the density of agents at site l. In other words, $\rho^{(1)}(A_l)$ is the probability of finding an agent at site l, while $\rho^{(1)}(0_l)$ is the probability of finding site l vacant. Our lattice and initial conditions are translationally invariant so that c_l represents the density of agents at any site l and we will henceforth drop the subscript notation. The assumption of translational invariance also simplifies our analysis since it means that we do not explicitly consider boundary effects.

For our initial condition, the distribution function with $k = 2$ depends only on the distance between two lattice sites, $r = |l - m| \geq \Delta$. The correlation function (Mai et al. 1993, 1994) can be defined as

$$F\big(|l - m|\big) = \frac{\rho^{(2)}(A_l, A_m)}{\rho^{(1)}(A_l)\rho^{(1)}(A_m)}. \tag{2}$$

The correlation function has a relatively straightforward physical interpretation. Setting $F(|l-m|) = 1$ implies that the occupancy status of sites l and m are independent.

Alternatively, if $F(|l-m|) \neq 1$ the occupancy status of sites l and m are correlated and the mean–field assumption is violated. If $F(r) > 1$, pairs of agents at distance r are more abundant than the mean-field density, whereas if $F(r) < 1$ pairs of agents at distance r are less abundant than the mean-field density. Intuitively, we expect that the occupancy status of distant sites will be uncorrelated so that we have an asymptotic condition $F(|l-m|) \to 1$ as $|l-m| \to \infty$.

Following the approach outlined in Baker and Simpson (2010), we now develop expressions for the time rate of change of the 1- and 2-point distribution functions. For the 1-point distribution functions we have

$$\frac{\mathrm{d}\rho^{(1)}(A_l)}{\mathrm{d}t} = P_m \sum_n \frac{\alpha_{n,l}}{4}\left[\rho^{(2)}(0_l, A_n) - \rho^{(2)}(A_l, 0_n)\right] + P_p \sum_n \frac{\alpha_{n,l}}{4}\rho^{(2)}(0_l, A_n), \tag{3}$$

where $\alpha_{n,l} = 1$ if l and n are nearest neighbour sites and $\alpha_{n,l} = 0$ otherwise. The summation in (3) is considered to be over n, which represents all sites on the lattice except site l. Each term on the right of (3) corresponds to a potential motility or proliferation event that would alter the occupancy of site l. For example, each positive term proportional to P_m represents the potential configuration where site l is vacant and one of the nearest neighbour lattice sites is occupied so that the agent in question could step to site l thereby increasing the density of site l. A similar interpretation of the remaining terms on the right of (3) can be given. Baker and Simpson (2010) show that (3) simplifies to

$$\frac{\mathrm{d}c}{\mathrm{d}t} = P_p c\big(1 - F(\Delta)c\big). \tag{4}$$

If we assume that the occupancies of neighbouring sites are independent, and $F(\Delta) \equiv 1$, Eq. (4) reduces to the logistic equation, $\mathrm{d}c/\mathrm{d}t = P_p c(1-c)$, which is the standard mean-field representation of this discrete process (Simpson et al. 2010b). We note that the logistic equation is independent of the motility rate, P_m, and we will discuss the consequences of this in Sects. 3.3 and 3.4.

To solve (4) without invoking the mean-field assumption, we must also solve for $F(\Delta)$, and we achieve this by considering the time evolution of the 2-point distribution functions. Here, we have

$$\begin{aligned}\frac{\mathrm{d}\rho^{(2)}(A_l, A_m)}{\mathrm{d}t} &= P_m \sum_{n\neq l} \frac{\alpha_{n,m}}{4}\left[\rho^{(3)}(A_l, 0_m, A_n) - \rho^{(3)}(A_l, A_m, 0_n)\right] \\ &\quad + P_m \sum_{n\neq m} \frac{\alpha_{n,l}}{4}\left[\rho^{(3)}(0_l, A_m, A_n) - \rho^{(3)}(A_l, A_m, 0_n)\right] \\ &\quad + P_p\left[\sum_{n\neq l} \frac{\alpha_{n,m}}{4}\rho^{(3)}(A_l, 0_m, A_n) + \sum_{n\neq m} \frac{\alpha_{n,l}}{4}\rho^{(3)}(0_l, A_m, A_n)\right] \\ &\quad + P_p \frac{\alpha_{l,m}}{4}\left[\rho^{(2)}(A_l, 0_m) + \rho^{(2)}(0_l, A_m)\right].\end{aligned} \tag{5}$$

Each term on the right of (5) corresponds to a potential motility or proliferation event that would alter the probability of finding a pair of agents at sites l and m. For example, the first term on the right of (5) represents a triplet configuration where site l is occupied, site m is vacant, and one of the nearest neighbour sites of m is occupied so that if that nearest neighbour agent moves into site m, a new pair of agents at l and m will be formed. All remaining terms on the right of (5) can be interpreted in a similar way. It is relevant to note that the time rate of change of the pair density, $\rho^{(2)}(A_l, A_m)$, can be either increased or decreased through motility events in the discrete model since agent movement can either lead to the formation of new pairs of agents, or the destruction of existing pairs of agents. Alternatively, the time rate of change of the pair density, $\rho^{(2)}(A_l, A_m)$, can only ever increase through proliferation events in the discrete model since agent proliferation can lead to the formation of new pairs but never leads to the destruction of existing pairs of agents.

Equation (5) can be combined with (2) and simplified to show how the correlation functions at each distance evolve in time, namely,

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}t}\big(|l-m|\big) = {} & \frac{P_m}{2}\sum_{n\neq l}\alpha_{n,m}\big[F\big(|l-n|\big) - F\big(|l-m|\big)\big] \\ & - 2P_p\big[1-cF(\Delta)\big]F\big(|l-m|\big) + \frac{P_p}{2c}\alpha_{l,m}\big[1-cF(\Delta)\big] \\ & + \frac{P_p}{2c^2}\sum_{n\neq l}\alpha_{n,m}\rho^{(3)}(A_l, 0_m, A_n). \end{aligned} \tag{6}$$

To close (6), we need to express the remaining $\rho^{(3)}$ terms as $\rho^{(2)}$ terms and we do this by using the Kirkwood superposition approximation (ksa), which can be written as (Hansen and McDonald 2006; Singer 2004)

$$\rho^{(3)}(A_l, A_m, A_n) = \frac{\rho^{(2)}(A_l, A_m)\rho^{(2)}(A_l, A_n)\rho^{(2)}(A_m, A_n)}{\rho^{(1)}(A_l)\rho^{(1)}(A_m)\rho^{(1)}(A_n)}. \tag{7}$$

