



Research



Cite this article: Martinson WD, Volkening A, Schmidtchen M, Venkataraman C, Carrillo JA.

2024 Linking discrete and continuous models of cell birth and migration. *R. Soc. Open Sci.* **11**: 232002.

<https://doi.org/10.1098/rsos.232002>

Received: 22 December 2023

Accepted: 10 May 2024

Subject Category:

Mathematics

Subject Areas:

applied mathematics/mathematical modelling/
theoretical biology

Keywords:

non-local interactions, zebrafish, self-organization, aggregation equations

Author for correspondence:

José A. Carrillo

e-mail: carrillo@maths.ox.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7303179>.

Linking discrete and continuous models of cell birth and migration

W. Duncan Martinson¹, Alexandria Volkening²,
Markus Schmidtchen³, Chandrasekhar Venkataraman⁴
and José A. Carrillo¹

¹Mathematical Institute, University of Oxford, Oxford, UK

²Department of Mathematics, Purdue University, West Lafayette, IN, USA

³Institute of Scientific Computing, Technische Universität Dresden, Dresden, Germany

⁴Department of Mathematics, University of Sussex, Brighton, UK

WDM, 0000-0002-3590-606X; AV, 0000-0003-3401-5094;
MS, 0000-0001-9662-3111; CV, 0000-0003-4039-2056;
JAC, 0000-0001-8819-4660

Self-organization of individuals within large collectives occurs throughout biology. Mathematical models can help elucidate the individual-level mechanisms behind these dynamics, but analytical tractability often comes at the cost of biological intuition. Discrete models provide straightforward interpretations by tracking each individual yet can be computationally expensive. Alternatively, continuous models supply a large-scale perspective by representing the ‘effective’ dynamics of infinite agents, but their results are often difficult to translate into experimentally relevant insights. We address this challenge by quantitatively linking spatio-temporal dynamics of continuous models and individual-based data in settings with biologically realistic, time-varying cell numbers. Specifically, we introduce and fit scaling parameters in continuous models to account for discrepancies that can arise from low cell numbers and localized interactions. We illustrate our approach on an example motivated by zebrafish-skin pattern formation, in which we create a continuous framework describing the movement and proliferation of a single cell population by upscaling rules from a discrete model. Our resulting continuous models accurately depict ensemble average agent-based solutions when migration or proliferation act alone. Interestingly, the same parameters are not optimal when both processes act simultaneously, highlighting a rich difference in how combining migration and proliferation affects discrete and continuous dynamics.

1. Introduction

Self-organization of individual agents is a key feature of life. It occurs ubiquitously throughout the natural world, from the macroscopic example of bird flocking [1–4] to the microscopic phenomenon of cell sorting during development [5–9]. The degree to which members of a group coordinate their movement, proliferation and competition accounts for pattern diversity across biological scales. Alongside experimental approaches, mathematical models can help identify the underlying behaviours that give rise to specific collective dynamics. However, a trade-off often exists between tractability and detail when building models of pattern formation, due in part to the multiscale nature of biological systems. Consequently, better quantitative characterization of the relationship between analytically tractable models and more biologically representative approaches will improve our understanding of self-organization throughout nature.

Here, we help address this open challenge using pigment cell dynamics in zebrafish patterns as a motivation. The zebrafish (*Danio rerio*) is a popular model organism for studying pattern formation, as dark stripes and gold interstripes emerge in its skin during development [10–13]. As we show in figure 1, these stripes result from the coordination of interactions among several types of cells, including black melanophores and gold (dense) xanthophores [10,16–20]. Experiments that perturb stripes—e.g. by laser ablation [16,21]—demonstrate how cell–cell signalling and external cues contribute to the creation of alternative motifs such as spots or labyrinths. A rich diversity of mutant patterns, including widened or curvy stripes, also emerge when cell interactions are altered due to genetic mutations [19,22].

Data-driven mathematical models can help uncover the drivers of zebrafish pattern formation and other biological phenomena exhibiting self-organization by identifying important phase transitions, isolating the effects of specific processes such as cell division, and providing hypotheses that can guide the design of *in vivo* experiments [23–26]. Different modelling frameworks yield insight at the population or individual level, depending on how they represent members of a group. One modelling approach involves tracking how the position of each individual changes in time. These so-called ‘discrete’ systems include centre-based models [27,28], cellular automata [29,30], cellular Potts models [31,32] and vertex models [33,34]. Within the setting of zebrafish patterning, agent-based models (ABMs) have been developed that restrict cells to occupy certain locations ‘on-lattice’ [35–37] or allow them to roam freely, ‘off-lattice’, in the domain [14,38–40]. Due to their ability to work on the same length scales as empirical data, ABMs provide an intuitive connection to experiments and allow for detailed predictions about how interactions between agents drive group behaviours. However, ABMs can be prohibitive to simulate when the number of individuals is large, and understanding their long-time behaviour under alternative rules and parameters relies on extensive computation [41].

A second modelling approach uses continuous functions to represent the ‘average’ density of agents in a collective, with their dynamics governed by a partial differential equation (PDE) in space and time. Continuous models, including reaction–diffusion equations, Boltzmann-like kinetic equations and integro-differential equations (IDEs), typically cannot resolve individuals and, instead, track the ensemble average (EA) behaviour of a population. However, these models are more amenable to mathematical analysis and more readily provide insight into long-term behaviour than discrete frameworks do [42,43]. For example, changes in patterning may arise because of Turing-like instabilities [44–46] or due to alterations in physically based interactions such as cell–cell adhesion [7,38,47–50]. In the case of zebrafish patterns, researchers have applied a wide swath of continuous models—including reaction–diffusion equations [16,21,35,51,52] and non-local PDEs [7,53–55]—to better understand cell dynamics.

Despite the differences between discrete and continuous approaches, it is possible to establish a mathematical link between these representations in the limit of infinite individuals. This procedure, known as ‘coarse-graining’, derives differential equations from a given discrete model and yields information about its EA behaviour [56–62]. For example, the authors in [63–65] derive logistic IDEs from stochastic processes that describe the birth and death of individuals undergoing Darwinian evolution in the limit of large numbers. Coarse-grained descriptions become inaccurate when relatively few individuals are present, however, as is the case in many biological contexts such as pattern formation in zebrafish. Many approaches also neglect potentially important spatial correlations between cells—caused, for instance, by division or competition—that may play a critical role in pattern dynamics [52,66–71]. While it is possible to go beyond this ‘mean-field’ setting by deriving continuous models that respect higher-order correlations [72–76] and hard-core interactions [77], these methods still rely on simplifying assumptions that introduce errors between the discrete and continuous frameworks. Controlling these errors in a biological setting is an important objective.

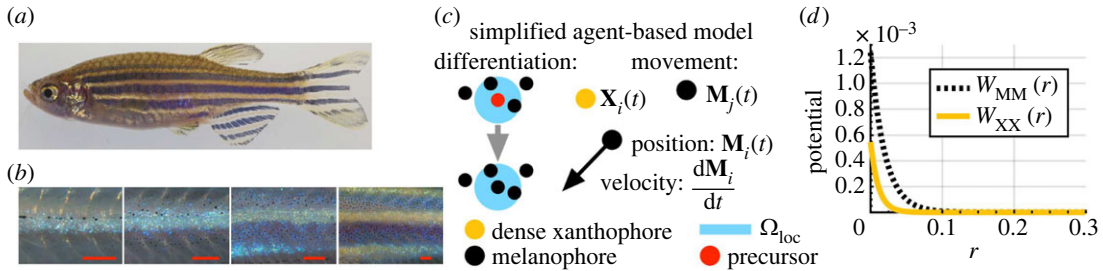


Figure 1. Motivating biological example and model. (a) Wild-type zebrafish feature stripe patterns in their skin. These patterns consist of several types of brightly coloured pigment cells. (b) Over the course of a few months, these cells organize sequentially into stripes and interstripes from the centre of the fish body outward [10]. (c) For the purposes of this paper, we focus on a single population of black melanophores or gold dense xanthophores, using a simplified version of the model from [14]. The agent-based model (ABM) [14] that motivates our work describes how patterns arise through cell differentiation, competition and movement. In our simplified version of the ABM [14], we assume new cells appear at randomly selected locations based on short-range activation; this models cell differentiation from uniformly distributed precursors (red position) [14], and we also refer to this as ‘proliferation’ or ‘birth’ in this paper. (We describe the cell differentiation rules in the full model [14] in more detail in electronic supplementary material, figure S1.) In both our work and the ABM from [14], cell movement is deterministic and governed by ordinary differential equations (ODEs). (d) These ODEs account for cell–cell repulsion through potential functions, which describe melanophore–melanophore (W_{MM}) and xanthophore–xanthophore (W_{XX}) interactions as a function of their pairwise distance r . Red scale bar is $250\ \mu\text{m}$ in (b). Image (a) adapted from Fadeev *et al.* [15] and licensed under CC-BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>). Image (b) adapted from Frohnhofer *et al.* [10] and licensed under CC-BY 3.0 (<https://creativecommons.org/licenses/by/3.0/>); published by The Company of Biologists Ltd.

We help tackle this problem by developing a pipeline to minimize spatio-temporal discrepancies between continuous models and individual-based data in settings with biologically relevant, dynamic cell numbers. The method relies on fitting parameters that effectively dilate the time variable of continuous-model solutions. Our approach can be used to describe biological self-organization in systems whose macroscopic description can be derived or inferred. We apply our approach to a case study motivated by stripe formation in zebrafish skin, as this allows us to illustrate its utility and interpretability for experimentally measurable quantities with biologically meaningful spatial and temporal units. In §2, we describe the ABM that we simulate to generate synthetic individual-based data associated with self-organizing phenomena. In its full form, the ABM motivating our work [14] admits pattern formation via non-local rules for cell birth, death and movement that are inspired by the underlying biology of zebrafish-skin patterns (figure 1) [14,40]. To focus on the presentation of our pipeline, however, we simplify some biological complexity by reducing this model [14] to focus on a single cell type—melanophores or dense xanthophores. (We plan to extend this pipeline to multiple cell types in future work.) We then detail the corresponding continuous descriptions and our method for matching their solutions to EA ABM data even in scenarios with finite and time-dynamic cell numbers. We present our results for black melanophores in §3, and—as a means of demonstrating the generality of our methodology—apply the same approach to dense xanthophores in electronic supplementary material, S6.

2. Mathematical models and methods

In §2.1, we develop our ABM for cell migration and derive its continuous counterpart. Subsequently, in §2.2, we introduce our discrete model for cell birth and develop a corresponding continuous IDE model. We present our full models of migration and proliferation in §2.3. Lastly, we present our approach to estimating scaling parameters in our continuous models from EA ABM data in §2.4. Following previous ABMs [14,40] of pattern formation in zebrafish, we assume (i) migration is governed by conservative forces between pigment cells and (ii) non-local interactions inform cell birth in a two-dimensional (2D) plane [14,39,40]. Throughout this paper, we refer to:

- $\Omega \subset \mathbb{R}^2$ = domain of the simulation with spatial units of millimetres (mm),
- $\mathbb{R}^2 \ni \mathbf{M}_i(t)$ = coordinates of the centre of the i th melanophore at t days in our discrete models,
- $\mathbb{N} \ni N_M(t)$ = total number of melanophores present at time t days and
- $\mathbb{R}_{\geq 0} \ni M(\mathbf{x}, t)$ = density of melanophores at position \mathbf{x} and time t in cells mm^{-2} ,

with the exception of [figure 7](#) where we consider a one-dimensional (1D) domain; there $M(\mathbf{x}, t)$ is the number density of melanophores in cells mm^{-1} . Because it appears several times, we define the indicator function $1_{\{\text{condition}\}}$ here as

$$1_{\{\text{condition}\}}(x) = \begin{cases} 1, & \text{if } x \text{ satisfies the rule specified by 'condition',} \\ 0, & \text{otherwise,} \end{cases} \quad (2.1)$$

where ‘condition’ depends on the model rule and cell interaction, as we discuss next.

2.1. Models of migration

Our ABM for cell movement tracks the position, $\mathbf{M}_i(t)$, of each cell, indexed by $i \in \{1, \dots, N_M(t)\}$, at time $t \geq 0$. The movement of each melanophore depends on its interactions with surrounding melanophores, following an overdamped version of Newton’s second law. The forces are assumed to be conservative, i.e. they may be written as the gradient of a potential. This leads to the following system:

$$\frac{d\mathbf{M}_i}{dt} \approx \mathbf{F}_{\text{int}}^{(i)} = - \sum_{j=1, j \neq i}^{N_M(t)} \nabla W_{\text{MM}}^c(\mathbf{M}_i - \mathbf{M}_j). \quad (2.2)$$

Here, $\mathbf{F}_{\text{int}}^{(i)}$ is the net force arising from all cell–cell interactions according to the potential $W_{\text{MM}}(\mathbf{r})$, where \mathbf{r} is the inter-particle displacement. The potential can encode cell–cell repulsion and adhesion, depending on the sign of its gradient along the direction between two cell centres. Many choices for $W_{\text{MM}}(\mathbf{r})$ are possible, including harmonic, power-law, Morse and Lennard-Jones potentials, among others. Here, we use exponential potentials given by

$$W_{\text{MM}}(\mathbf{r}) = R_{\text{MM}} e^{-|\mathbf{r}|/\omega_{\text{MM}}} - A_{\text{MM}} e^{-|\mathbf{r}|/d_{\text{MM}}}, \quad (2.3)$$

in accordance with prior ABMs of pattern formation in zebrafish [[14,40](#)]; see [figure 1c](#) and [table 1](#) for parameter values and their biological interpretations. To model cells communicating through cellular extensions or dendrites [[19,20](#)], secreted signals [[80](#)] or cell–cell contact [[81](#)], we assume forces on cells are zero beyond some cut-off distance d_{max} . We represent this using the notation $\nabla W_{\text{MM}}^c(\mathbf{r}) = \nabla W_{\text{MM}}(\mathbf{r}) 1_{\{|\mathbf{r}| < d_{\text{max}}\}}(\mathbf{r})$, and set $N_M(t) = N_M$ when there is no cell birth. We remark that the model given by equation (2.2) is deterministic in the sense that it does not include Brownian motion.

