Connecting gene expression to cellular movement: A transport model for cell migration

Alexis Grau Ribes ⬤,[*] Yannick De Decker, and Laurence Rongy

Nonlinear Physical Chemistry Unit, Faculté des Sciences, Université libre de Bruxelles (ULB), Brussels, Belgium

The adhesion properties and the mobility of biological cells play key roles in the propagation of cancer. These properties are expected to depend on intracellular processes and on the concentrations of chemicals inside the cell. While most existing reaction-diffusion models for cell migration consider that cell mobility and proliferation rate are constant or depend on an external diffusing species, they do not include the gene expression dynamics taking place in moving cells that affect cellular transport. In this work, we propose a multiscale model where mobility and proliferation depend explicitly on the cell's internal state. We focus more specifically on the case of cellular mobility in epithelial tissues. Wound-healing experiments have demonstrated that the loss of a key protein, E-cadherin, results in a significant increase in both mobility and invasiveness of epithelial cells, with dramatic consequences on cancer progression. We can reproduce the results of these experiments under various genetic conditions with a single set of parameters.

I. INTRODUCTION

Detachment and dissemination of cancer cells from primary tumors are crucial steps of metastatic colony and malignant progression [1,2]. The cell-biological program governing these steps is called epithelial-to-mesenchymal transition (EMT) [3,4]. During EMT, adhesive cells lose contact and acquire migratory and invasive abilities [5]. The loss of cell adhesion is due to the downregulation of the transmembrane protein E-cadherin occurring during the transition. This protein mediates cell contact and its loss is associated with tumor invasiveness [6–8]. The EMT is orchestrated by several transcription factors interacting with each other in a complex gene regulation network [9,10]. Among these, SNAIL and ZEB1 directly target CDH1, the gene that encodes E-cadherin, and also repress several other proteins establishing cell-cell junctions. In addition to transcription factors, potent extracellular factors have the ability to initiate the EMT program, among which TGFβ [11], IL-6, which triggers differentiation in inflammatory conditions [12] and exosome-secreted microRNAs (miR-34, miR-21, etc.) [13–16].

In this context, wound-healing assays are simple *in vitro* experiments providing quantitative data on cell migration and proliferation and can be used to characterize the influence of a given gene on these phenomena [17]. The experimental procedure consists in removing a section of a confluent cell monolayer and in observing the dynamics of the regeneration of the sheet of cells. This method was used to demonstrate that the downregulation of E-cadherin could promote malignant progression of prostate and breast cancers [18,19]. We note that cell dissemination is, however, not always accompanied by E-cadherin loss [20].

Reaction-diffusion-based models are often used to describe the evolution of cell density in wound-healing assays. Fisher's equation is, for example, commonly used to capture both individual cell diffusion and proliferation, the two processes being assumed to govern wound closure [21]. Depending on the model, the cellular diffusivity is constant [22–25], function of the cell density [26–29], or depends on an external factor [30,31].

With a constant diffusion coefficient, the solution of Fisher's equation is a traveling wave moving at constant velocity depending on the cell diffusivity. This velocity and data for cell proliferation were used to measure cell diffusivity [22]. Cai *et al.* considered cell contact inhibition using a diffusivity function depending on cell density [26]. Khain *et al.* also used wound-healing data to fit the parameters of their discrete model which considers cell-cell adhesion and the number of nearest neighbors [29]. These three models are reaction-diffusion models in which the only variable is the cell density. More complex models based on evolution equations for the cell density and for an externally diffusing chemical factor produced by the cell were also developed. In those models, the cell proliferation rate and/or the cell diffusivity are assumed to depend linearly on the chemical factor [30,31].