The ksa allows us to write

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}t}\big(|l-m|\big) = {} & \frac{P_m}{2}\hat{F}\big(|l-m|\big) - 2P_p\big[1-cF(\Delta)\big]F\big(|l-m|\big) \\ & + \frac{P_p}{2c}\alpha_{l,m}\big[1-cF(\Delta)\big] \\ & + \frac{P_p}{2(1-c)}\big[1-cF(\Delta)\big]\big[1-cF\big(|l-m|\big)\big] \\ & \times\left[\hat{F}\big(|l-m|\big) + \sum_{n\neq l}\alpha_{n,m}F\big(|l-m|\big)\right], \end{aligned} \tag{8}$$

where $\hat{F}(|l-m|)$ is the discrete lattice Laplacian

$$\hat{F}\big(|l-m|\big) = \sum_{n\neq l}\alpha_{n,m}\big[F\big(|l-n|\big) - F\big(|l-m|\big)\big]. \tag{9}$$

To make predictions using the moment dynamics model, we solve (4) for c and (8) for $F(r)$ simultaneously (Baker and Simpson 2010). We solve a truncated system by solving (8) for $F(r)$ considering all appropriate increasing lattice distances

$\Delta \leq r \leq r_{\max}$ with $F(r) \equiv 1$ for $r > r_{\max}$ (Baker and Simpson 2010). For all results reported here we solved the truncated system by setting $r_{\max} = 5\Delta$ and re-solved the system with $r_{\max} = 10\Delta$ to ensure that the different truncations give similar results. In practice, our model involves relatively short-range correlations, and we find that setting $r_{\max} = 5\Delta$ is sufficient to capture these details. We note that the equivalent mathematical expression for (8) in Baker and Simpson (2010) contained a typographical error which has been corrected here. This typographical error did not affect the results in Baker and Simpson (2010), and it does not affect the results presented here.

3 Results

3.1 Visual Observations of Monolayer Development

Snapshots of the growth processes are shown in Fig. 1. Although images were recorded over a period of 48 hours, we found that it became increasingly difficult to distinguish between individual cells within the monolayer towards the end of this time period as the cell density increased significantly. To alleviate this, we choose to focus on data from the first 25 hours of the experiment where the cell density is sufficiently low that we can accurately identify the location of individual cells within the growing population. Images in Fig. 1(a)–(c) show the distribution of 3T3 fibroblast

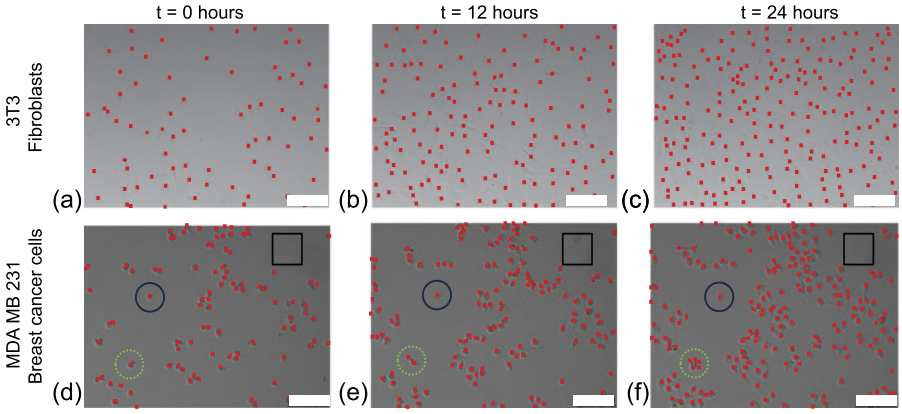

Fig. 1 Snapshots of the population of 3T3 fibroblast cells (**a**)–(**c**) and the population of MDA MB 231 breast cancer cells (**d**)–(**f**) at $t = 0$, 12 and 24 hours, as indicated. In each experiment, 25,000 cells were initially placed approximately uniformly into the 15.6 mm diameter wells of a 24-well tissue culture plate. The snapshots were taken at a random location within the well and our snapshots at different times focus on the same randomly-located window of dimension 640 μm × 480 μm. The location of individual cells within the window is indicated by a *red square* and the scale bar corresponds to 100 μm. Certain properties of the growth process of the 231 cells are indicated. The *solid square* illustrates a region of space that is initially vacant and remains vacant over the 24 hour period, the *solid circle* indicates the location of an individual 231 breast cancer cell that remains relatively stationary over the 24 hour period, and the *dotted circle* indicates the location of an individual 231 breast cancer cell that undergoes several proliferation events and relatively few motility events so that we observe the development of a patch or cluster of cells over time (Color figure online)

cells during the first 24 hours. To highlight the position of each cell within the growing population, we have superimposed a red square onto the centre of each cell. These three snapshots show the same field of view, which indicates that the cell density increases rapidly during this period. Importantly, we see that the spatial distribution of these cells appears to be uniform with no obvious development of patchiness or clustering over the 24 hours period shown in Fig. 1.

Images in Fig. 1(d)–(f) show the distribution of MDA MB 231 breast cancer cells during the first 24 hours. We superimposed a red square onto the centre of each cell to highlight the location of individual 231 cells amongst the growing population. These three snapshots show the same field of view, which indicates that, as for the 3T3 fibroblast cells, the number of the 231 cells increases rapidly during this period. Unlike the 3T3 cells, however, the spatial distribution of 231 cells is not uniform. In particular, we see that certain regions in the image that are initially vacant, such as the region illustrated with the black square (Fig. 1(d)), remains vacant at $t = 12$ and $t = 24$ hours (Fig. 1(e)–(f)). Unlike the 3T3 cells, it is possible to visually identify a particular 231 cell at $t = 0$, such as the cell highlighted in the solid circle (Fig. 1(d)), which can be clearly identified at $t = 12$ and $t = 24$ hours, indicating that this cell is relatively immobile (Fig. 1(e)–(f)). The mechanisms leading to the formation of clusters can also be quantitatively identified from the images in Fig. 1 since we can identify certain cells, such as the cell highlighted in the dotted circle (Fig. 1(d)), that proliferate several times during the period over which the images were recorded. Since the original mother cell and the new daughter cells undergo relatively little movement, we observe a discrete cluster, or patch of cells forming over time as a result (Fig. 1(e)–(f)).

In summary, our visual interpretation of the MDA MB 231 and 3T3 fibroblast cell growth process suggests that the 231 population displays significant spatial patterning whereas the 3T3 population remains relatively uniform. We will now apply both a standard mean-field model and the more detailed moment dynamics model to both these data sets to explore the significance of the role of the spatial patterning and cluster development in these experiments.

3.2 Estimating the Random Motility Rate

Our visual inspection of the data in Fig. 1 suggests that the 3T3 fibroblast cells are relatively motile whereas the MDA MB 231 breast cancer cells are less motile. To quantify this difference we estimate the motility of both cell types by analysing many individual cell trajectories. To characterise the intrinsic motility rate in the absence of cell-to-cell crowding effects, we analyse trajectories of isolated cells that do not interact with other cells over the duration of that particular trajectory. Since isolated cells were more easily identified within the low density experiments, where 5,000 cells were placed into each tissue culture well, we focus on trajectories from these low density experiments rather than the higher density experiments where 25,000 cells were placed into each tissue culture well.