The associated continuous model describes the melanophore density, $M(\mathbf{x}, t)$. Integrating $M(\mathbf{x}, t)$ over a bounded region yields the total number of melanophores within that area at time t . Following the coarse-graining procedure in [[60,82,83](#)], an outline of which is found in electronic supplementary material, §1.1, we obtain the PDE below:

$$\frac{\partial M}{\partial t} = \alpha_{\text{MM}} \nabla \cdot (M \nabla W_{\text{MM}}^c \star M), \quad (2.4)$$

where the force ∇W_{MM}^c is the same as in equation (2.2) and \star is the convolution operator [[14,40](#)]. The parameter α_{MM} in equation (2.4) is not inherent to the coarse-graining procedure; rather, it accounts for possible differences between the discrete and continuous models. Indeed, simulating equation (2.4) with the parameter values listed in [table 1](#) and $\alpha_{\text{MM}} = 1$ (the value expected from the mean-field approximation) does not always capture the ABM dynamics; see [figure 2](#).

The individual and EA ABM results demonstrate that cells disperse until they are about 55–115 μm apart at $t = 150$ days. The PDE with $\alpha_{\text{MM}} = 1$, however, predicts that cells travel about 250 μm further in the same time period. Additionally, the PDE cell density is lower than the EA ABM density near the centre of the domain, implying that cells are more separated there. The continuous solution at earlier times more closely resembles the EA ABM result at $t = 150$ days, however, which suggests that PDE solutions evolve at a faster time scale than that of the discrete model. The parameter α_{MM} effectively dilates the time variable, such that solutions travel α_{MM} times more quickly. Thus, a non-unitary value of α_{MM} is likely to produce a better match between the discrete and continuous solutions. To our knowledge, the value of α_{MM} cannot be derived *a priori*. Instead, we develop an approach for estimating its value based on ABM data in §2.4, and will pursue an analytic derivation for this parameter in future work.

Table 1. Model and simulation parameters used throughout the paper. We note that $N_{\text{bir}} \in \{10, 25, 50, 100, 150, 200, 250\}$ realizations for 2D simulations and $N_{\text{bir}} \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ realizations for 1D simulations. As we discuss in the electronic supplementary material, S2, $N_{\text{sim}} = 5 \times 10^3$ realizations for most EA ABM solutions of the cell movement model and $N_{\text{sim}} = 10^3$ for all EA ABM results of the cell birth and combined models. We set $N_{\text{hist}} = 30$ voxels for EA ABM solutions of the cell birth and combined models, and $N_{\text{hist}} = N_{\text{bin}}$ for the cell movement model. (See figure captions for our N_{bin} values.) The values of R_{MM} (and R_{XX} , see electronic supplementary material, S1.3) were reported as repulsion strengths (i.e. $R_{\text{MM}}/\omega_{\text{MM}}$) in [39].

parameter	value	description and motivation
R_{MM}	$0.00124 \text{ mm}^2 \text{ day}^{-1}$	strength of melanophore repulsion potential in equation (2.3); based on [14,39]
A_{MM}	$0 \text{ mm}^2 \text{ day}^{-1}$	strength of melanophore adhesion potential in equation (2.3); based on [14,39]
ω_{MM}	0.02 mm	melanophore repulsion interaction range in equation (2.3); based on [14,39]
a_{MM}	0.012 mm	melanophore adhesion interaction range in equation (2.3); based on [14,39]
d_{max}	0.2 mm	maximum cell interaction distance in equation (2.2); based on [14,39]
d_{loc}	0.075 mm	maximum interaction range for cell birth in equation (2.6); based on [14] and chosen slightly larger than measurements of cell–cell distances [78,79]
N_{bir}	varies	number of positions selected uniformly at random per day for possible cell proliferation (e.g. differentiation from precursors) in equations (2.5) and (2.7)
c^-	1 cell	lower bound for the number of cells in a short-range neighbourhood for cell proliferation in equations (2.5) and (2.7)
c^+	6 cells	upper bound for the number of cells in a short-range neighbourhood for birth in equations (2.5) and (2.7); based on estimations of data [16,78] in [14]
t_{final}	150 or 2000 days	simulation end time (150 days in 2D and 2000 days in 1D)
Δt_{move}	0.01 or 0.1 days	time step for numerical implementation of equation (2.2) and electronic supplementary material, equation (S3)
Δt_{bir}	1 day	time step for numerical implementation of cell birth in equations (2.5) and (2.7)
Δt_{PDE}	0.05 days	time step for numerical implementation of equations (2.4), (2.8) and electronic supplementary material, equation (S6)
Δt_{record}	1 day	time step for recording data from model simulations
N_{sim}	varies	number of ABM realizations for computing EA cell densities
N_{bin}	varies	spatial discretization step for solving our continuous models
N_{hist}	N_{bin} or 30 voxels	spatial discretization step for binning simulation results for comparison

2.2. Models of cell birth

Our ABM for cell birth consists of stochastic, discrete-time rules which we adapt and simplify from [14] (motivation for these rules may be found in electronic supplementary material, S1.2). Specifically, at each time step (i.e. day) in a simulation, we select $N_{\text{bir}} \in \mathbb{N}$ locations uniformly at random from Ω and evaluate them synchronously for possible cell birth. Each selected location, \mathbf{z} , represents the position of a precursor cell that may differentiate into a melanophore based on the signals that it receives. The conditions for melanophore birth in the ABM [14] depend on both neighbouring melanophores and dense xanthophores, as we show in electronic supplementary material, figure S1. Since we restrict to one population in this paper, we simplify the rules from [14]; see electronic supplementary material, S1.2, for details. In particular, a new melanophore emerges at position \mathbf{z} according to the rule

$$\left. \begin{array}{l} \underbrace{\sum_{i=1}^{N_M} \mathbf{1}_{\{\mathbf{M}_i \in \Omega_{\text{loc}}^z\}}(\mathbf{M}_i) \geq 1}_{\text{short-range activation}} \quad \text{and} \\ \underbrace{\sum_{i=1}^{N_M} \mathbf{1}_{\{\mathbf{M}_i \in \Omega_{\text{loc}}^z\}}(\mathbf{M}_i) < c_{\text{ABM}}^+}_{\text{overcrowding prevention}} \rightarrow \text{melanophore appears at } \mathbf{z}, \end{array} \right\} \quad (2.5)$$

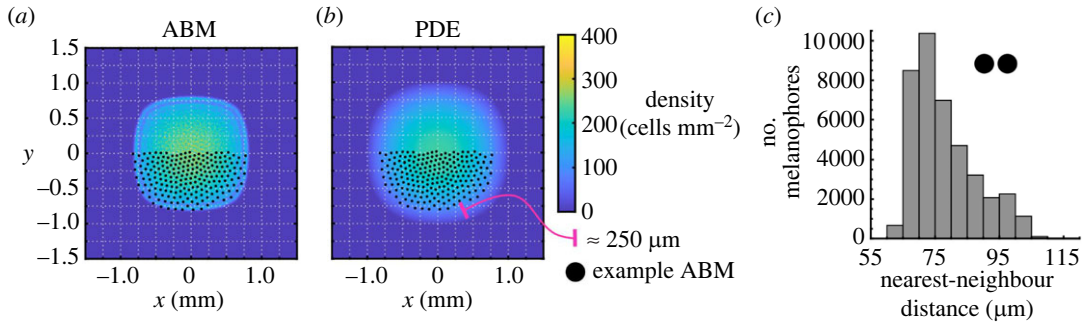


Figure 2. The PDE for cell migration does not accurately describe the EA ABM result when its scaling parameter, α_{MM} , is set to unity. (a) To compute our EA ABM result, we solve equation (2.2) using an initial condition of 400 melanophores placed uniformly at random in a $1 \times 1 \text{ mm}^2$ square, group cell positions in a 240×240 histogram, and average such data over 10^4 ABM realizations. (b) We compute the corresponding PDE solution by simulating equation (2.4) with $\alpha_{MM} = 1$ from a uniform density of 400 cells mm^{-2} in the same square region. The ABM and PDE solutions use the same potential (given by equation (2.3) with parameters in table 1). We overlay an example ABM realization for comparison; the results demonstrate that the support of the PDE is larger than that of the ABM by about 200–250 μm . Because melanophore–melanophore distances have been measured to be roughly 50 μm *in vivo* [78] and stripes are only about 7–12 cells wide [16], this is a large difference. (c) The distribution of nearest-neighbour distances across 100 ABM realizations demonstrates that cell–cell separation ranges from roughly 60 to 100 μm . Based on visual inspection of the graphs, nearest-neighbour distances appear inversely proportional to the EA cell density. In (a–c), we show results at $t = 150$ days.

where $c_{ABM}^+ \in \mathbb{N}$, and

$$\Omega_{loc}^z = \text{disc centred at } z \text{ with radius } d_{loc}. \quad (2.6)$$

According to equation (2.5), new cells appear near existing melanophores until the maximum number of cells—namely, c_{ABM}^+ —in Ω_{loc}^z , the interaction region between cells, is reached; see table 1 for parameter values. While equation (2.5) is deterministic, stochasticity enters our ABM through our N_{bir} randomly selected positions $\{z\}$. Similar stochastic rules can also be used to model cell death, as in [14,40], although we do not consider them here.

We do not know of existing methods for rigorously deriving continuous models of cell birth from off-lattice ABMs with this noise structure. We therefore adopt a phenomenological modelling approach, in which we create a continuous model whose governing equations mimic the stochastic interaction rules. We reason that the number density of cells in this setting must increase at a constant rate (proportional to N_{bir}) when continuous versions of the overcrowding and short-range activation restrictions are met, since this occurs at the individual level in the ABM. Furthermore, we represent the density restrictions with an indicator function using the integral of the number density over Ω_{loc} as an argument, since the latter quantity yields the total number of cells within that region. This leads to the following model:

$$\frac{\partial M}{\partial t}(\mathbf{x}, t) = \gamma N_{bir} 1 \left\{ 1 \leq \int_{\Omega_{loc}^{\mathbf{x}}} M(\mathbf{y}, t) d\mathbf{y} < c^+ \right\}(\mathbf{x}, t), \quad (2.7)$$

where c^+ is the continuous equivalent of the density-limiting parameter c_{ABM}^+ in equation (2.5); N_{bir} has the same value as in our corresponding ABM; and $\gamma \in \mathbb{R}^+$ is a parameter that effectively dilates the time variable in a similar way as α_{MM} in the cell movement model. The units of γ must be inversely proportional to those of the domain size in order to make the dimensions of equation (2.7) consistent. Its value is unknown; however, we can employ a phenomenological argument to determine an expected value by integrating equation (2.7) over the whole domain. This yields an upper bound on the number of cells born per unit time of $\gamma N_{bir} |\Omega|$, hence one expects $\gamma \approx |\Omega|^{-1}$ to maintain a maximum rate of N_{bir} cells born per day as in our ABM. We note, however, that this argument does not take into account possible clustering or other spatial correlations that can occur in the discrete setting, which may change the values of γ and c^+ from their expected values. While we could address this by allowing both parameters to depend on the proportion of the domain in which the birth conditions are fulfilled, we leave this extension for a future study and simply estimate uniform values for γ and c^+ by fitting to EA ABM data, as we do for α_{MM} in the movement-only model. We overview our approach for estimating the values of c^+ and γ in §2.4.