Although the latter models consider an external chemical factor, none of the aforementioned approaches include the gene expression dynamics taking place in cells, which is at the very origin of changes in mobility. In particular, no model is available in which cell diffusivity is function of the E-cadherin level. Here, we present a multiscale modeling approach that combines the description of gene expression dynamics, cellular mobility, and proliferation. We develop a discrete spatial model capable of showing the impact of E-cadherin downregulation on the wound-healing process. Moreover, E-cadherin regulation by a chemical factor is included, as well as the secretion and diffusion of this factor in the extracellular matrix and its capture by adjacent cells. Available wound-healing experimental data quantifying the effect of E-cadherin on cell mobility are used to fit the parameters of the model with unique values, independent of the genetic conditions, and
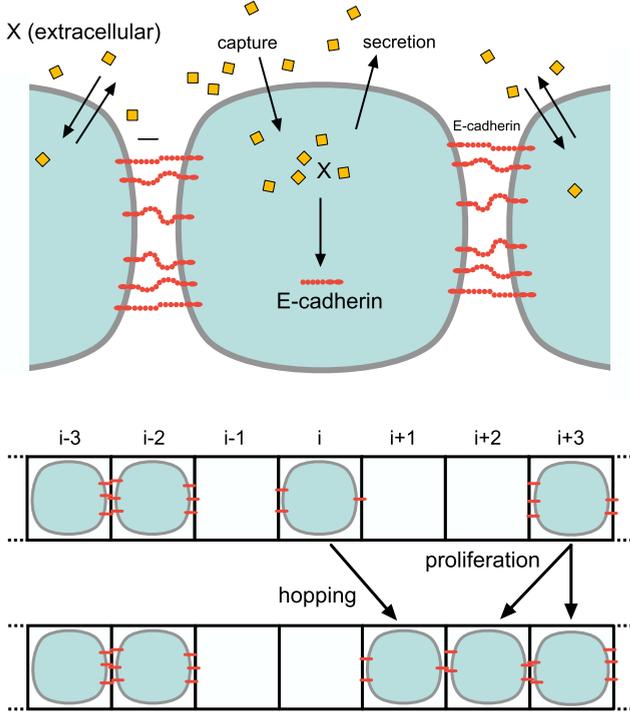
*agraurib@ulb.ac.be

FIG. 1. Schematic of the system. E-cadherin mediates cell-cell adhesion and is positively regulated by an external chemical factor X. Each site of the system can be empty or occupied by a cell. Cell hopping and cell proliferation are the two processes governing cell dynamics. A hopping process and a proliferation process are illustrated here on a 1D lattice.

to link gene expression dynamics to cell migratory abilities. Furthermore, we show how the presented model is capable of taking into account the spatial impact of cell division using a nonlocal proliferation term, the influence of neighboring cells on cell mobility, and finally the effect of E-cadherin on cell proliferation rate.

II. MODEL FORMULATION

Collective cell spreading taking place during wound-healing assays can be described as cells moving on a lattice made of N_0 nodes. Each site of the lattice can be occupied by a single cell or either empty and each cell can carry, secrete, or capture some biochemical species. Cell occupation can change over time either by diffusive or proliferation processes (see Fig. 1). A cell can hop from site i to a neighboring site i' if site i is occupied and site i' is empty. After the transition, site i is empty and site i' is occupied. The transition frequency associated to this process is expected to depend on the level of E-cadherin in the cell. For a cell to proliferate, the same condition of an occupied site i next to an empty site i' has to be fulfilled. After this process, both sites i and i' are occupied. E-cadherin, the protein mediating cell adhesion, is produced and degraded in each cell. Its synthesis is activated by a regulating factor also produced and degraded in the cell. This regulating factor is secreted in the extracellular matrix where it can be transported to other cells. This species could be an

exosome-secreted microRNA intervening in E-cadherin gene expression network [16,32–34].

The aforementioned processes are discrete events that occur with specific probabilities. The most natural level of description of such a system is thus stochastic. From the master equation ruling the evolution of the underlying probability distribution, we obtain a mean-field evolution equation for the mean cell occupation of node i, n_i, comprised between 0 and 1 as well as for the mean concentrations per site of E-cadherin, E_i, of the regulating factor in the cell X_i and in the extracellular matrix X_i^{ext} :

$$\frac{dn_i}{dt} = \frac{1}{2}[\Gamma_{i+1}\, n_{i+1}\, (1-n_i) + \Gamma_{i-1}\, n_{i-1}\, (1-n_i)$$
$$- \Gamma_i\, n_i\, (1-n_{i+1}) - \Gamma_i\, n_i\, (1-n_{i-1})$$
$$+ \Theta_{i+1}\, n_{i+1}\, (1-n_i) + \Theta_{i-1}\, n_{i-1}\, (1-n_i)], \quad (1)$$