Since cells are initially placed into the tissue culture well approximately uniformly, we do not expect, and we do not observe on average, any biased mo-

tion. To characterise the motility we estimate the squared displacement of both the x-coordinate and the y-coordinate of each trajectory

$$x^2(t) = \big(x(t) - x(0)\big)^2, \qquad y^2(t) = \big(y(t) - y(0)\big)^2, \tag{10}$$

where $x(t)$ is the x-coordinate of the location of the tagged cell after time t, and $y(t)$ is the y-coordinate of the tagged cell after time t. An estimate of the random motility coefficient (diffusivity), in each orthogonal direction, is obtained using a standard approach (Hughes 1995) by fitting a straight line to the data

$$x^2(t) = 2D_x t, \qquad y^2(t) = 2D_y t, \tag{11}$$

where D_x and D_y are the diffusivities in the orthogonal x and y directions, respectively (Cai et al. 2007; Hughes 1995). We note that this standard approach makes the simplifying assumption that the diffusivities are constant and do not change over time (Massey et al. 2012). Data in Fig. 2(b)–(c) show the $x^2(t)$ and $y^2(t)$ values for the particular cell shown in Fig. 2(a), obtained over a period of 10 hours. If we treat the trajectory of the cell in Fig. 2(a) alone, the straight line regression for this one trajectory, shown in Fig. 2(b)–(c), gives $D_x = 133.4\ \mu m^2/hour$, and $D_y = 70.7\ \mu m^2/hour$. We emphasise that these estimates correspond to data from one trajectory only and that the trajectory of one cell corresponds to one experimental realisation of the stochastic process. To obtain an average estimate of diffusivity, we repeated the same process for another 39 trajectories. In total, we measured $x^2(t)$ and $y^2(t)$ for 40 individual trajectories, giving us 40 estimates each of D_x and D_y. The duration for

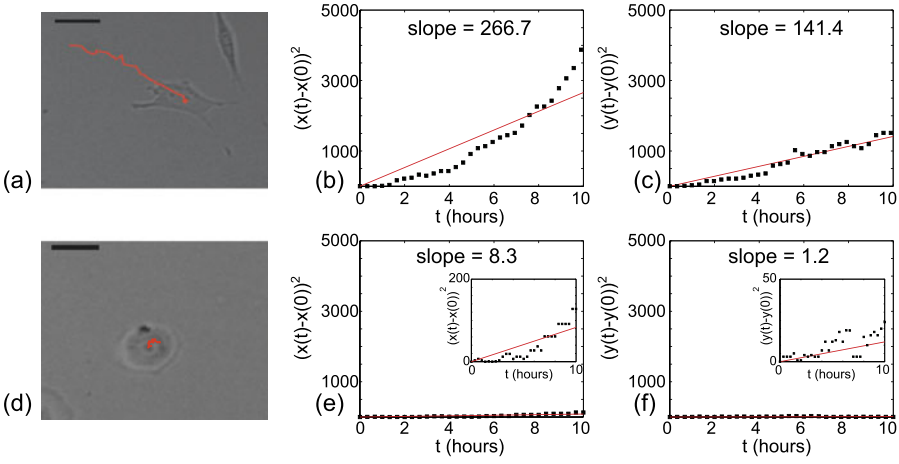

Fig. 2 Motility estimates of 3T3 fibroblast cells and MDA MB 231 breast cancer cells. An image of an isolated 3T3 cell is shown in (**a**) together with the observed trajectory of that cell recorded over a period of 10 hours. The scale bar is 25 μm indicating that $\Delta \approx 25$ μm. Results in (**b**)–(**c**) show the squared displacement data, $x^2(t)$ and $y^2(t)$, respectively, corresponding to the trajectory of the cell shown in (**a**). An image of an isolated 231 cell is shown in (**d**). The scale bar is 25 μm indicating that $\Delta \approx 25$ μm. Results in (**e**)–(**f**) show the squared displacement data, $x^2(t)$ and $y^2(t)$, respectively, corresponding to the trajectory of the cell shown in (**d**). Note that the squared displacement data in (**e**)–(**f**) is shown at the same scale as the equivalent data for the 3T3 trajectory in (**b**)–(**c**) to emphasise that the 231 cells are much less motile than the 3T3 cells. The details of the 231 squared displacement data, $x^2(t)$ and $y^2(t)$, together with the straight line regression, is shown in an appropriately re–scaled inset in (**e**)–(**f**) (Color figure online)

which we analysed each trajectory was different since some trajectories took a long time before that particular cell collided with others, whereas some trajectories encountered a cell-to-cell collision after a relatively short period. The shortest trajectory we considered was 3.9 hours, and the average duration of all 40 trajectories was 10.8 hours. Averaging our estimates over the sample of 40 trajectories gave us $\langle D_x(3T3)\rangle = 107.7 \pm 27.3\ \mu m^2/hour$ and $\langle D_y(3T3)\rangle = 99.7 \pm 18.8\ \mu m^2/hour$, where we have chosen to characterise the variability in our estimates using the standard error. These diffusivity estimates indicate that there is very little difference between the observed random motility in the two orthogonal directions, as we might have anticipated. Therefore, we will characterise the motility of 3T3 cells using a constant diffusivity, D, that is independent of direction which we obtain by averaging our 80 estimates to give $\langle D(3T3)\rangle = 103.7 \pm 23.0\ \mu m^2/hour$.

The diffusivity of the 231 breast cancer cells was estimated using the same procedure. Data in Fig. 2(e)–(f) show $x^2(t)$ and $y^2(t)$ for the cell shown in Fig. 2(d), which was analysed over a period of 10 hours. The straight line regression for this one particular trajectory gives $D_x = 4.2\ \mu m^2/hour$ and $D_y = 0.6\ \mu m^2/hour$, suggesting that this particular cell is far less motile than the 3T3 cell shown in Fig. 2(a). Averaging over 40 equivalent trajectories gives us $\langle D_x(231)\rangle = 6.1 \pm 1.4\ \mu m^2/hour$ and $\langle D_y(231)\rangle = 6.3 \pm 1.4\ \mu m^2/hour$, which again indicates that there is very little difference between the observed motility of cells in the x and the y, as expected. Therefore, we average the entire data set and treat the diffusivity as a constant that is independent of direction to give $\langle D(231)\rangle = 6.2 \pm 1.4\ \mu m^2/hour$. The shortest trajectory for the MDA MB 231 cells was 9 hours, and the average duration of all 40 trajectories was 24.1 hours.