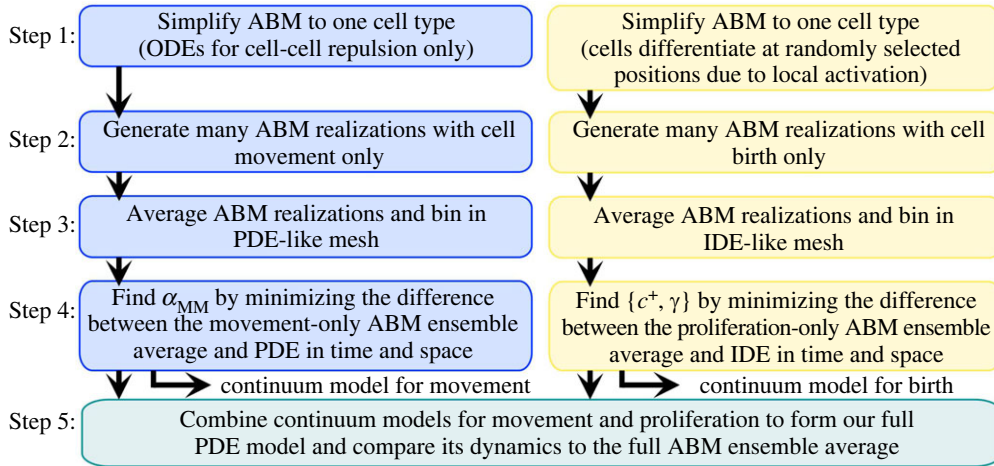


Figure 3. Our modular pipeline for matching the solutions of continuous and discrete models and identifying how cell movement and birth interact in both settings. We first isolate the discrete-model terms from [14] corresponding to movement (left column) and birth (right column) and simplify them to consider only one cell type. We then produce multiple realizations of our ABMs, sorting the cell locations into a grid of $N_{\text{hist}} \times N_{\text{hist}}$ voxels to yield the EA discrete-model results. We simulate our continuous model for cell movement (respectively, cell birth) and compare it on the same spatial mesh, with values of α_{MM} (respectively, c^+ and γ) obtained from a least-squares optimization approach; see electronic supplementary material, S2, for details. Finally, we combine the fitted movement and birth models to produce our full continuous model. While an extension of this pipeline to fit all three parameters simultaneously is straightforward, fitting separately allows us to better understand the effects of cell movement and birth in discrete and continuous frameworks.

2.3. Full models of cell movement and birth

We combine our descriptions of cell movement and proliferation to form our full discrete and continuous models. For our full ABM, we move cells according to equation (2.2) and then introduce new agents based on equation (2.5) at each simulated day; see electronic supplementary material, S2, for details. For our continuous model, we combine the terms related to movement and birth, such that the cell density evolves according to

$$\frac{\partial M}{\partial t}(\mathbf{x}, t) = \alpha_{\text{MM}} \nabla \cdot (M \nabla W_{\text{MM}}^c \star M) + \gamma N_{\text{bir}} 1_{\left\{1 \leq \int_{\Omega_{\text{loc}}} M(\mathbf{y}, t) d\mathbf{y} < c^+\right\}}(\mathbf{x}, t), \quad (2.8)$$

where the parameters α_{MM} , γ , N_{bir} and c^+ have the same interpretations as in §§2.1 and 2.2. Importantly, by assuming that these parameters have the same interpretations, we are assuming that migration and proliferation are additive, so that combining them has no extra influence. Our fitting approach for these parameters, discussed below, allows us to evaluate this choice and better understand the interplay of these two mechanisms in discrete and continuous settings.

2.4. Parameter estimation procedure

We identify the values of the parameters— α_{MM} in equation (2.4) and $\{\gamma, c^+\}$ in equation (2.7)—by minimizing the sum of squared differences (hereafter referred to as the ‘ L^2 error’) between the continuous and EA discrete solutions over time and space. Because we are able to model the isolated processes of cell birth and movement separately, or consider them acting simultaneously, there are two ways of estimating parameters: by fitting all three parameters simultaneously to data from the combined model, or by fitting them in a modular fashion by considering cell movement and birth in isolation from each other. For the remainder of this paper, we adopt a modular approach because it allows us to probe the particular effects of cell movement and birth in detail (see figure 3 for an overview), and we present a study of simultaneous estimation in electronic supplementary material, S3. In particular, by using our modular parameter values in the combined PDE model, we can investigate their interplay and better understand the additive effects of individual-level mechanisms on the accuracy of continuous models. As we discuss in electronic supplementary material, S3, our modular approach may also supply additional information that can improve parameter estimation.

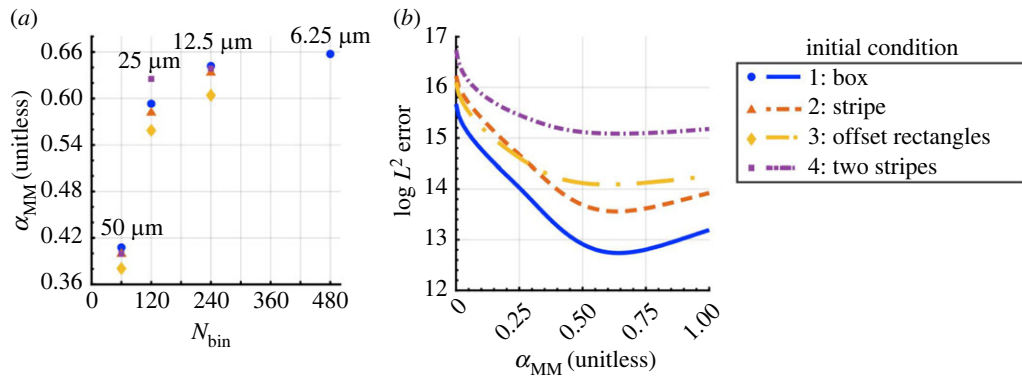


Figure 4. The optimal PDE scaling parameter for movement depends on the mesh resolution but appears to converge. (a) A scatter plot of the numerically optimized value of α_{MM} from equation (2.4) as a function of the mesh resolution, N_{bin} , demonstrates that this scaling parameter is correlated with the mesh resolution, but appears to converge at sufficiently high (i.e. $N_{bin} \geq 240$) detail. (We omit 95% confidence intervals because these are so narrow that they are difficult to see.) (b) Plotting the $\log L^2$ error, given by equation (2.9), as a function of α_{MM} with $N_{bin} = 240$ for each initial condition suggests that the values in (a) are optimal. As we show in figure 2c, melanophores are typically separated by more than $60 \mu\text{m}$ in our movement-only ABM simulations—over four times the voxel width in a grid with $N_{bin} = 240$. See §2.4 and electronic supplementary material, S2, for numerical details.

For example, we show in electronic supplementary material, S3, that α_{MM} and γ are not uniquely identifiable if they are fit only to the combined EA ABM data, whereas figure 4b and figure 7f suggest we can uniquely identify them with a modular approach. We overview our method for parameter estimation below; for parameter values, see table 1. We refer to electronic supplementary material, S2–S3, for further details about our implementation of the pipeline in addition to alternative choices that could be taken in parameter estimation (such as fitting to earlier times, less refined spatial data etc.).

We consider biologically meaningful time scales (i.e. days), length scales (i.e. mm), and cell densities and stress empirical units throughout our results. This choice supports future studies that may treat pattern formation with multiple cell types. Throughout our simulations, we consider a domain of size $3 \times 3 \text{ mm}^2$ (with one 1D exception in figure 7). We implement four initial conditions to extract common features of cell interactions from different geometric scenarios. The first involves a square region of melanophores in the centre of the domain (*Box*), which mimics the symmetry of the domain. For the second initial condition, we place a single stripe of melanophores (*Stripe*), which is motivated by the typical patterning observed in wild-type zebrafish. For the third initial condition, we consider two rectangular regions of melanophores (*Offset rectangles*), which take into account non-standard geometries and the meeting of two disjoint melanophore populations. Finally, the last initial condition we consider involves two melanophore stripes (*Two stripes*), which explores the interactions between two disjoint melanophore populations with biologically realistic sharp fronts. (See electronic supplementary material, figure S6, for a summary of these initial conditions.) We initialize individual ABM simulations by sampling cell positions uniformly in these regions for each respective initial condition, and initialize our continuous models by setting the cell density uniformly equal to the estimated biological density of $400 \text{ cells mm}^{-2}$ [14].

In Step 2 of our pipeline in figure 3, we solve our discrete models with an explicit approach. Specifically, we solve equation (2.2) with an explicit forward Euler scheme. To model differentiation from uniformly distributed precursor cells [14], we solve the birth-only ABM by selecting N_{bir} sites in the domain uniformly at random at a fixed time step (here, 1 day) and placing a new cell at each position that meets the conditions given by equation (2.5). (Following the approach in [14,39,40], we evaluate all N_{bir} locations for potential cell proliferation at the same time. This synchronous evaluation means that it is possible, though uncommon, for more than c_{ABM}^+ cells to be present in a local neighbourhood, and the choice of parameters in our model, based on the ABMs [14,39,40], accounts for this possibility.) We solve our ABM combining migration and birth by simulating equation (2.2) and then implementing cell birth as above, with migration evaluated using a shorter time step than the time step for birth events.

To compare ABM results directly with the cell density from our continuous models, we obtain an EA distribution by simulating many ABM realizations (here, between 10^3 and 10^4 simulations), sorting all the cell locations into a histogram of $N_{hist} \times N_{hist}$ voxels (or $N_{hist} \times 1$ voxels in 1D), and normalizing by the

number of simulations and the voxel area for each day simulated; see Step 3 of our pipeline in figure 3. Other ways of relating ABM and PDE results are also possible, for example by introducing Gaussian kernels at each cell location [84]. However, as far as the construction of the histogram is concerned, we expect comparable results for particles and localized Gaussian kernels. Furthermore, we show in electronic supplementary material, S3, that histogram voxel size used to bin EA ABM data and the final time used to fit the continuous equations only play a minor role in affecting the parameter values we obtain, at least for the cell movement model.

As part of Step 4 of our pipeline in figure 3, we need to solve our continuous models, and we do so with explicit approaches. Specifically, we apply a first-order finite volume scheme for the migration model (equation (2.4)), a forward Euler method for the continuous cell birth model (equation (2.7)), and a combined finite volume/forward Euler scheme for the full continuous framework (equation (2.8)). (More details about the particular time steps used to simulate the discrete and continuous models are in electronic supplementary material, S2.) We simulate the continuous models on an $N_{\text{bin}} \times N_{\text{bin}}$ mesh and, to match with EA ABM solutions, record the average cell density at each day on a (possibly coarser) grid of $N_{\text{hist}} \times N_{\text{hist}}$ voxels.

We compute continuous model parameters by minimizing the L^2 error between the continuous and EA ABM results across time. Notably, this nonlinear least-squares problem is equivalent to maximum likelihood parameter estimation when the densities produced from the ABM simulations are independent, identically distributed normal random variables with constant variance and mean equal to the continuous solution. The L^2 error that we minimize is given by

$$\begin{aligned} e_{L^2}^2 &= \|M_{\text{cts}} - M_{\text{ABM}}\|_{L^2}^2 = \int_{t=0}^{t=t_{\text{final}}} \int_{\Omega} (M_{\text{cts}}(\mathbf{x}, t) - M_{\text{ABM}}(\mathbf{x}, t))^2 \, \mathbf{d}\mathbf{x} \, dt \\ &\approx \Delta t_{\text{record}} \Delta x \Delta y \sum_{n=0}^{N_T} \sum_{i=1}^{N_{\text{hist}}} \sum_{j=1}^{N_{\text{hist}}} (M_{\text{cts},ij}^{(n)} - M_{\text{ABM},ij}^{(n)})^2, \end{aligned} \quad (2.9)$$

where Δt_{record} denotes the time steps at which data are collected; Δx and Δy are the spatial step sizes of the histogram used to compare the EA ABM and PDE data; $M_{\text{cts},ij}^{(n)}$ is the continuous-model solution at time t_n and position (x_i, y_j) ; and $M_{\text{ABM},ij}^{(n)}$ is the corresponding EA ABM result. For the birth-only model, we consider fitting to either the L^2 error as above or simply the difference in the total cell count of the two datasets (we verify in electronic supplementary material, S2, table S4, that fitting to the L^2 error produces similar parameter estimates). When we consider cell birth, we simulate our models with different values of N_{bir} and estimate parameters by minimizing the sum of the errors across these N_{bir} values. We fit parameters related to cell proliferation sequentially—that is, we determine the optimal value for c^+ before γ . We verify in 1D that sequential and simultaneous estimation does not lead to significant difference in parameter values; see electronic supplementary material, S3.

3. Results

We now present our results linking discrete and continuous models of cell migration (§2.1), birth (§2.2), and migration and birth (§2.3). We first isolate each interaction process, separately identifying the values of α_{MM} in equation (2.4) and $\{\gamma, c^+\}$ in equation (2.7). As we note in §2.4, this choice allows us to extract the distinct effects of each mechanism. We then determine how this simplification affects the ability of the full continuous framework, given by equation (2.8), to approximate EA ABM solutions. By considering different initial conditions (we discuss the details and motivations for these in electronic supplementary material, S4), we demonstrate the robustness of our fitting procedure. Our results show how the time scales of proliferation and movement in our continuous model may depend on numerical implementation and the frequency of stochastic cell birth controlled by N_{bir} . Moreover, our modular fitting approach highlights important considerations to account for in more general systems where agents are moving and changing in number.