$$\frac{dE_i}{dt} = n_i v_E\, \frac{X_i^0}{K + X_i^0} - k_E\, E_i$$
$$+ \frac{1}{2}[\Gamma_{i+1}\, n_{i+1}\, (1-n_i)E_{i+1}$$
$$+ \Gamma_{i-1}\, n_{i-1}\, (1-n_i)E_{i-1}$$
$$- \Gamma_i\, n_i\, (1-n_{i+1})E_i - \Gamma_i\, n_i\, (1-n_{i-1})E_i$$
$$+ \Theta_{i+1}\, n_{i+1}\, (1-n_i)E_{i+1}$$
$$+ \Theta_{i-1}\, n_{i-1}\, (1-n_i)E_{i-1}], \quad (2)$$

$$\frac{dX_i}{dt} = v_X\, n_i - k_X\, X_i - k_{\text{out}}\, X_i + k_{in}\, X_i^{\text{ext}}\, n_i$$
$$+ \frac{1}{2}[\Gamma_{i+1}\, n_{i+1}\, (1-n_i)X_{i+1}$$
$$+ \Gamma_{i-1}\, n_{i-1}\, (1-n_i)X_{i-1}$$
$$- \Gamma_i\, n_i\, (1-n_{i+1})X_i - \Gamma_i\, n_i\, (1-n_{i-1})X_i$$
$$+ \Theta_{i+1}\, n_{i+1}\, (1-n_i)X_{i+1}$$
$$+ \Theta_{i-1}\, n_{i-1}\, (1-n_i)X_{i-1}], \quad (3)$$

and

$$\frac{dX_i^{\text{ext}}}{dt} = k_{\text{out}}\, X_i - k_{in}\, X_i^{\text{ext}}\, n_i$$
$$+ \Gamma_X\, \left(X_{i+1}^{\text{ext}} + X_{i-1}^{\text{ext}} - 2\, X_i^{\text{ext}}\right). \quad (4)$$

These evolution equations are derived from the microscopic rules governing cell movement (see Appendix A). We discuss here the evolution equations in a one-dimensional (1D) system for clarity but generalization to two-dimensional (2D) lattices is straightforward. The first two terms in square brackets in Eq. (1) represent the hopping processes from neighboring sites to site i, with $n_{i+1}\,(1-n_i)$ and $n_{i-1}\,(1-n_i)$ quantifying the probability to have the required cell configurations for the process to happen. Similarly, the third and fourth terms in square brackets describe the hopping processes from site i to adjacent sites. Γ_i stands for the hopping frequency at site i and is expected to depend on the level of E-cadherin. As illustrated in Fig. 1, E-cadherin is a transmembrane protein responsible for cell adhesion. The hopping frequency must therefore be maximal when a cell does not carry any E-cadherin and decrease with increasing

E-cadherin concentration. We choose to write the hopping frequency as

$$\Gamma_i = \Gamma_0 \, e^{-\alpha E_i^0}, \tag{5}$$

with Γ_0 being the hopping frequency of a cell with no E-cadherin, α an adhesion parameter, and $E_i^0 = E_i/n_i$ the concentration of E-cadherin per cell. The last two terms in square brackets of Eq. (1) account for the proliferation of neighboring cells. We first make the assumption that the proliferation frequency Θ_i remains constant ($\Theta_i = \Theta_0$) and will discuss this hypothesis at the end of this article.

The first two terms of Eq. (2) represent the synthesis and degradation of E-cadherin, with v_E the maximum synthesis rate of E-cadherin and k_E its rate constant for degradation. The activation of E-cadherin synthesis by the regulating factor is assumed to obey Michaelis-Menten kinetics [35] with the corresponding constant K and depends on the concentration of the regulating factor per cell $X_i^0 = X_i/n_i$. The level of E-cadherin can also vary due to changes in cell occupation [third term of Eq. (2)]. For instance, the term $\Gamma_{i+1} \, n_{i+1} \, (1 - n_i) \, E_{i+1}$ represents the variation of E-cadherin level at site i, due to the hopping of a cell located in $i+1$ and whose E-cadherin concentration is E_{i+1}. Similarly, the term $\Theta_{i+1} \, n_{i+1} \, (1 - n_i) \, E_{i+1}$ represents how E-cadherin concentration at site i increases when a cell located on site $i+1$ gives birth to a cell at position i.