Our estimates of diffusivity can be converted into an equivalent motility rate, P_m, since we have $D = (P_m \Delta^2)/4$ (Baker and Simpson 2010) giving $P_m(3T3) = 0.66 \pm 0.15\ hour^{-1}$ and $P_m(231) = 0.04 \pm 0.01\ hour^{-1}$. To obtain these estimates we assume that the diameter of both cells types is $\Delta \approx 25\ \mu m$, which is consistent with the images in Fig. 2(a) and (d). Comparing the motility rates indicates that the 231 cells are more than an order of magnitude less motile than the 3T3 cells. This major difference in cell motility will have no consequence when we apply the standard mean-field model to the cell growth process since the standard mean-field model, Eq. (4) with $F(\Delta) = 1$, is independent of P_m. Alternatively, we expect that the difference in motility rate will have a significant influence when we apply the moment dynamics model, given by (4) and (8), since the dynamics of the correlation functions depends on P_m.

3.3 Using Experimental Data to Estimate the Proliferation Rate: Traditional Mean-Field Model

We will first estimate the growth rate, P_p, for both cell types using a traditional mean-field model. To be consistent with the discrete model described in Sect. 2.2, we assume that the carrying capacity density corresponds to the maximum packing of circles on a square lattice. This implies that the 640 μm × 480 μm inspection window in Fig. 1 will approximately accommodate a maximum of $K = 500$ cells of diameter 25 μm. With this estimate, the non-dimensional density of cells is given by $c(t) =$

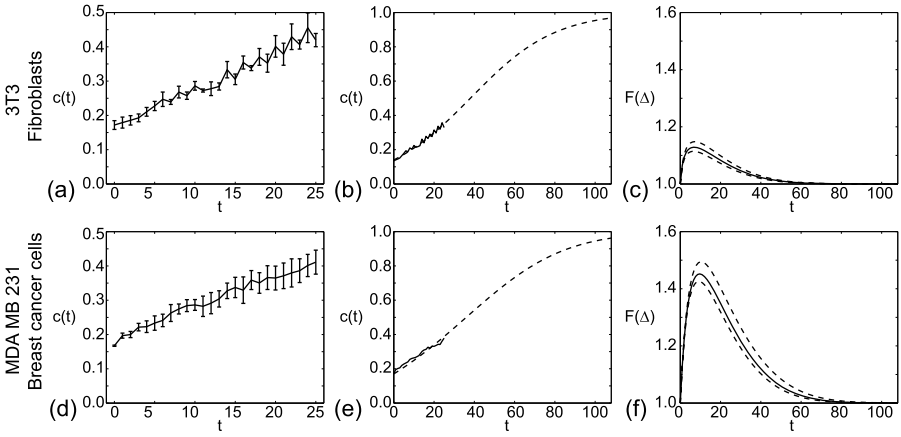

Fig. 3 Experimental density data in (**a**) and (**d**) shows the increase in cell density over the first 25 hours of the experiment for the 3T3 fibroblast cells and MDA MB 231 breast cancer cells, respectively. Results are presented as an average calculated for three identically-prepared experimental observations, and the error bar corresponds to the standard error. The density data in (**a**) and (**d**) are included in (**b**) and (**e**) (*solid lines*) together with the solution of (4) and (8) with P_p chosen to minimise the least-squares error between the experimental data and the solution of the moment dynamics model during the interval $t \leq 25$ hours. The calibrated model was also used to predict the density profile over a longer interval, $t \leq 108$ hours (*dotted*), to show the expected behaviour of the system as the density increases towards confluence. The profile in (**b**) corresponds to the moment dynamics model with $P_p(3T3) = 0.056$ hour^{-1} and $P_m(3T3) = 0.66$ hour^{-1}. The profile in (**e**) corresponds to the solution of the moment dynamics models with $P_p(231) = 0.069$ hour^{-1} and $P_m(231) = 0.04$ hour^{-1}. Profiles in (**c**) and (**f**) show the evolution of $F(\Delta)$ from the solution of (4) and (8) for the 3T3 and 231 data, respectively. Results are presented as a mean (*solid line*) together with an upper and lower bound (*dotted lines*), where the upper and lower bounds are obtained by solving the moment dynamics model with the parameters chosen to be the mean plus or minus one standard error. Profiles in (**c**) correspond to $P_p(3T3) = 0.056$ hour^{-1} and $P_m(3T3) = 0.66 \pm 0.15$ hour^{-1}. Results in (**f**) correspond to $P_p(231) = 0.069$ hour^{-1} and $P_m(231) = 0.04 \pm 0.01$ hour^{-1}. All numerical solutions of the moment dynamics model correspond to $r_{\max} = 5\Delta$ and $h = 0.01$ hours

$Q(t)/K$, where $Q(t)$ is the number of cells present in the inspection window at time t. Data in Fig. 3(a) and (d) show the increase in non-dimensional cell density with time for the first 25 hours of the experiment. The density curves represent averaged cell density data and were constructed using three identically-prepared experimental data sets. These data indicate that the density increases approximately 2.5 times during this time interval for both cell types.

The traditional mean-field model of monolayer development without clustering effects is given by (4), with $F(\Delta) \equiv 1$, which has the solution

$$c(t) = \frac{c(0)\mathrm{e}^{P_p t}}{1 + c(0)(\mathrm{e}^{P_p t} - 1)}. \tag{12}$$

To estimate the growth rate we fitted (12) to the density data in Fig. 3(a) and (d) for the first 25 hours of the experiment. The fitting was done by setting $c(0) = 0.18$ for the 3T3 fibroblast cells and $c(0) = 0.17$ for the 231 cells, as given by the average

data in Fig. 3(a) and (d). For each cell type we estimated the proliferation rate such that it minimised the least-squares error,

$$E = \sqrt{\frac{\sum_{i=1}^{k}[c_{\mathrm{e}}(t_i) - c_{\mathrm{m}}(t_i)]^2}{k}}, \tag{13}$$

where $c_{\mathrm{e}}(t_i)$ is the experimentally-observed density at time t_i, $c_{\mathrm{m}}(t_i)$ is the density predicted by the model at time t, and k is the number of observation points. We found that choosing $P_p(3\mathrm{T3}) = 0.055$ hour^{-1} and $P_p(231) = 0.055$ hour^{-1} gave a minimum value for E.

3.4 Estimating the Proliferation Rate: Moment Dynamics Model

To model the monolayer formation without invoking a mean-field assumption we apply (4) and (8) to the data in Fig. 3(a) and (d). We solve (4) and (8) numerically using a fourth order Runge–Kutta method with constant step size h (Baker and Simpson 2010) and we make the simplifying assumption that the initial distribution of cells are located at random giving $F(r) = 1, \forall r$ at $t = 0$.