3.1. Cell migration

We estimate α_{MM} , the scaling parameter that controls the dynamics of melanophore movement. Figure 4a presents the values of α_{MM} that minimize the L^2 error between the continuous solution of equation (2.4) and EA ABM results for our four initial conditions (see §2.4 and electronic supplementary material, S2–

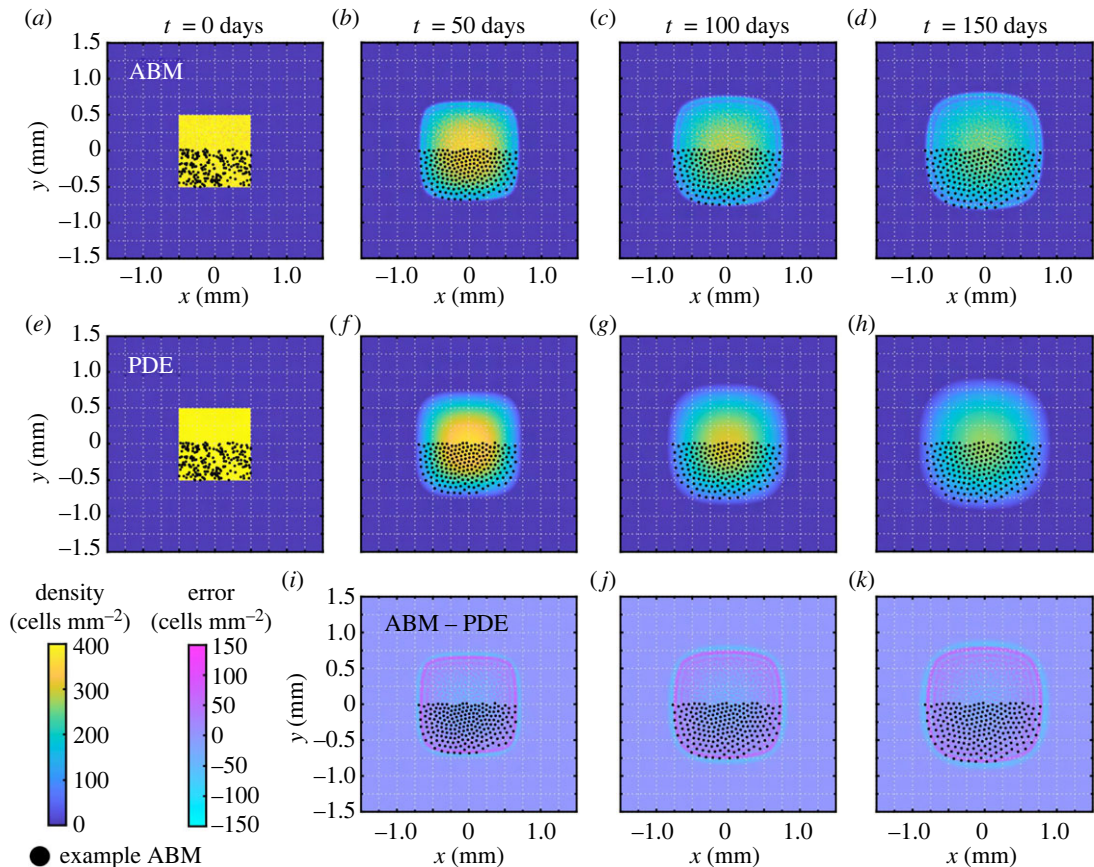


Figure 5. Melanophore movement models with our *Box* initial condition. We present (a–d) snapshots of the EA cell density (cells mm^{-2}) across 10^4 ABM realizations, (e–h) the corresponding PDE results using the optimal value of α_{MM} for a mesh resolution of $N_{\text{bin}} = 240$, and (i–k) the error between the PDE and EA ABM densities. We overlay cell positions for one example ABM simulation (black points) as a visual guide; we show roughly half of the region occupied by these cells. A difference in cell density of $150 \text{ cells mm}^{-2}$ in a given voxel corresponds to about 0.0234 cells for this choice of mesh resolution. We find that the average pointwise errors (over voxels where at least one of the EA ABM or PDE solutions is non-zero) are about 28 cells mm^{-2} , 17 cells mm^{-2} , 17 cells mm^{-2} and 17 cells mm^{-2} at $t = 0, 50, 100$ and 150 days, respectively (these values correspond, respectively, to roughly 7%, 4%, 4% and 4% of the maximum cell density of $400 \text{ cells mm}^{-2}$).

S4). In each case, the optimal value of α_{MM} is positively correlated with our PDE mesh resolution, i.e. greater values of α_{MM} are associated with larger $N_{\text{bin}} = N_{\text{hist}}$ values. This unitless parameter appears to converge to around 0.60–0.66 as the mesh resolution increases. There is at most a 2.5% relative difference between the values of α_{MM} that we find when $N_{\text{bin}} = 240$ versus when $N_{\text{bin}} = 480$ for our *Box* initial condition. These results suggest that α_{MM} is independent of the mesh resolution when the latter contains at least 240×240 voxels, corresponding to a mesh spacing of $12.5 \mu\text{m}$. As we show in figure 2c, melanophores tend to separate by between 60 and $100 \mu\text{m}$ in our ABM results, so this mesh spacing is less than one quarter of the typical distance between agents.

At each mesh resolution in figure 4a, the estimated optimal value of α_{MM} does not appear to depend greatly on the initial condition. For example, in the case of a mesh with $N_{\text{bin}} = N_{\text{hist}} = 240$, the maximum relative difference between the four parameter values is at most 6.5%. This similarity suggests that there is an inherent time scale at which migratory melanophore–melanophore interactions occur. Figure 4b, which presents the log L^2 error for $N_{\text{bin}} = 240$ as a function of α_{MM} , further supports this conclusion. Although the errors associated with different initial conditions can vary by an order of magnitude, the minimum value of each (roughly convex) curve appears nearly identical and is located near the values shown in figure 4a.

Figure 5 presents snapshots of the EA ABM results across 10^4 realizations of equation (2.2) and the optimized PDE solution associated with the *Box* initial condition. The first row shows the expansion in time of the EA ABM support, i.e. the area occupied by the cells, due to melanophore–melanophore repulsion. For more intuition, we superimpose the cell positions from one ABM realization on our number-density results in this figure and throughout the paper. In all cases, we crop out approximately the upper half of cell positions. Visual inspection of cell positions in figure 5 suggests

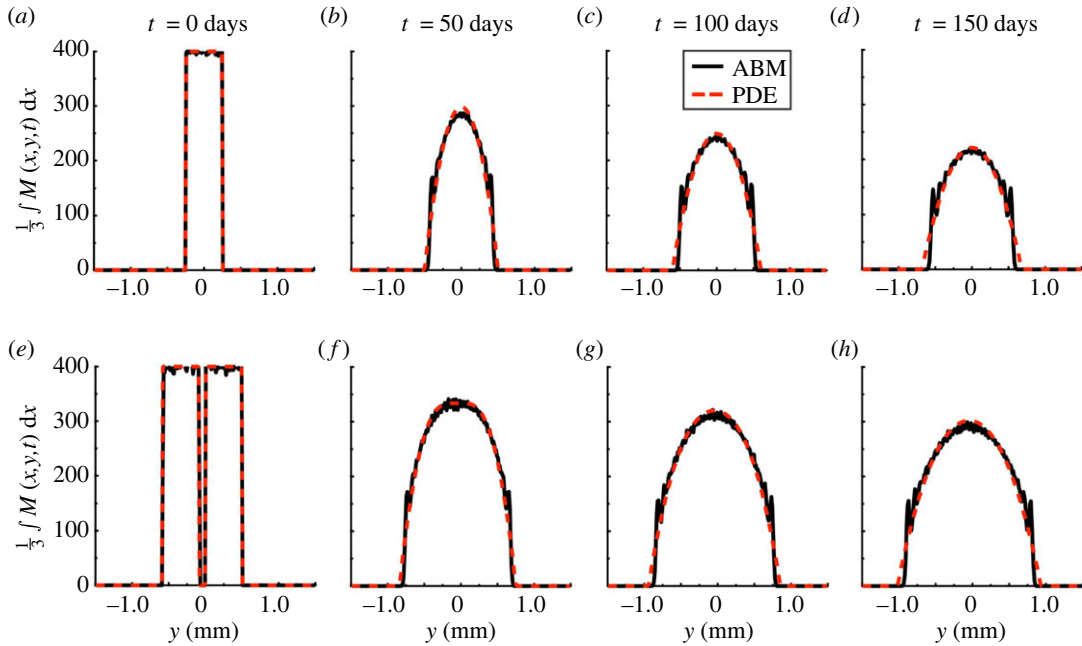


Figure 6. Melanophore movement models with our *Stripe* and *Two stripes* initial conditions. (a–d) We present snapshots of the column-averaged cell density (cells mm^{-2} , black solid line), generated from 10^3 ABM realizations for the *Stripe* case, alongside the corresponding PDE solution (dashed red line) under a mesh resolution of $N_{\text{bin}} = 240$ and our optimized value of α_{MM} . (e–h) Similarly, we show snapshots of the column-averaged density, generated from 10^3 ABM realizations for the *Two stripes* case, and the corresponding PDE solution. The 2D solutions are nearly uniform in the x -direction (data not shown).

that melanophore–melanophore distances increase near the edge of the collective. Similarly, the speed at which the support expands appears to slow down for the EA ABM result, consistent with melanophores experiencing weaker forces from comparatively distant cells in this region.

We also observe in [figure 5a–d](#) that a band of high cell density emerges around the edge of the support which surrounds a ring-like region of low density. These bands may result from the combined effects of cell–cell repulsion and the fine mesh resolution that we use to sort agent positions in the EA solution. Repulsion causes cells at the edge of the collective to travel towards empty regions, while more centrally located agents move more slowly due to the balance of forces from their neighbours. When repulsion separates cells by distances greater than the mesh resolution, we expect regions of low density within the solution support to appear. These oscillatory bands should become less evident when the repulsive potentials in [figure 1d](#) exhibit shallower gradients, as this permits cells to cluster more closely, or when coarser histograms with fewer bins are used to visualize the EA ABM data. As we discuss in electronic supplementary material, S6, the forces acting on xanthophores are about an order of magnitude smaller than those for melanophores, and we indeed observe less pronounced bands there. Notably, fitting to EA ABM data on coarser histograms leads to similar parameter estimates; see electronic supplementary material, S3, for details.

We present snapshots of the continuous model, equation (2.4), under our estimated value of α_{MM} in [figure 5e–h](#). This PDE solution captures the dynamics of our example ABM realization significantly better than the case in [figure 2b](#), when $\alpha_{\text{MM}} = 1$. However, unlike the EA ABM result, the PDE does not exhibit bands of high and low cell density. This discrepancy can be further appreciated in [figure 5i–k](#), which presents snapshots of the pointwise difference between the PDE and EA ABM solutions. Here, positive values indicate that the discrete solution is larger than the continuous one. The lack of bands in the PDE setting is likely because the mean-field assumption used to derive the continuous system is invalid where density is low. We do not expect this discrepancy to be as pronounced in models that include cell birth, as this mechanism increases density; see §3.3. Moreover, the PDE support expands more quickly than that of the ABM. This result is likely due to our choice of error function to fit α_{MM} . Specifically, this parameter is biased towards values that produce accurate approximations in the bulk as these regions have a larger contribution to the L^2 norm. Since we have already determined that the assumptions underlying the continuous model break down in low density regions, however, we choose to fit to the bulk of the cell density and focus on the L^2 difference.

To demonstrate that our observations for the *Box* case are consistent across initial conditions, we compare the EA ABM and PDE dynamics for the *Stripe* and *Two stripes* initial conditions in [figure 6](#); electronic

supplementary material, figure S7, presents results for the *Offset rectangles* initial condition. In figure 6a–d, the column-averaged PDE solution, i.e. the solution average over the x variable, has a larger support than that of the EA ABM and does not exhibit oscillatory bands. (Comparing column averages is justified because both results are nearly uniform along the x -axis.) Nevertheless, the continuous solution closely approximates the EA ABM density, particularly in regions where the latter is high. For example, we find that the average pointwise error (over voxels where at least one of the EA ABM or PDE solution is non-zero) is equal to about 40 cells mm^{-2} at $t=0$ days, 24 cells mm^{-2} at $t=50$ days, 23 cells mm^{-2} at $t=100$ days and 23 cells mm^{-2} at $t=150$ days (these correspond to roughly 10%, 6%, 6% and 6% of the maximum cell density of 400 cells mm^{-2} , respectively). Both solutions invade empty space in time, and the speed of this travelling wavefront appears to slow as cells become more diffuse. For the *Two stripes* initial condition in figure 6e–h, the ABM and PDE predict that cells move into the initially empty space between stripes to approach a characteristic profile also observed in the one-stripe case. The EA ABM model does not appear to form oscillatory bands in the interstripe region, corroborating our hypothesis that these bands are more likely to arise near the edge of the solution support. In this case, the average pointwise error between the continuous and discrete data is roughly 18%, 11%, 11% and 11% of the maximum cell density at $t=0$, 50, 100 and 150 days, respectively.

3.2. Cell birth

We identify the density-limiting parameter c^+ and growth rate γ in our IDE model, equation (2.7), by comparing with agent-based data from equation (2.5). Importantly, the dynamics of discrete-model proliferation, unlike cell migration, involve stochasticity beyond the initial condition. To gain intuition, we thus start with 1D simulations: for each value of $N_{\text{bir}} \in \{1, 2, \dots, 10\}$, we compute the EA of 10^3 ABM realizations from an initial condition in which a single melanophore is placed at the origin in a 1D domain. In figure 7a, we show the EA result for $N_{\text{bir}}=1$ and the corresponding IDE model solution with the optimal values of γ and c^+ in figure 7b. The continuous solution appears to have a smaller radius of support than the EA ABM result at every time point; see figure 7c. This result holds across all N_{bir} values in figure 7d. While the IDE predicts a piecewise linear growth of the total number of cells, the corresponding EA ABM result increases linearly before slowly saturating as the domain fills, as we depict in figure 7e. This behaviour likely arises from our overcrowding condition that prevents cell densities from exceeding c^+ . As the domain fills with cells, it becomes less likely to select a location \mathbf{z} that satisfies the overcrowding condition in the ABM. This reduces the population growth rate at later times. By contrast, the IDE model specifies that the support increases by the same amount at each time step until it reaches the domain boundaries. As we discuss in §4, capturing discrete model behaviour more accurately at higher cell numbers may require replacing γ in our IDE with a density-dependent function.

Our 1D simulations provide a baseline case to test our estimation process. As we note in §2.4, we employ a sequential procedure, first fitting c^+ with $\gamma = |\Omega|^{-1}$ and then estimating γ with c^+ fixed. In the 1D case, this leads to optimal values $c^+ = 7.592$ cells and $\gamma = 0.2822$. If we instead estimate both parameters simultaneously, we find $c^+ = 7.430$ cells and $\gamma = 0.2902$. This is a difference of about 2.1% in c^+ and 2.8% in γ , suggesting that sequential estimation reduces computational complexity without strongly affecting parameter values. To understand if a coarser discrepancy measure based only on cell numbers at each time is sufficient, we also fit c^+ and γ by minimizing the squared difference in the total cell numbers over time; see electronic supplementary material, S2 and table S4, for the resulting parameter values. The corresponding parameter estimates differ from the density-based case by approximately 1.2% for c^+ and γ , suggesting both error measures are reasonable. Both approaches also appear to exhibit similar sensitivity as parameters are varied (compare figure 7; electronic supplementary material, figure S2).

Figure 8 and electronic supplementary material, figure S8, respectively, show that proliferation in 2D broadens the solution support from the *Box* and *Offset rectangles* initial conditions over time, and the IDE model accurately captures the total cell mass of the ABM system for all N_{bir} values considered. Our estimated optimal values of c^+ and γ for these two initial conditions differ by about 2.5% and 0.31%, respectively, suggesting that our estimation procedure is robust to the initial condition. We also highlight that a region of higher density forms at the edge of the initial condition's support for both the ABM and IDE in figure 8a–h. Indeed, if \mathbf{z} is near the support boundary, Ω_{loc}^z covers only a fraction of the occupied domain, thereby meeting both conditions for birth. Conversely, the cell density at the centre of the domain is comparatively low throughout time because the total number of cells contained within discs of size $|\Omega_{\text{loc}}^z|$ is already close to the threshold c^+ . Interestingly, as in the 1D case

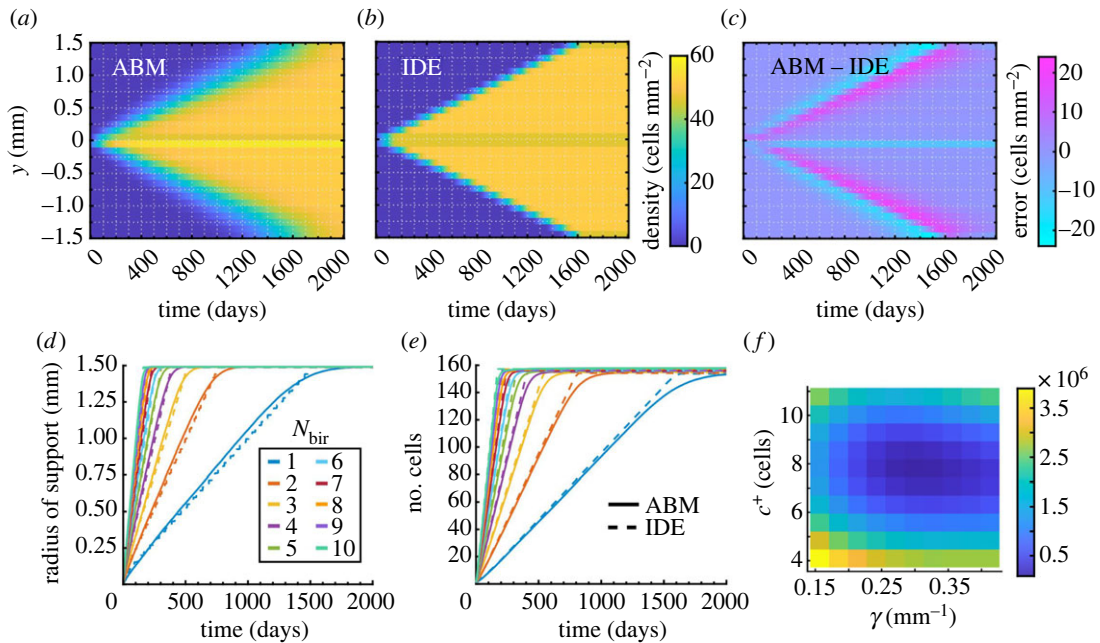


Figure 7. Melanophore birth models with a baseline initial condition of one cell at $y = 0$ in a 1D domain. Results in (a–c) are for $N_{\text{bir}} = 1$ positions day $^{-1}$. (a) We compute the EA ABM result by simulating 10^3 realizations of our ABM birth model, equation (2.5), and binning cell positions in a histogram with 0.1 mm-wide voxels (i.e. $N_{\text{hist}} = 30$). (b) We use a finer mesh resolution to solve our corresponding IDE model (equation (2.7)) before transferring results to the same histogram in (a) to perform parameter estimation. Here, we show our IDE solution produced with optimal parameter values $c^+ = 7.592$ cells and $\gamma = 0.2822$. (c) The difference between our discrete and continuous results highlights that the ABM support is wider than the PDE support. (d) This is also visible by comparing their mean radii of support in time. To compute the mean radius of support for the ABM at a given time, we find the most distant melanophore from $y = 0$ for each simulation and average across these values. (e) Cell mass grows linearly in both models at first, but stochastic effects coupled with our overcrowding condition drive down the growth rate of the ABM as the domain fills with cells. (f) Plotting the squared L^2 space–time difference between the discrete and continuous densities, summed over all N_{bir} values considered (namely $N_{\text{bir}} = 1, \dots, 10$), as a function of the density-limiting parameter c^+ and birth-rate scaling parameter γ highlights its convex shape in c^+ and lesser sensitivity to γ . We compute this L^2 difference using a time step of 10 days here, and our results are based on 10^3 simulations for each N_{bir} value; see electronic supplementary material, S3, for parameter values under alternative choices in our estimation process.

with only proliferation, the ABM EA support is larger than that of the IDE solution, the reverse of the behaviour that we observed for cell migration in figure 5.

3.3. Cell movement and proliferation

To obtain a full continuous model, we may substitute our estimated values of the migration scaling parameter α_{MM} , density-limiting parameter c^+ and birth-rate scaling parameter γ into equation (2.8). However, comparing this model to the dynamics of our full ABM shows that migration and proliferation have interwoven effects. To illustrate this phenomenon, we present a PDE solution with our optimal values of α_{MM} , c^+ and γ from §§3.1 and 3.2 at $t = 70$ days in figure 9a. We observe that this PDE model produces a significantly higher cell density than its discrete counterpart in figure 9b. This discrepancy occurs regardless of the value of N_{bir} , which influences the speed of cell birth. Related to this, we note that the long-time cell density in our ABM results is much lower when both mechanisms operate simultaneously than it is when only birth occurs; compare figure 8d and figure 9b. On the other hand, the inclusion of movement does not influence the long-time density of the continuous model solution; see the colourbar in figure 8h in comparison to the one in figure 9. Although we do not furnish these observations with an analytical explanation here, they demonstrate an interesting difference in how ‘adding’ mechanisms or terms impact PDE and ABM dynamics.

One approach to addressing these discrepancies is to refit all three scaling parameters (α_{MM} , γ and c^+) simultaneously, and we present the results of this approach in electronic supplementary material, S3. (Indeed, we show there that the errors produced with a simultaneous estimation approach can be

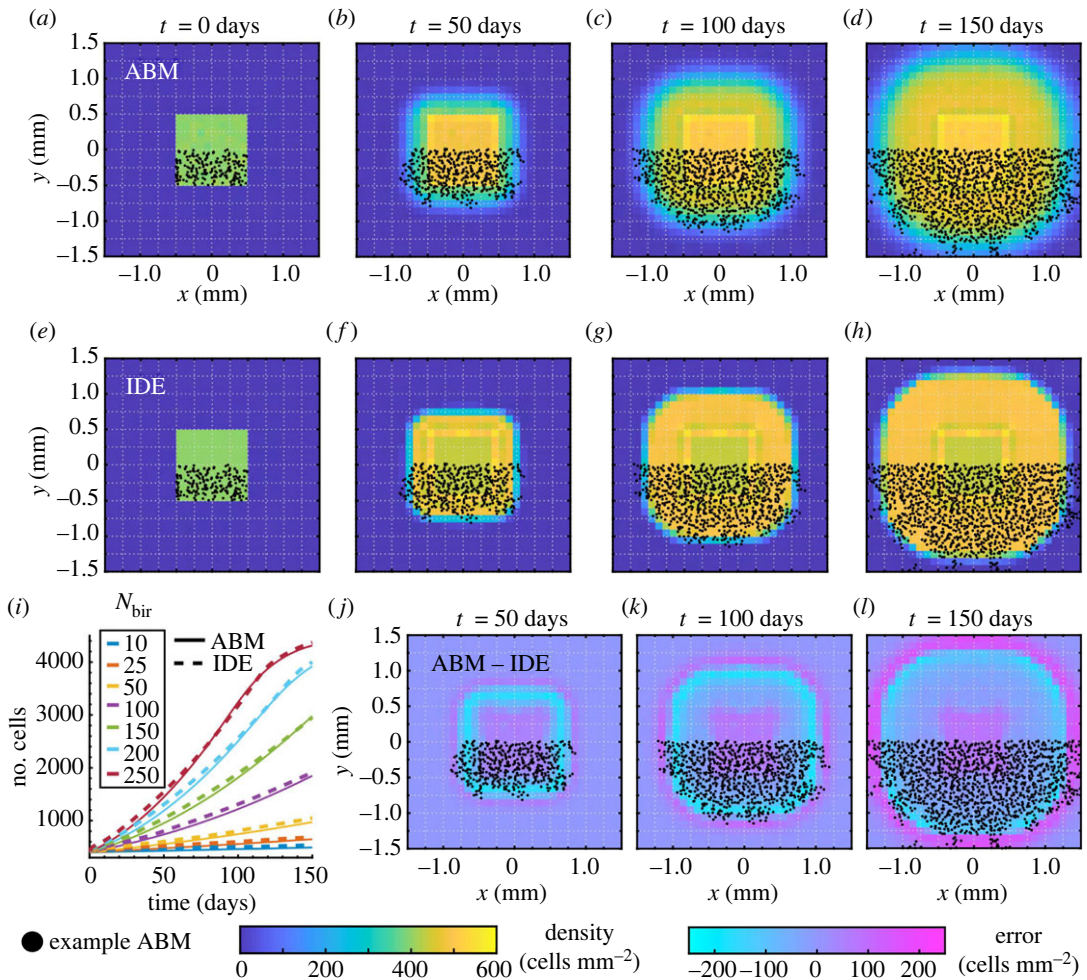


Figure 8. Melanophore proliferation models with our *Box* initial condition. Results in (a–h) and (j–l) are for $N_{\text{bir}} = 150$ positions day^{-1} . We compute (a–d) the EA ABM result across 10^3 simulations, and (e–h) the solution of our IDE model with optimal parameters $c^+ = 8.564$ and $\gamma = 0.1274$. (i) Our continuous model captures the mean number of cells in our ABM simulations for different N_{bir} values well across time. (j–l) As in the 1D case in figure 7, the difference between the IDE and EA ABM results demonstrates that the ABM support extends beyond the IDE support. To provide more intuition, we overlay roughly half of the cell positions from an example ABM simulation in (a–h) and (j–l). See electronic supplementary material, figure S8 and table S4, for corresponding simulations using our *Offset rectangles* initial condition.

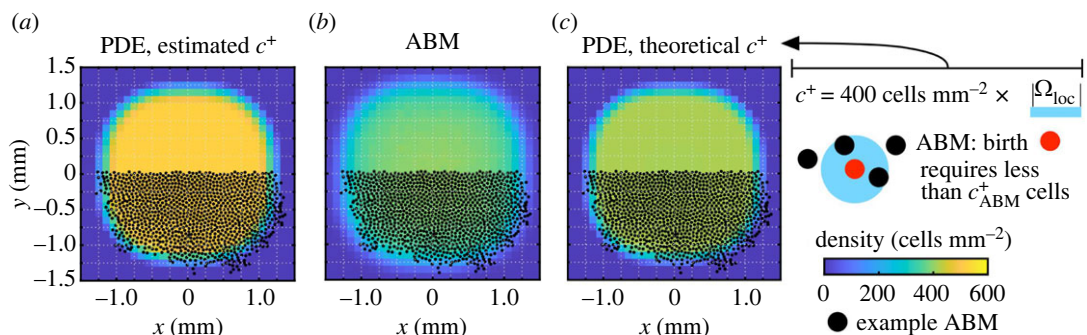


Figure 9. A modular approach to fitting parameters for cell movement and birth does not account for the interplay between these two mechanisms. We show results in (a–c) at $t = 100$ days for $N_{\text{bir}} = 150$ positions day^{-1} . (a) The solution of our full PDE model (equation (2.8)) with the values of α_{MM} and $\{c^+, \gamma\}$ that we fit based on ABM simulations of cell movement and birth, respectively, captures the support of the ABM EA result, but not its density. (b) In comparison, the density for the full discrete model is roughly $400 \text{ cells } \text{mm}^{-2}$. (c) By integrating this density, which is based on empirical estimates of melanophore–melanophore distances [14,78], over an Ω_{loc} -region, we find that $c^+ \approx 7.0686$ cells. With this value of c^+ , alongside the values of α_{MM} and γ that we estimated for migration and birth individually, our PDE produces cell densities that more accurately represent the ABM dynamics.