Using the numerical solutions of (4) and (8), we found the proliferation rate, P_p, that minimised E in (13) for the data in Fig. 3(a) and (d). For the 3T3 fibroblast cells, this gave $P_p(3\mathrm{T3}) = 0.056$ hour^{-1}, which is just 2 % larger than the proliferation rate predicted using the traditional mean–field model, Eq. (12). For the 231 cells, we found an optimal proliferation rate of $P_p(231) = 0.069$ hour^{-1}, which is 25 % greater than the proliferation rate obtained by fitting the traditional mean-field model in Sect. 3.3.

Results in Fig. 3(b) and (e) show the evolution of the density, $c(t)$, as predicted by the solution of (4) and (8) using the proliferation rate chosen to provide the best match to the observed data. The solution of the moment dynamics model is shown over a period of 108 hours which is a sufficient time for the population density to approach the carrying capacity, $c(t) \approx 1$. The moment dynamics solution is superimposed on the observed density data collected over the first 25 hours of the experiment, confirming that our choices of $P_p(3\mathrm{T3}) = 0.056$ hour^{-1} and $P_p(231) = 0.069$ hour^{-1} provide excellent matches to the observed growth data. The corresponding evolution of the correlation function, $F(\Delta)$, is shown in Fig. 3(c) and (f). Here we see that $F(\Delta)$ for the 3T3 fibroblast data remains relatively close to unity for all time. This is consistent with our finding of a relatively large value of P_m for the 3T3 cells since a rapid rate of agent motility in the discrete model can act to reduce the density of pairs of agents that are produced by proliferation events, as illustrated in (5). In contrast, the evolution of $F(\Delta)$ for the 231 data increases well above unity before decaying to unity relatively slowly. This increase in the correlation function, $F(\Delta)$, for the 231 data is a consequence of the relatively small motility rate for these cells. The reduced motility means that the density of pairs of agents produced by proliferation in the equivalent discrete process is reduced slowly by relatively infrequent motility events.

Comparison of the evolution of the correlation function, $F(\Delta)$, in Fig. 3(c) and (f) is consistent with the differences we observe in our estimates of P_p using the mean-field and moment dynamics models. The 3T3 correlation function (Fig. 3(c)) remains relatively close to unity for all time which means that the moment dynamics model is

reasonably well approximated by the mean-field model since equation (4) simplifies to the standard mean-field logistic equation when $F(\Delta) \equiv 1$. This is also consistent with our observation that the calibrated moment dynamics model predicts a proliferation rate that is just 2 % larger than the mean-field model for the 3T3 fibroblast cell data. In this case, the difference in the growth rate predicted by the moment dynamics model and the traditional mean-field model is practically negligible, indicating that there is little advantage in using the moment dynamics model. Conversely, the correlation function for the 231 cells (Fig. 3(f)) increases to a maximum of approximately 1.45 and before decaying to unity relatively slowly, suggesting that we have a significant difference between the mean-field and the moment dynamics models for this data set. This is consistent with our parameter estimates which indicate that the growth rate in the moment dynamics model is 25 % larger than the proliferation rate predicted by the traditional mean-field model.

Figure 3(c) and (f) also provides an indication of the sensitivity of the results of the moment dynamics model to our estimates of the motility rate. The solid line in Fig. 3(c) and (f) shows the evolution of the correlation function for our estimates of the motility rates, $P_m(3T3) = 0.66$ hour^{-1} and $P_m(231) = 0.04$ hour^{-1}. We also superimpose lower and upper bounds of the correlation function, obtained by resolving (4) and (8) using the upper bounds ($P_m(3T3) = 0.81$ hour^{-1}, $P_m(231) = 0.05$ hour^{-1}) and lower bounds ($P_m(3T3) = 0.51$ hour^{-1}, $P_m(231) = 0.03$ hour^{-1}) of the motility rates, while holding all other parameters constant. We found that the solution of the moment dynamics model for the non-dimensional density, $c(t)$, was less sensitive to the variability in the motility rate than the correlation function, $F(\Delta)$. Therefore, we do not show upper and lower bounds for the non-dimensional density data in Fig. 3(b) and (e).

3.5 Spatial Statistics

Our results in Sects. 3.3 and 3.4 indicate that the development of the 231 monolayer is associated with significant short-range correlations and clustering that are neglected by the usual mean-field logistic description. Conversely, the development of the 3T3 fibroblast monolayer is relatively unaffected by short range correlations, and consequently we observe that the moment dynamics description of the growth rate is very similar to the standard mean-field description. We now provide an independent, but complementary, analysis of the spatial distribution of cells in the experiments. To do this, we consider a rectangular domain, of dimensions $X \times Y$, populated by $Q(t)$ cells, each of which occupies an average area s. We divide the spatial domain into M equally-sized bins, and at each time t count the number of cells in bin j to give $b(j,t)$. With this information, we can calculate the expected number of cells per bin, $Q(t)/M$, and the variance of the number of cells per bin. A scaled variance, or index (Binder and Landman 2011; Binder et al. 2011, 2012; Hackett-Jones et al. 2011; Phelps and Tucker 2006), can then be defined as

$$I(t) = \frac{M}{Q(t)^2(M-1)} \sum_{j=1}^{M} \left(b(j,t) - \frac{Q(t)}{M} \right)^2, \tag{14}$$

where $I(t) \in [0, 1]$. A value of $I(t) = 1$ corresponds to maximum segregation where all cells lie in one bin, whereas $I(t) = 0$ corresponds to a perfectly uniform distribution where each bin contains exactly the expected number of cells, $Q(t)/M$. Although the perfectly uniform distribution is theoretically possible, it is not often realised in practice. Instead, a more realistic situation occurs when each object is equally likely to reside in any bin which is known as the complete spatial randomness (CSR) state (Binder and Landman 2011; Phelps and Tucker 2006). Exact expressions for the index at the CSR limit have been derived previously (Binder and Landman 2011), and it has been shown that when the size of the object is much less than the size of the domain, $s \ll XY$, we can approximate the index at the CSR limit (Binder and Landman 2011; Binder et al. 2011, 2012; Hackett-Jones et al. 2011; Phelps and Tucker 2006) by the simplified expression

$$\hat{I}(t) \approx \frac{1}{Q(t)} - \frac{s}{XY}. \tag{15}$$

To make use of (14) and (15), we can take a given distribution of objects and a given spatial partition to calculate $I(t)$ and $\hat{I}(t)$. Averaging these results over several experimental realisations of the process gives $\langle I(t)\rangle$ and $\langle \hat{I}(t)\rangle$. Comparing estimates of $\langle I(t)\rangle$ and $\langle \hat{I}(t)\rangle$ provides a quantitative measure of the departure from the CSR state. Distributions with $\langle I(t)\rangle \leq \langle \hat{I}(t)\rangle$ are close to the CSR state (Hackett-Jones et al. 2011) whereas distributions with $\langle I(t)\rangle > \langle \hat{I}(t)\rangle$ are characterised by clusters or patches in the spatial distribution.