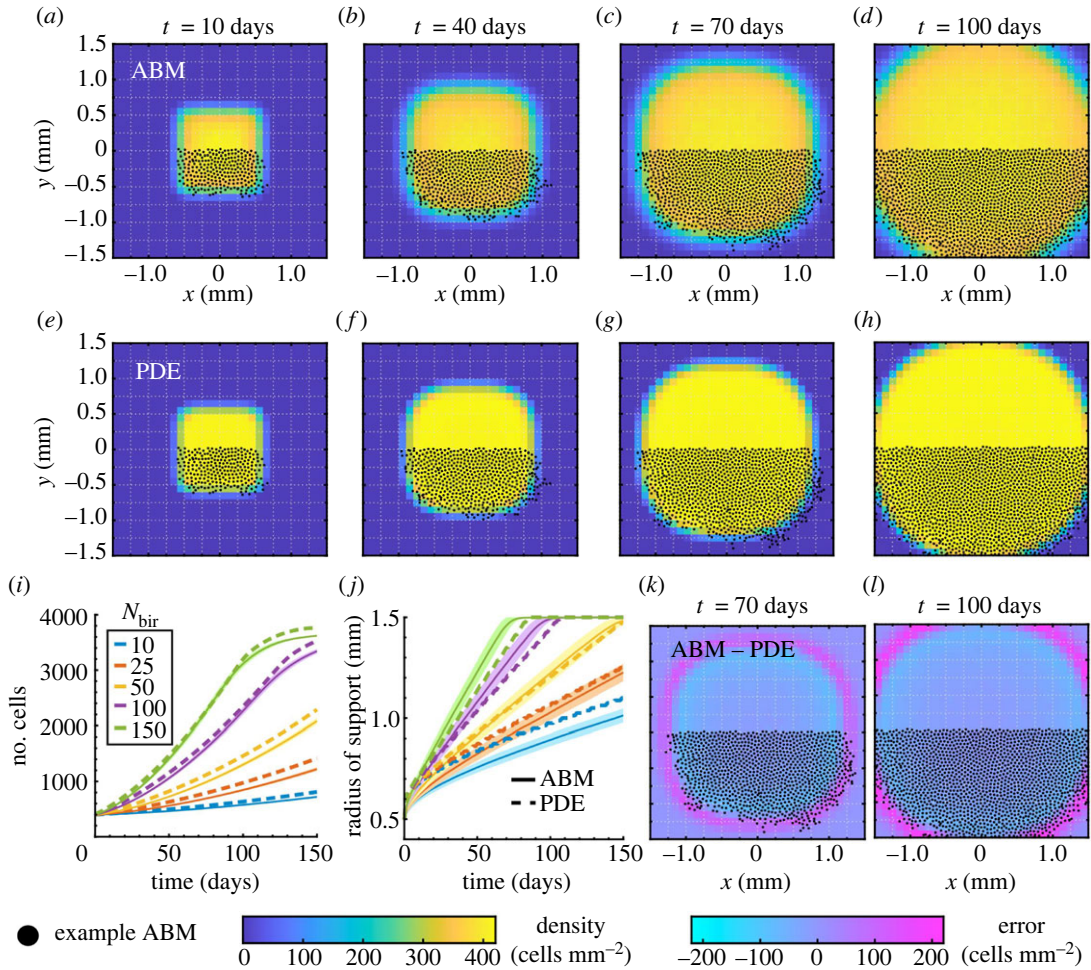


Figure 10. Melanophore movement and birth models with our *Box* initial condition. Results in (a–h) and (j–l) are for $N_{\text{bir}} = 150$ positions day^{-1} . We (a–d) compute the EA ABM result using 10^3 simulations, and (e–h) generate the PDE solution of equation (2.8) with $c^+ = 7.0686$ cells and the values of α_{MM} and γ that we estimated in §§3.1 and 3.2, respectively. (i) The time evolution of the PDE cell mass agrees well with the mean number of cells for the ABM under different N_{bir} values. (j) Depending on the time scales of migration and birth, the approximate PDE radius of support overtakes or trails the corresponding EA ABM result. We compute the radius of support for each ABM realization by finding the most distant cell from the origin at each time step; we then average these values across our simulations. In the PDE case, we find the furthest voxel with non-zero density from the origin based on the L^∞ distance, after setting the density to zero if it is below single-digit precision of 10^{-7} . (k, l) We show the difference between the PDE and EA ABM solutions from (a–h) at two sample times. We overlay cell positions from one ABM simulation to illustrate how the continuous and discrete solutions are related. In (i, j), shaded regions denote plus or minus one standard deviation of the EA ABM solution.

relatively small, although the model parameters may not all be identifiable.) Because we are interested in understanding the interplay of individual-based mechanisms of proliferation and movement in continuous models, however, we instead take a simpler theoretical approach. Namely, we note that the parameter c^+ is largely responsible for controlling the maximum cell density over long time periods. (We determine this by integrating equation (2.8) over space and identifying the steady-state dynamics; this analysis reveals that equilibrium is reached when the density within any neighbourhood Ω_{loc}^x is below c^+ .) In order to limit the maximum density to our estimated empirical value of $400 \text{ cells mm}^{-2}$ [14,78], we let $c^+ = 400 |\Omega_{\text{loc}}^x| \approx 7.0686$ cells. As we show in figure 9c, using this value of c^+ , alongside our previously fitted values of α_{MM} and γ , produces PDE densities that are much closer to the corresponding ABM results. We thus fix $c^+ = 7.0686$ cells for the remainder of this paper, which allows us to highlight the time dynamics of our full PDE model in comparison to the EA ABM result with *Box* and *Offset rectangles* initial conditions in figures 10 and 11, respectively.

Figure 10j, which depicts the time evolution of the estimated radius of support for the PDE and EA ABM results, shows that a reasonably accurate continuous description of the combined model can be obtained by using the scaling parameters obtained from a modular approach. The supports of the ABM and PDE

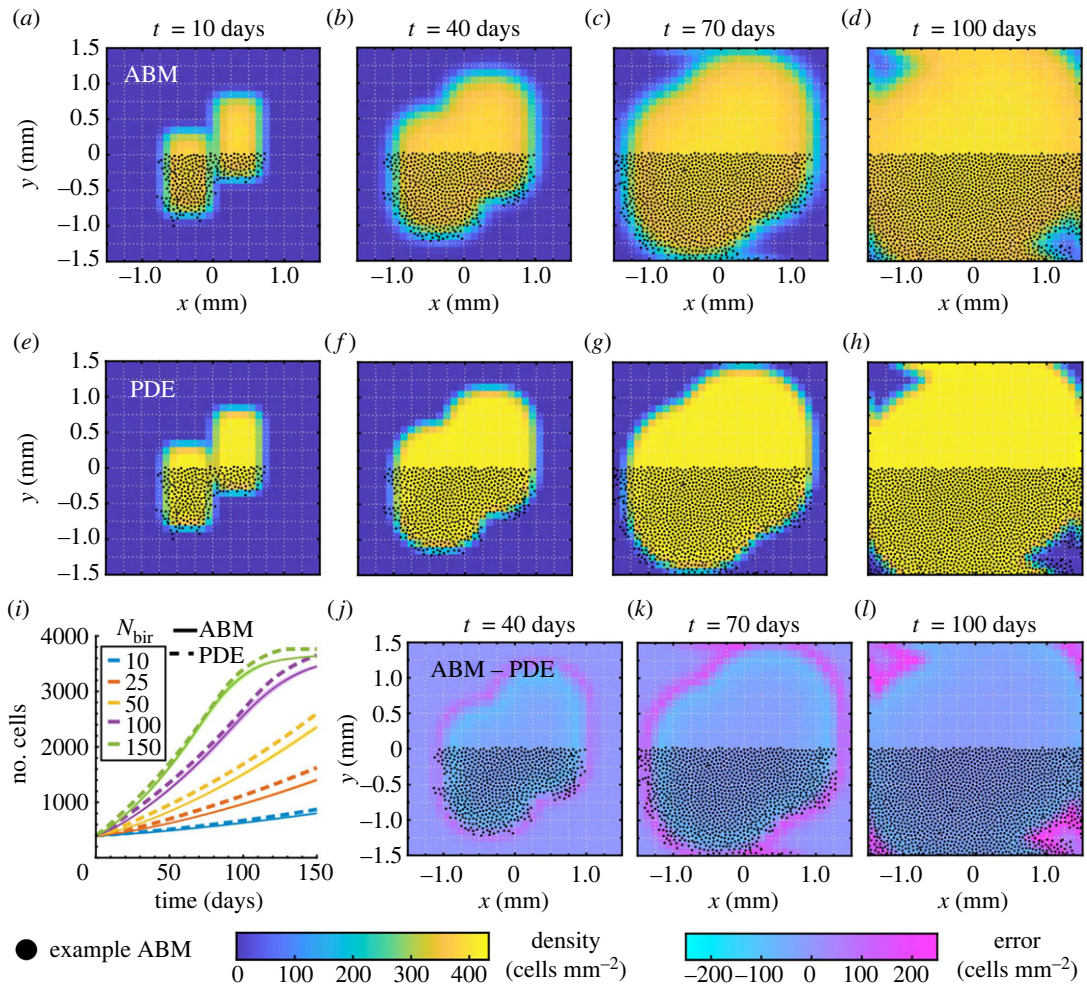


Figure 11. Melanophore movement and birth models with our *Offset rectangles* initial condition to verify that our fitting procedure is robust in non-standard geometries. Results in (a–h) and (j–l) are for $N_{\text{bir}} = 150$ positions day⁻¹. As in figure 10, we show (a–d) the EA ABM result across 10^3 simulations; (e–h) the corresponding PDE solution with $c^+ = 7.0686$ cells and the values of α_{MM} and γ that we estimated in §§3.1 and 3.2, respectively; (i) the PDE cell mass in time compared to the mean number of cells across 10^3 ABM simulations for different N_{bir} values; and (j–l) the difference between the PDE and EA ABM solutions from (a–h). To provide more intuition, we overlay some cell positions from one ABM simulation.

solutions both increase at roughly the same rate, although the degree to which the solutions agree can be affected by N_{bir} : when this parameter is small, the PDE solution travels at a faster rate than the ABM solution, whereas the opposite occurs when this value is large (greater than $N_{\text{bir}} = 50$ positions day⁻¹). At intermediate values of N_{bir} (i.e. $N_{\text{bir}} = 50$ positions day⁻¹), however, the ABM and PDE solution curves are almost identical. Figure 10i, which presents the number of cells over time, yields similar observations: the ABM and PDE solutions exhibit similar dynamics over the time period investigated here and there are certain values of N_{bir} for which the solution curves are nearly identical. Figure 11 further demonstrates that these observations do not depend on the specific choice of initial condition. Figures 10 and 11 demonstrate that combining movement with proliferation also dissipates the oscillatory bands that we observed for movement alone in figure 5. This is likely because the stochastic addition of cells in the birth model disrupts the regular cell spacing created by the movement model. Furthermore, the EA ABM and PDE solutions exhibit similar characteristic profiles without regions of high cell density around the edge of the initial condition support, in contrast to the birth-only model (figure 8).

4. Discussion

We presented a procedure for constructing experimentally interpretable continuous models of cell migration and birth in biologically relevant settings of low numbers of individuals and localized

interactions which may lie outside the validity of the mean-field regime. Specifically, we introduced and estimated scaling parameters in continuous models to account for realistic—i.e. relatively small and changing—numbers of cells with localized interactions. We applied this methodology to an illustrative, simplified example inspired by zebrafish pattern formation, in which we used a reduced ABM to generate individual-level data with biologically meaningful spatial and temporal units. Non-local rules for cell birth and migration, based on the ABM [14], informed our discrete and continuous descriptions and allowed us to transfer biological length scales and units to the macroscopic setting. Throughout our work, we stressed matching the spatio-temporal behaviour of our continuous and discrete models. We adopted a modular approach by estimating parameters in cases with either movement or birth before considering both mechanisms simultaneously. This allowed us to examine the specific contributions of each mechanism to self-organization and provided insight into their interplay in discrete and continuous settings.