Results in Fig. 4(a) and (d) illustrate a snapshot of the distribution after 12 hours for the 3T3 and 231 cells, respectively. We will now describe how to calculate $\langle I(t)\rangle$ and $\langle \hat{I}(t)\rangle$ for these kinds of images. It is convenient to perform the spatial analysis in terms of pixels rather than the physical spatial dimensions since the images are electronically stored in dimensions of pixels. To do this, we convert the average cell area, $s = \pi 25^2/4$ μm^2, into pixels. The area of the spatial domain ($\approx$308,000 μm^2) is the same for both types of experimental images. However, the resolution of the images for the two cell types is different (3T3: 3.61920×10^5 pixels, 231: 2.03580×10^5 pixels), giving us two different values of $\hat{s}$ measured in pixels.

Figure 4(b) and (e) highlights the location of each cell in Fig. 4(a) and (d) superimposed on a regular partition of the domain. The profiles in Fig. 4(c) and (f) shows the temporal variation in $\langle I(t)\rangle$ and $\langle \hat{I}(t)\rangle$ for the 3T3 and 231 data, averaged over three independent realisations, respectively. For both cell types, at each time point, the average value of $\langle I(t)\rangle$ has been calculated for two different partitions. In each case, the finer partition contains approximately twice as many bins as the coarser partition. Our estimates of $\langle I(t)\rangle$ for both cell types appear to be approximately independent of the number of bins used in the partition as shown in Fig. 4(c) and (f). Comparing the average index values and the corresponding CSR limit shows that the average index for the 3T3 fibroblast cell distribution is slightly less than the CSR limiting value, indicating that the distribution of 3T3 cells is close to the CSR state. Conversely, the average index for the 231 cells is above the CSR limit, indicating that the distribution of 231 cells is nonuniform. This difference confirms our visual interpretation of our results in Fig. 1 and is also consistent with our modelling results in Sects. 3.3 and 3.4.

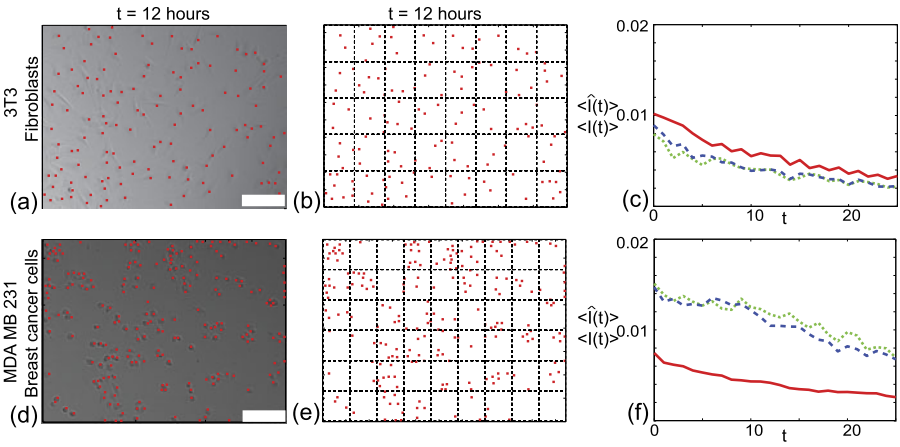

Fig. 4 Spatial statistics analysis. Images in (**a**) and (**d**) show the distribution of 3T3 fibroblast and MDA MB 231 cells after $t = 12$ hours, respectively. The scale bar corresponds to 100 μm. Results in (**b**) and (**e**) show the location of each cell in (**a**) and (**d**), together with the partition of the domain used to calculate $I(t)$. The partition in (**b**) corresponds to $M = 40$ equally-sized bins whereas the partition in (**e**) corresponds to $M = 54$ equally-sized bins. Profiles in (**c**) and (**f**) show $\langle I(t)\rangle$ and $\langle \hat{I}(t)\rangle$ for the 3T3 and 231 cells, respectively. The *red* (*solid*) *curves* in (**c**) and (**f**) correspond to the CSR limiting value, $\langle \hat{I}(t)\rangle$, averaged over three images at each time point. The *green* (*dotted*) *curves* in (**c**) and (**f**) correspond to the index, $\langle I(t)\rangle$, averaged over three images at each time point, for a partition with $M = 40$ and $M = 54$, respectively. The *blue* (*dashed*) *curves* in (**c**) and (**f**) correspond to the index, $\langle I(t)\rangle$, averaged over three images at each time point, for a finer partition with $M = 80$ and $M = 90$, respectively. The size of the 3T3 cells is $\hat{s} = 577$ pixels, and the size of the 231 cells is $\hat{s} = 325$ pixels (Color figure online)

4 Discussion and Conclusion

Traditional models of collective cell behaviour typically invoke a mean-field assumption implying that individuals within the population interact with each other in proportion to their average mean-field density (Law and Dieckmann 2000; Law et al. 2003). Invoking the mean-field assumption, such as in the application of reaction–diffusion equations (Maini et al. 2004a, 2004b; Sengers et al. 2007; Sherratt and Murray 1990; Swanson et al. 2003; Tremel et al. 2009), implies that the dynamics of the system is unaffected by spatial structure, such as the formation of clusters or patchiness within the population (Bolker and Pacala 1997; Law and Dieckmann 2000; Law et al. 2003; Simpson et al. 2010a). Here, we report a series of experiments describing the formation of a monolayer from an initially sparse and approximately uniform population of cells. Two different cell types are used and our experimental results illustrate that some cells, such as the 3T3 fibroblast population, forms a monolayer where the cells remain approximately spatially uniform as the population density increases. Alternatively, other cell types, such as the MDA MB 231 breast cancer cells, develop a monolayer which involves significant clustering.

We described a discrete model of cell proliferation and motility with volume exclusion effects and briefly recalled how the average properties of the discrete system can be described using a moment dynamics model. We also point out that our moment dynamics model reduces to the usual logistic growth model when correlations between the occupancy of lattice sites are neglected. Our results illustrate that care is

required when applying a mean-field model to experimental data. We show that it is possible to calibrate the standard mean-field logistic growth curve to both the 3T3 and 231 data: the simple calibration procedure provides us with an estimate of the proliferation rate; however, this procedure alone does not provide us with any opportunity for testing whether the mean-field assumption inherent in the logistic growth model is valid. To provide additional information, we also apply the moment dynamics model to the same data. Calibrating the moment dynamics model to both data sets showed that the mean-field growth rate and the moment dynamics growth rate for the 3T3 fibroblast population are almost identical. The similarity of the mean-field and moment dynamics estimates indicates that the mean-field model is a reasonable description of 3T3 monolayer formation. Alternatively, we found that the mean-field growth rate for the 231 cell population underestimated the moment dynamics growth rate by 25 %. The large difference between the mean-field and moment dynamics growth rates for the 231 cells indicates that the mean-field model might not be appropriate to describe the 231 monolayer formation. This is consistent with our visual interpretation of the 231 growth process which appears to involve significant clustering. To support our visual interpretation of the experimental images, we also used a statistical measure which confirms the absence of clusters in the 3T3 population and their presence in the 231 cell population.

The difference between the mean-field and moment dynamics growth rates for the 231 breast cancer cell population highlights two key points that can be easily overlooked when standard ode and pde models of collective cell behaviour are applied to experimental observations. First, our results indicate that standard mean-field models, such as the logistic equation, might not be the most appropriate model to apply to a given data set. This is of particular interest given that many experimental observations of collective cell behaviour indicate that cell populations often involve significant clustering (Hackett-Jones et al. 2012; Lah and Key 2012; Tamm et al. 1994). Second, although standard mean-field models can be inappropriate for describing processes involving clustering or patchy dynamics, these models can still be calibrated to patchy or clustered data sets to produce a good match to the observed data. Without any alternative modelling and interpretive techniques, such as the moment dynamics model and spatial statistics analysis presented here, it is difficult, if not impossible, to assess whether the mean-field assumption is reasonable.

Although recent theoretical work has shown how to make modelling predictions of collective cell behaviour without the standard mean-field assumption (Baker and Simpson 2010; Simpson and Baker 2011; Johnston et al. 2012), we are unaware of any previous study that has applied moment dynamics models to experimental observations of collective cell behaviour. Therefore, this study is a first attempt to demonstrate the possibility of using a moment dynamics model to interpret experimental observations with the aim of examining the validity of the mean-field assumption for applications in collective cell behaviour. Although we have chosen to begin our work with a relatively straightforward experimental system describing the development of a two-dimensional monolayer, our study indicates that certain refinements to the experimental system and modelling approach are possible and could lead to new insight into the role of the mean-field assumption. For example, future experimental investigations would benefit from using a fluorescent stain of the cell nucleus

(Suzuki et al. 1997) so we can identify the location of individual cells as the density increases toward confluence. Similarly, our moment dynamics model assumes that the motion of individual cells is unaffected by cell-to-cell adhesion and it would also useful to include experimental markers of cell-to-cell adhesion (Shiozaki et al. 1991; Takeichi 1991) so that we can assess whether this assumption is reasonable or whether is necessary to incorporate adhesion effects into the moment dynamics model (Johnston et al. 2012). Our work could also be extended by considering spatially-variable initial conditions, such as in a scratch assay (Maini et al. 2004a, 2004b) or a barrier assay (Simpson et al. 2013) where a spatially-variable moment dynamics model (Simpson and Baker 2011) could be applied. Finally, our mathematical modelling platform could be extended by taking a lattice-free approach (Codling et al. 2008). Very recent progress has been made towards developing ode and pde descriptions of collective cell behaviour with volume exclusion using a lattice-free framework and an effective mean–field assumption (Bruna and Chapman 2012; Dyson et al. 2012; Plank and Simpson 2012); however, we are unaware of any similar developments using a moment dynamics model of a lattice-free mechanism.

Acknowledgements We appreciate the advice and assistance of Associate Professor David Leavesley, Dr. Kerry Manton and Dr Leo de Boer. This research is supported by the Australian Research Council Discovery Project DP120100551, and the 2011 International Exchange Scheme funded by the Royal Society.

References

Baker, R. E., & Simpson, M. J. (2010). Correcting mean-field approximations for birth–death–movement processes. *Phys. Rev. E*, *82*, 041905.

Binder, B. J., & Landman, K. A. (2011). Quantifying evenly distributed states in exclusion and nonexclusion processes. *Phys. Rev. E*, *83*, 041914.

Binder, B. J., Hackett-Jones, E. J., Tuke, J., & Landman, K. A. (2011). A modified Pólya urn process and an index for spatial distributions with volume exclusion. *ANZIAM J.*, *53*, 122–133.

Binder, B. J., Landman, K. A., Newgreen, D. F., Simkin, J. E., Takahashi, Y., & Zhang, D. (2012). Spatial analysis of multi-species exclusion processes: application to neural crest cell migration in the embryonic gut. *Bull. Math. Biol.*, *74*, 474–490.

Bolker, B., & Pacala, S. W. (1997). Using moment equations to understand stochastically-driven spatial pattern formation in ecological systems. *Theor. Popul. Biol.*, *52*, 179–197.

Bruna, M., & Chapman, S. J. (2012). Excluded-volume effects in the diffusion of hard spheres. *Phys. Rev. E*, *85*, 011103.

Cai, A. Q., Landman, K. A., & Hughes, B. D. (2007). Multi-scale modeling of a wound-healing cell migration assay. *J. Theor. Biol.*, *245*, 576–594.

Cailleau, R., Young, R., Olivé, M., & Reeves WJ, Jr. (1974). Breast tumor cell lines from pleural effusions. *J. Natl. Cancer Inst.*, *53*, 661–674.

Codling, E. A., Plank, M. J., & Benhamou, S. (2008). Random walk models in biology. *J. R. Soc. Interface*, *5*, 813–834.

Dangerfield, C. E., Ross, J. V., & Keeling, M. J. (2008). Integrating stochasticity and network structure into an epidemic model. *J. R. Soc. Interface*, *6*, 761–774.

Deroulers, C., Aubert, M., Badoual, M., & Grammaticos, B. (2009). Modeling tumor cell migration: from microscopic to macroscopic models. *Phys. Rev. E*, *79*, 031917.

Dyson, L., Maini, P. K., & Baker, R. E. (2012). Macroscopic limits of individual-based models for motile cell populations with volume exclusion. *Phys. Rev. E*, *86*, 031903.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, *81*, 2340–2631.

Hackett-Jones, E. J., Landman, K. A., Newgreen, D. F., & Zhang, D. (2011). On the role of differential adhesion in gangliogenesis in the enteric nervous system. *J. Theor. Biol.*, *287*, 148–159.

Hackett-Jones, E. J., Davies, K. J., Binder, B. J., & Landman, K. A. (2012). Generalized index for spatial data sets as a measure of complete spatial randomness. *Phys. Rev. E*, *85*, 061908.

Hansen, J.-P., & McDonald, I. R. (2006). *Theory of simple liquids* (3rd edn.). London: Elsevier/Academic Press.

Hughes, B. D. (1995). *Random walks and random environments* (Vol. 1). Oxford: Oxford University Press.

Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*. Chichester: Wiley.

Johnston, S. T., Simpson, M. J., & Baker, R. E. (2012). Mean-field descriptions of collective migration with strong adhesion. *Phys. Rev. E*, *85*, 051922.

Keeling, M. J., Rand, D. A., & Morris, A. J. (1997). Correlation models for childhood epidemics. *Proc. R. Soc. Lond. B*, *264*, 1149–1156.