We observed that the solutions of our continuous models expand at a different rate than EA ABM results and feature smoother profiles. Indeed, inaccuracies in mean-field descriptions for ‘intermediate’ numbers of individuals appear to be common in other biological phenomena described by simpler dynamics such as Fisher-KPP-type equations, cf. [85,86]. In fact, both references analytically derive corrections to the wave speed, a procedure we cannot adopt due to our use of off-lattice models. However, this substantiates our introduction of scaling parameters to handle the discrepancy. By introducing and estimating parameters that rescale the time variable, we produced more accurate descriptions of agent-based movement or birth. However, when we used the same parameter values in a continuous model of both cell migration and birth in §3.3, the PDE did not produce close estimates of the full ABM. Specifically, our full continuous model yielded larger long-time densities than the EA ABM results, motivating us to re-estimate the threshold value c^+ with a theoretical approach. This generated a more faithful continuous description and highlighted that the effects of movement and proliferation are not simply additive. We thus stress that parameters must be fitted to data in which all mechanisms of interest act simultaneously, in order to capture their interplay. This is particularly crucial for contexts such as cancer biology, where cell migration, proliferation and death are known to play critical roles in tumour progression and immune response [87].

Our results highlight how choices in numerical implementation affect parameter estimates and suggest several directions for future work that may improve our approach. For example, the optimal value of our parameter controlling the time scale of cell migration (α_{MM}) appears to be independent of the initial condition and the mesh resolution that we used to construct PDE solutions, provided the latter is sufficiently refined. One drawback of our current approach, however, is that we may need to estimate α_{MM} and γ for each new choice of discrete rules governing migration and cell birth, respectively, because these rules perturb the short-range interactions between relatively small numbers of cells. This naturally leads to the question of whether an analytic expression can be derived for these parameters. Several coarse-graining techniques that take into account higher-order correlations between cells in on- or off-lattice models may produce scaling factors similar to those introduced in this paper, but these may only apply to certain classes of ABMs [88]. Alternatively, it may be possible to estimate scaling parameters of continuous models by investigating the convergence of EA ABM results to features of their solutions such as the speed of solution propagation, as in [85,86] for on-lattice models; by accounting for the dynamics of the two-particle distribution, as in [73,74]; or by exploring scaling relationships as in [89]. Adapting these approaches to our setting is an interesting avenue for future work.

Additionally, our continuous models more accurately represent ABM results within the bulk of the solution support because the L^2 norm more strongly penalizes discrepancies there. In the future, other norms, such as the L^∞ error, could be used to match the solution supports given by our discrete and continuous models. Replacing our birth-rate scaling parameter γ with a density-dependent function—through either rigorous derivation or an equation-learning approach [90]—is another exciting future direction. In particular, because cell proliferation in the ABM involves selecting positions uniformly at random from the domain each day, the chance that we select a location \mathbf{z} that permits birth appears to depend in a nonlinear way on the solution support. More generally, our computational study does not provide theoretical explanations for our parameter values, and we plan to build on the intuition that we established here to develop these arguments in the future.

To simplify our initial study, we considered the dynamics of one cell population (i.e. melanophores in the main text and xanthophores in electronic supplementary material, S6), but pattern formation in zebrafish skin involves multiple cell types and longer-range interactions, as we highlight in electronic supplementary material, figure S1. Future work may extend our pipeline to construct more realistic

continuous models with multiple cell types and interaction neighbourhoods. Related to this, the initial conditions that we designed allowed us to make one-to-one comparisons between discrete- and continuous-model densities, but this may not always be possible. More realistic zebrafish models (i.e. [14,37,39,40]) produce patterns that are more complicated than our box and stripe motifs. This means that ensemble-averaging stochastic ABM realizations may not retain information about the length scales inherent in patterns. For such cases, fitting parameters based on summary statistics (e.g. pair-correlation functions [91], pattern-simplicity scores [92] or persistent-homology approaches [93]) may be more useful, and we plan to address this in future work. These and other directions move us toward constructing interpretable, analytically tractable continuous models of self-organization, increasing our understanding of biological pattern formation more broadly.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. Data and relevant code for this research work are stored in GitHub: https://github.com/wdmartinson/Self-Organization-One-Species/tree/first_version and have been archived within the Zenodo repository: <https://doi.org/10.5281/zenodo.11160930> [94]. Electronic supplementary material, figures and videos are available on Figshare at https://figshare.com/projects/Linking_discrete_and_continuous_models_of_cell_birth_and_migration_in_one_population/171234. Raw data used to generate figures in the main text and electronic supplementary material have been archived in the Dryad repository <https://doi.org/10.5061/dryad.s4mw6m9cb> [95].

Supplementary material is available online [96].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. W.D.M.: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing; A.V.: conceptualization, data curation, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing; M.M.: conceptualization, data curation, investigation, methodology, software, validation, writing—original draft, writing—review and editing; C.V.: conceptualization, data curation, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing; J.A.C.: conceptualization, funding acquisition, methodology, project administration, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. J.A.C. and W.D.M. were supported by the Advanced Grant Non-local-CPD (Non-local PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronization) of the European Research Council Executive Agency (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 883363). J.A.C. is also partially supported by the 'Maria de Maeztu' Excellence Unit IMAG, reference CEX2020-001105-M, funded by MCIN/AEI/10.13039/501100011033/. W.D.M., J.A.C. and A.V. would also like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Mathematics of Movement, where work on this paper was undertaken. This work was supported by EPSRC grant no. EP/R014604/1. C.V. acknowledges support from the Dr Perry James (Jim) Browne Research Centre on Mathematics and its Applications (University of Sussex).

Acknowledgements. We are grateful to Shigeru Kondo for helpful discussions during an early stage of this research.

References

- Mogilner A, Edelstein-Keshet L. 1999 A non-local model for a swarm. *J. Math. Biol.* **38**, 534–570. (doi:10.1007/s002850050158)
- D'Orsogna MR, Chuang YL, Bertozzi AL, Chayes LS. 2006 Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Phys. Rev. Lett.* **96**, 104302. (doi:10.1103/PhysRevLett.96.104302)
- Cucker F, Smale S. 2007 Emergent behavior in flocks. *IEEE Trans. Autom. Control* **52**, 852–862. (doi:10.1109/TAC.2007.895842)
- Carrillo JA, Fornasier M, Toscani G, Vecil F. 2010 Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, pp. 297–336. New York, NY: Springer.
- Amack JD, Manning ML. 2012 Knowing the boundaries: extending the differential adhesion hypothesis in embryonic cell sorting. *Science* **338**, 212–215. (doi:10.1126/science.1223953)
- Burger M, Francesco MD, Fagioli S, Stevens A. 2018 Sorting phenomena in a mathematical model for two mutually attracting/repelling species. *SIAM J. Math. Anal.* **50**, 3210–3250. (doi:10.1137/17M1125716)
- Carrillo JA, Murakawa H, Sato M, Togashi H, Trush O. 2019 A population dynamics model of cell-cell adhesion incorporating population pressure and density saturation. *J. Theor. Biol.* **474**, 14–24. (doi:10.1016/j.jtbi.2019.04.023)
- Buttenschön A, Edelstein-Keshet L. 2020 Bridging from single to collective cell migration: a review of models and links to experiments. *PLoS Comput. Biol.* **16**, e1008411. (doi:10.1371/journal.pcbi.1008411)
- Tsai TYC, Garner RM, Megason SG. 2022 Adhesion-based self-organization in tissue patterning. *Annu. Rev. Cell. Dev. Biol.* **38**, 349–374. (doi:10.1146/annurev-cellbio-120420-100215)
- Frohnhofer HG, Krauss J, Maischein HM, Nüsslein-Volhard C. 2013 Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish. *Development* **140**, 2997–3007.
- Irion U, Nüsslein-Volhard C. 2019 The identification of genes involved in the evolution of color patterns in fish. *Curr. Opin. Genet. Dev.* **57**, 31–38. (doi:10.1016/j.gde.2019.07.002)
- Parichy DM. 2021 Evolution of pigment cells and patterns: recent insights from teleost fishes. *Curr. Opin. Genet. Dev.* **69**, 88–96. (doi:10.1016/j.gde.2021.02.006)
- Kondo S, Watanabe M, Miyazawa S. 2021 Studies of Turing pattern formation in zebrafish skin. *Phil. Trans. R. Soc. A* **379**, 20200274. (doi:10.1098/rsta.2020.0274)

14. Volkening A, Sandstede B. 2015 Modelling stripe formation in zebrafish: an agent-based approach. *J. R. Soc. Interface* **12**, 20150812. (doi:10.1098/rsif.2015.0812)
15. Fadeev A, Krauss J, Frohnhöfer HG, Irion U, Nüsslein-Volhard C. 2015 Tight Junction Protein 1a regulates pigment cell organisation during zebrafish colour patterning. *eLife* **4**, e06545. (doi:10.7554/eLife.06545)
16. Nakamasu A, Takahashi G, Kanbe A, Kondo S. 2009 Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. *Proc. Natl Acad. Sci. USA* **106**, 8429–8434. (doi:10.1073/pnas.0808622106)
17. Singh AP, Schach U, Nüsslein-Volhard C. 2014 Proliferation, dispersal and patterned aggregation of iridophores in the skin prefigure striped colouration of zebrafish. *Nat. Cell Biol.* **16**, 604–611. (doi:10.1038/ncb2955)
18. Gur D *et al.* 2020 *In situ* differentiation of iridophore crystalloids underlies zebrafish stripe patterning. *Nat. Commun.* **11**, 6391. (doi:10.1038/s41467-020-20088-1)
19. Hamada H, Watanabe M, Lau HE, Nishida T, Hasegawa T, Parichy DM, Kondo S. 2014 Involvement of Delta/Notch signaling in zebrafish adult pigment stripe patterning. *Development* **141**, 318–324. (doi:10.1242/dev.099804)
20. Inaba M, Yamanaka H, Kondo S. 2012 Pigment pattern formation by contact-dependent depolarization. *Science* **335**, 677–677. (doi:10.1126/science.1212821)
21. Yamaguchi M, Yoshimoto E, Kondo S. 2007 Pattern regulation in the stripe of zebrafish suggests an underlying dynamic and autonomous mechanism. *Proc. Natl Acad. Sci. USA* **104**, 4790–4793. (doi:10.1073/pnas.0607790104)
22. Kroll F *et al.* 2021 A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife* **10**, e59683. (doi:10.7554/eLife.59683)
23. Byrne HM. 2010 Dissecting cancer through mathematics: from the cell to the animal model. *Nat. Rev. Cancer* **10**, 221–230. (doi:10.1038/nrc2808)
24. Gompfer G *et al.* 2020 The 2020 motile active matter roadmap. *J. Phys. Condens. Matter* **32**, 193001. (doi:10.1088/1361-648X/ab6348)
25. Stillman NR, Mayor R. 2023 Generative models of morphogenesis in developmental biology. *Semin. Cell Dev. Biol.* **147**, 83–90. (doi:10.1016/j.semcdb.2023.02.001)
26. Volkening A. 2020 Linking genotype, cell behavior, and phenotype: multidisciplinary perspectives with a basis in zebrafish patterns. *Curr. Opin. Genet. Dev.* **63**, 78–85. (doi:10.1016/j.gde.2020.05.010)
27. Alert R, Trepat X. 2020 Physical models of collective cell migration. *Annu. Rev. Condens. Matter Phys.* **11**, 77–101. (doi:10.1146/annurev-conmatphys-031218-013516)
28. Metzcar J, Wang Y, Heiland R, Macklin P. 2019 A review of cell-based computational modeling in cancer biology. *JCO Clin. Cancer Inform.* **2**, 1–13. (doi:10.1200/CC.18.00069)
29. Deutch A. 2015 Cellular automaton models for collective cell behaviour. In *Cellular automata and discrete complex systems* (ed. J Kari), pp. 1–10. Berlin, Germany: Springer. (doi:10.1007/978-3-662-47221-7_1)
30. Deutch A, Dormann S. 2017 *Cellular automaton modeling of biological pattern formation, characterization, examples, and analysis*. New York, NY: Springer.
31. Hirashima T, Rens EG, Merks RMH. 2017 Cellular Potts modeling of complex multicellular behaviors in tissue morphogenesis. *Dev. Growth Differ.* **59**, 329–339. (doi:10.1111/dgd.12358)
32. Rens EG, Edelstein-Keshet L. 2019 From energy to cellular forces in the cellular Potts model: an algorithmic approach. *PLoS Comput. Biol.* **15**, e1007459. (doi:10.1371/journal.pcbi.1007459)
33. Alt S, Ganguly P, Salbreux G. 2017 Vertex models: from cell mechanics to tissue morphogenesis. *Phil. Trans. R. Soc. B* **372**, 20150520. (doi:10.1098/rstb.2015.0520)
34. Fletcher A, Osterfield M, Baker R, Shvartsman S. 2014 Vertex models of epithelial morphogenesis. *Biophys. J.* **106**, 2291–2304. (doi:10.1016/j.bpj.2013.11.4498)
35. Bullara D, De Decker Y. 2015 Pigment cell movement is not required for generation of Turing patterns in zebrafish skin. *Nat. Commun.* **6**, 6971. (doi:10.1038/ncomms7971)
36. Moreira J, Deutch A. 2005 Pigment pattern formation in zebrafish during late larval stages: a model based on local interactions. *Dev. Dyn.* **232**, 33–42. (doi:10.1002/dvdy.20199)
37. Owen JP, Kelsch RN, Yates CA. 2020 A quantitative modelling approach to zebrafish pigment pattern formation. *eLife* **9**, e52998. (doi:10.7554/eLife.52998)
38. Caicedo-Carvajal CE, Shinbrot T. 2008 *In silico* zebrafish pattern formation. *Dev. Biol.* **315**, 397–403. (doi:10.1016/j.ydbio.2007.12.036)
39. Volkening A, Abbott MR, Chandra N, Dubois B, Lim F, Sexton D, Sandstede B. 2020 Modeling stripe formation on growing zebrafish tailfins. *Bull. Math. Biol.* **82**, 56. (doi:10.1007/s11538-020-00731-0)
40. Volkening A, Sandstede B. 2018 Iridophores as a source of robustness in zebrafish stripes and variability in *Danio* patterns. *Nat. Commun.* **9**, 3231. (doi:10.1038/s41467-018-05629-z)
41. Osborne JM, Fletcher AG, Pitt-Francis JM, Maini PK, Gavaghan DJ. 2017 Comparing individual-based approaches to modelling the self-organization of multicellular tissues. *PLoS Comput. Biol.* **13**, e1005387. (doi:10.1371/journal.pcbi.1005387)
42. Murray JD. 2002 *Mathematical biology: I. An introduction*. New York, NY: Springer.
43. Perthame B. 2006 *Transport equations in biology*. Berlin, Germany: Springer Science & Business Media.
44. Turing AM. 1952 The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B* **237**, 37–72. (doi:10.1098/rstb.1952.0012)
45. Maini P, Painter K, Chau HP. 1997 Spatial pattern formation in chemical and biological systems. *J. Chem. Soc.* **93**, 3601–3610. (doi:10.1039/a702602a)
46. Marciniak-Czochra A, Karch G, Suzuki K. 2017 Instability of Turing patterns in reaction-diffusion-ODE systems. *J. Math. Biol.* **74**, 583–618. (doi:10.1007/s00285-016-1035-z)
47. Armstrong NJ, Painter KJ, Sherratt JA. 2006 A continuum approach to modelling cell–cell adhesion. *J. Theor. Biol.* **243**, 98–113. (doi:10.1016/j.jtbi.2006.05.030)
48. Murakawa H, Togashi H. 2015 Continuous models for cell–cell adhesion. *J. Theor. Biol.* **374**, 1–12. (doi:10.1016/j.jtbi.2015.03.002)
49. Carrillo JA, Huang Y, Schmidtchen M. 2018 Zoology of a nonlocal cross-diffusion model for two species. *SIAM J. Appl. Math.* **78**, 1078–1104. (doi:10.1137/17M1128782)
50. Carrillo JA, Colombi A, Scianna M. 2018 Adhesion and volume constraints via nonlocal interactions determine cell organisation and migration profiles. *J. Theor. Biol.* **445**, 75–91. (doi:10.1016/j.jtbi.2018.02.022)
51. Gaffney EA, Seirin Lee S. 2015 The sensitivity of Turing self-organization to biological feedback delays: 2D models of fish pigmentation. *Math. Med. Biol.* **32**, 57–79. (doi:10.1093/mamm/bdq017)
52. Konow C, Li Z, Shepherd S, Bullara D, Epstein IR. 2021 Influence of survival, promotion, and growth on pattern formation in zebrafish skin. *Sci. Rep.* **11**, 9864. (doi:10.1038/s41598-021-89116-4)
53. Bloomfield JM, Painter KJ, Sherratt JA. 2011 How does cellular contact affect differentiation mediated pattern formation? *Bull. Math. Biol.* **73**, 1529–1558. (doi:10.1007/s11538-010-9578-4)
54. Kondo S. 2017 An updated kernel-based Turing model for studying the mechanisms of biological pattern formation. *J. Theor. Biol.* **414**, 120–127. (doi:10.1016/j.jtbi.2016.11.003)
55. Painter KJ, Bloomfield JM, Sherratt JA, Gerisch A. 2015 A nonlocal model for contact attraction and repulsion in heterogeneous cell populations. *Bull. Math. Biol.* **77**, 1132–1165. (doi:10.1007/s11538-015-0080-x)
56. Giacomin G, Lebowitz JL. 1997 Phase segregation dynamics in particle systems with long range interactions I. Macroscopic limits. *J. Stat. Phys.* **87**, 37–61. (doi:10.1007/BF02181479)
57. Painter KJ, Hillen T. 2002 Volume-filling and quorum-sensing in models for chemosensitive movement. *Can. Appl. Math. Quart.* **10**, 501–543.
58. Penington CJ, Hughes BD, Landman KA. 2011 Building macroscale models from microscale probabilistic models: a general probabilistic approach for nonlinear diffusion and multispecies phenomena. *Phys. Rev. E* **84**, 041120. (doi:10.1103/PhysRevE.84.041120)
59. Hillen T, Painter KJ. 2013 Transport and anisotropic diffusion models for movement in oriented habitats. In *Dispersal, individual movement and spatial ecology* (eds M Lewis, P Maini, S Petrowskii), pp. 177–222. Berlin, Germany: Springer. (doi:10.1007/978-3-642-35497-7_7)