Lah, G. J., & Key, B. (2012). Novel roles of the chemorepellent axon guidance molecule RGMa in cell migration and adhesion. *Mol. Cell. Biol.*, *32*, 968–980.

Law, R., & Dieckmann, U. (2000). A dynamical system for neighbourhoods in plant communities. *Ecology*, *81*, 2137–2148.

Law, R., Murrell, D. J., & Dieckmann, U. (2003). Population growth in space and time: spatial logistic equations. *Ecology*, *84*, 252–262.

Law, R., Illian, J., Burslem, D. F. R. P., Gratzer, G., Gunatilleke, C. V. S., & Gunatilleke, A. U. N. (2009). Ecological information from spatial patterns of plants: insights from point process theory. *J. Ecol.*, *97*, 616–628.

Liggett, T. M. (1999). *Stochastic interacting systems: contact, voter and exclusion processes*. Berlin: Springer.

Mai, J., Kuzovkov, V. N., & von Niessen, W. (1993). A theoretical stochastic model for the $A + 1/2B_2 \rightarrow 0$ reaction. *J. Chem. Phys.*, *98*, 10017–10025.

Mai, J., Kuzovkov, V. N., & von Niessen, W. (1994). A general stochastic model for the description of surface reaction systems. *Physica A*, *203*, 298–315.

Maini, P. K., McElwain, D. L. S., & Leavesley, D. I. (2004a). Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. *Tissue Eng.*, *10*, 475–482.

Maini, P. K., McElwain, D. L. S., & Leavesley, D. (2004b). Travelling waves in a wound healing assay. *Appl. Math. Lett.*, *17*, 575–580.

Massey, S. C., Assanah, M. C., Lopez, K. A., Canoll, P., & Swanson, K. R. (2012). Glial progenitor cell recruitment drives aggressive glioma growth: mathematical and experimental modelling. *J. R. Soc. Interface*, *9*, 1757–1766.

Murray, J. D. (2002). *Mathematical biology I: an introduction* (3rd edn.). Heidelberg: Springer.

Murrell, D. J., Dieckmann, U., & Law, R. (2004). On moment closures for population dynamics in continuous space. *J. Theor. Biol.*, *229*, 421–432.

Penington, C. J., Hughes, B. D., & Landman, K. A. (2011). Building macroscale models from microscale probabilistic models: a general probabilistic approach for nonlinear diffusion and multispecies phenomena. *Phys. Rev. E*, *84*, 041120.

Penington, C. J., Korvasová, K., Hughes, B. D., & Landman, K. A. (2012). Collective motion of dimers. *Phys. Rev. E*, *86*, 051909.

Phelps, J. H., & Tucker, C. L. (2006). Lagrangian particle calculations of distributive mixing: limitations and applications. *Chem. Eng. Sci.*, *61*, 6826–6836.

Plank, M. J., & Simpson, M. J. (2012). Models of collective cell behaviour with crowding effects: comparing lattice-based and lattice-free approaches. *J. R. Soc. Interface*, *9*, 2983–2996.

Raghib, M., Hill, N. A., & Dieckmann, U. (2011). A multiscale maximum entropy moment closure for locally regulated space—time point process models of population dynamics. *J. Math. Biol.*, *62*, 605–653.

Sengers, B. G., Please, C. P., & Oreffo, R. OC. (2007). Experimental characterization and computational modelling of two-dimensional cell spreading for skeletal regeneration. *J. R. Soc. Interface*, *4*, 1107–1117.

Sharkey, K. J. (2008). Deterministic epidemiological models at the individual level. *J. Math. Biol.*, *57*, 311–331.

Sharkey, K. J. (2011). Deterministic epidemic models on contact networks: correlations and unbiological terms. *Theor. Popul. Biol.*, *79*, 115–129.

Sherratt, J. A., & Murray, J. D. (1990). Models of epidermal wound healing. *Proc. R. Soc. Lond. B*, *241*, 29–36.

Shiozaki, H., Tahara, H., Oka, H., Miyata, M., Kobayashi, K., Tamura, S., Iihara, K., Doki, Y., Hirano, S., Takeichi, M., & Mori, T. (1991). Expression of immunoreactive E-Cadherin adhesion molecules in human cancers. *Am. J. Pathol.*, *139*, 17–23.

Simpson, M. J., & Baker, R. E. (2011). Corrected mean-field models for spatially-dependent advection–diffusion–reaction phenomena. *Phys. Rev. E*, *83*, 051922.

Simpson, M. J., Landman, K. A., Hughes, B. D., & Fernando, A. E. (2010a). A model for mesoscale patterns in motile populations. *Physica A*, *389*, 1412–1424.

Simpson, M. J., Landman, K. A., & Hughes, B. D. (2010b). Cell invasion with proliferation mechanisms motivated by time-lapse data. *Physica A*, *389*, 3779–3790.

Simpson, M. J., Baker, R. E., & McCue, S. W. (2011). Models of collective cell spreading with variable cell aspect ratio: a motivation for degenerate diffusion models. *Phys. Rev. E*, *83*, 021901.

Simpson, M. J., Treloar, K. K., Binder, B. J., Haridas, P., Manton, K., Leavesley, D. I., McElwain, D. L. S., & Baker, R. E. (2013). Quantifying the roles of cell motility and cell proliferation in a circular barrier assay. *J. R. Soc. Interface*, *10*, 20130007.

Singer, A. (2004). Maximum entropy formulation of the Kirkwood superposition approximation. *J. Chem. Phys.*, *121*, 3657.

Suzuki, T., Fujikura, K., Higashiyama, T., & Takata, K. (1997). DNA staining for fluorescence and laser confocal microscopy. *J. Histochem. Cytochem.*, *45*, 49–53.

Swanson, K. R., Bridge, C., Murray, J. D., & Alvord, E. C. Jr. (2003). Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J. Neurol. Sci.*, *216*, 1–10.

Takeichi, M. (1991). Cadherin cell adhesion receptors as a morphogenetic regulator. *Science*, *251*, 1451–1555.

Tamm, I., Kikuchi, Y., Cardinale, I., & Krueger, J. G. (1994). Cell-adhesion-disrupting action of interleukin 6 in human ductal breast carcinoma cells. *Proc. Natl. Acad. Sci. USA*, *91*, 3329–3333.

Todaro, G. J., & Green, H. (1963). Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J. Cell Biol.*, *17*, 299–313.

Tremel, A., Cai, A. Q., Tirtaatmadja, N., Hughes, B. D., Stevens, G. W., Landman, K. A., & O'Connor, A. J. (2009). Cell migration and proliferation during monolayer formation and wound healing. *Chem. Eng. Sci.*, *64*, 247–253.

Young, W. R., Roberts, A. J., & Stuhne, G. (2001). Reproductive pair correlations and the clustering of organisms. *Nature*, *412*, 328–331.