60. Carrillo JA, Choi YP, Hauray M. 2014 The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds: an excursion through modeling, analysis and simulation* (eds A Muntean, F Toschi), pp. 1–46. Vienna, Austria: Springer. (doi:10.1007/978-3-7091-1785-9_1)
61. Chen Y, Kolokolnikov T. 2014 A minimal model of predator–swarm interactions. *J. R. Soc. Interface* **11**, 20131208. (doi:10.1098/rsif.2013.1208)
62. Bruna M, Burger M, Pietschmann JF, Wolfram MT. 2022 Active crowds. In *Active particles*, vol. 3 (eds N Bellomo, JA Carrillo, E Tadmor), pp. 35–73. Cham, Switzerland: Birkhäuser. (doi:10.1007/978-3-030-93302-9_2)
63. Champagnat N, Ferrière R, Méléard S. 2006 Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theor. Popul. Biol.* **69**, 297–321. (doi:10.1016/j.tpb.2005.10.004)
64. Champagnat N, Ferrière R, Méléard S. 2008 Individual-based probabilistic models of adaptive evolution and various scaling approximations. *Prog. Probab.* **59**, 75. (doi:10.1007/978-3-7643-8458-6_6)
65. Champagnat N, Ferrière R, Méléard S. 2008 From individual stochastic processes to macroscopic models in adaptive evolution. *Stoch. Models* **24**, 2–44. (doi:10.1080/15326340802437710)
66. Hansen JP, McDonald IR. 2006 *Theory of simple liquids*. London, UK: Academic Press.
67. Lushnikov PM, Chen N, Alber M. 2008 Macroscopic dynamics of biological cells interacting via chemotaxis and direct contact. *Phys. Rev. E* **78**, 061904. (doi:10.1103/PhysRevE.78.061904)
68. Simpson MJ, Landman KA, Hughes BD. 2009 Multi-species simple exclusion processes. *Phys. A: Stat. Mech.* **388**, 399–406. (doi:10.1016/j.physa.2008.10.038)
69. Simpson MJ, Baker RE. 2011 Corrected mean-field models for spatially dependent advection-diffusion-reaction phenomena. *Phys. Rev. E* **83**, 51922. (doi:10.1103/PhysRevE.83.051922)
70. Markham DC, Simpson MJ, Baker RE. 2013 Simplified method for including spatial correlations in mean-field approximations. *Phys. Rev. E* **87**, 62702. (doi:10.1103/PhysRevE.87.062702)
71. Wiecek R. 2023 Hydrodynamic limit of a stochastic model of proliferating cells with chemotaxis. *Kinet. Relat. Models* **16**, 373–393. (doi:10.3934/krm.2022032)
72. Binny RN, Plank MJ, James A. 2015 Spatial moment dynamics for collective cell movement incorporating a neighbour-dependent directional bias. *J. R. Soc. Interface* **12**, 20150228. (doi:10.1098/rsif.2015.0228)
73. Binny RN, James A, Plank MJ. 2016 Collective cell behaviour with neighbour-dependent proliferation, death and directional bias. *Bull. Math. Biol.* **78**, 2277–2301. (doi:10.1007/s11538-016-0222-9)
74. Middleton AM, Fleck C, Grima R. 2014 A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. *J. Theor. Biol.* **359**, 220–232. (doi:10.1016/j.jtbi.2014.06.011)
75. Bruna M, Chapman SJ, Robinson M. 2017 Diffusion of particles with short-range interactions. *SIAM J. Appl. Math.* **77**, 2294–2316. (doi:10.1137/17M1118543)
76. Johnston ST, Baker RE, Simpson MJ. 2017 A new and accurate continuum description of moving fronts. *New J. Phys.* **19**, 033010. (doi:10.1088/1367-2630/aa5bf5)
77. Bruna M, Chapman SJ. 2012 Excluded-volume effects in the diffusion of hard spheres. *Phys. Rev. E* **85**, 011103. (doi:10.1103/PhysRevE.85.011103)
78. Takahashi G, Kondo S. 2008 Melanophores in the stripes of adult zebrafish do not have the nature to gather, but disperse when they have the space to move. *Pigment Cell Melanoma Res.* **21**, 677–686. (doi:10.1111/j.1755-148X.2008.00504.x)
79. Parichy DM, Turner JM. 2003 Zebrafish *puma* mutant decouples pigment pattern and somatic metamorphosis. *Dev. Biol.* **256**, 242–257. (doi:10.1016/S0012-1606(03)00015-0)
80. Patterson LB, Bain EJ, Parichy DM. 2014 Pigment cell interactions and differential xanthophore recruitment underlying zebrafish stripe reiteration and *Danio* pattern evolution. *Nat. Commun.* **5**, 5299. (doi:10.1038/ncomms6299)
81. Walderich B, Singh AP, Mahalwar P, Nüsslein-Volhard C. 2016 Homotypic cell competition regulates proliferation and tiling of zebrafish pigment cells during colour pattern formation. *Nat. Commun.* **7**, 11462. (doi:10.1038/ncomms11462)
82. Golse F. 2003 The mean-field limit for the dynamics of large particle systems. *Journées Équations aux dérivées Partielles*.
83. Di Francesco M, Fagioli S. 2013 Measure solutions for non-local interaction PDEs with two species. *Nonlinearity* **26**, 2777. (doi:10.1088/0951-7715/26/10/2777)
84. Chiari G, Delitala ME, Morselli D, Scianna M. 2022 A hybrid modeling environment to describe aggregates of cells heterogeneous for genotype and behavior with possible phenotypic transitions. *Int. J. Non-Linear Mech.* **144**, 104063. (doi:10.1016/j.ijnonlinmec.2022.104063)
85. Brunet E, Derrida B. 1997 Shift in the velocity of a front due to a cutoff. *Phys. Rev. E* **56**, 2597–2604. (doi:10.1103/PhysRevE.56.2597)
86. Sontag A, Rogers T, Yates CA. 2023 Stochastic drift in discrete waves of nonlocally interacting particles. *Phys. Rev. E* **107**, 014128. (doi:10.1103/PhysRevE.107.014128)
87. Bull JA, Byrne HM. 2022 The hallmarks of mathematical oncology. *Proc. IEEE* **110**, 523–540. (doi:10.1109/JPROC.2021.3136715)
88. Bruna M, Chapman SJ, Robinson M. 2017 Diffusion of particles with short-range interactions. *SIAM J. Appl. Math.* **77**, 2294–2316. (doi:10.1137/17M1118543)
89. Franke F, Aland S, Böhme HJ, Voss-Böhme A, Lange S. 2022 Is cell segregation like oil and water: asymptotic versus transitory regime. *PLoS Comput. Biol.* **18**, e1010460. (doi:10.1371/journal.pcbi.1010460)
90. Nardini JT, Baker RE, Simpson MJ, Flores KB. 2021 Learning differential equation models from stochastic agent-based model simulations. *J. R. Soc. Interface* **18**, 20200987. (doi:10.1098/rsif.2020.0987)
91. Bull JA, Byrne HM. 2023 Quantification of spatial and phenotypic heterogeneity in an agent-based model of tumour-macrophage interactions. *PLoS Comput. Biol.* **19**, e1010994. (doi:10.1371/journal.pcbi.1010994)
92. Miyazawa S. 2020 Pattern blending enriches the diversity of animal colorations. *Sci. Adv.* **6**, eabb9107. (doi:10.1126/sciadv.abb9107)
93. McGuirl MR, Volkening A, Sandstede B. 2020 Topological data analysis of zebrafish patterns. *Proc. Natl Acad. Sci. USA* **117**, 5113–5124. (doi:10.1073/pnas.1917763117)
94. Martinson WD. 2024 wdmartinson/Self-Organization-One-Species: Self-Organization-One-Species-Code-for-Article. zenodo. (doi:10.5281/zenodo.11160930)
95. William M, Alexandria V, Markus S, Chandrasekhar V, José C. 2024 Data from: Linking continuous and discrete models of cell birth and migration. Dryad Digital Repository. (doi:10.5061/dryad.s4mw6m9cb)
96. Martinson WD, Volkening A, Schmidtchen M, Venkataraman C, Carrillo JA. 2024 Linking discrete and continuous models of cell birth and migration. Figshare. (doi:10.6084/m9.figshare.c.7303179)