The regulating factor is expressed and degraded in each cell, v_X being its synthesis rate and k_X its degradation constant. The rate constants for the secretion and capture of the regulating factor by the cell are k_{out} and k_{in}, respectively [see Eqs. (3) and (4)]. Finally, Γ_X is the hopping frequency of the extracellular species from one site to one of its first neighbors and is considered constant. Since microRNAs do not undergo degradation in exosomes, no degradation term is considered for the extracellular factor [36]. As for E-cadherin concentration, the amount of regulating factor inside the cell can change because of cell diffusion and cell proliferation [fifth term of Eq. (3)].

III. PROLIFERATION-INDUCED CELL PROPAGATION

In order to compare with existing models, we consider the continuous space limit of Eq. (1). We introduce the space coordinate $x = ia$ where a is the distance between two first neighbors. Equation (1) becomes (see Appendix B)

$$\frac{\partial n}{\partial t} = -\frac{\partial}{\partial x}\left[-D\frac{\partial n}{\partial x} - n(1-n)\frac{\partial D}{\partial x} \right]$$
$$+ \, \Theta \, n \, (1-n) + R \, (1-n)\frac{\partial^2 n}{\partial x^2}, \tag{6}$$

where $D = \frac{\Gamma a^2}{2}$ is a diffusion coefficient, Γ is the local E-cadherin-dependent hopping frequency [Eq. (5)], and $R = \frac{\Theta a^2}{2}$ is a proliferation-induced transport coefficient. We obtain a modified Fisher's equation where the first term is a nonideal diffusion term with an E-cadherin-dependent diffusion coefficient, the second term is a logistic growth term for the local cell proliferation, and the third term arises from the nonlocal character of cell proliferation. Indeed, cell proliferation leads
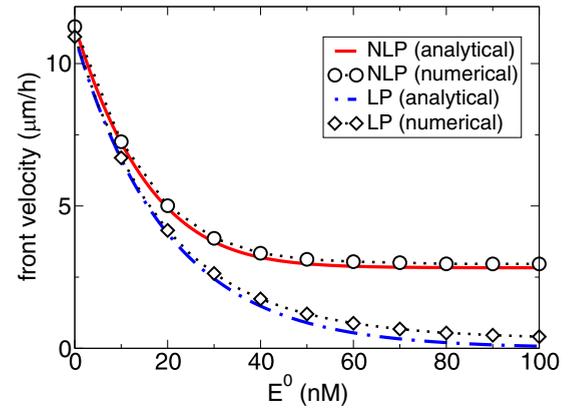

FIG. 2. Velocity of the front as a function of cellular E-cadherin concentration per cell E^0, obtained numerically (markers) and with analytical predictions (color lines) for nonlocal (NLP) and local proliferation (LP) as given by $v_{\min}^{\mathrm{NLP}}$ and $v_{\min}^{\mathrm{LP}}$, respectively. E-cadherin concentration is kept constant in space and time. A Heaviside function is used as initial condition, $a = 20 \ \mu\mathrm{m}$, $\Gamma_0 = 1.5 \ \mathrm{h}^{-1}$, $\alpha = 0.1 \ \mathrm{nM}^{-1}$, $\Theta = 0.1 \ \mathrm{h}^{-1}$ with converged time step $dt = 10^{-3}$ h.

to the creation of a new cell next to the mother cell which is not a local event.

Similarly to Fisher's equation, our model predicts traveling waves of cell density. The front velocity was obtained after numerical integration of Eq. (6) using finite differences method, Euler's explicit time integration scheme, and no flux boundary conditions. We have studied the influence of E-cadherin concentration on the front velocity, considering E_i^0 to be constant in space and time. We observe proliferation-induced front propagation even for high E-cadherin concentration and thus very low cell mobility, highlighting the importance of cell proliferation as an additional transport process (see Fig. 2). Indeed, cell proliferation is modeled in this work with nonlocal terms [see fifth and sixth terms in the square brackets of Eq. (1)]. The mechanism of this nonlocal proliferation is schematized in Fig. 1. For a cell in site i to proliferate, one of the neighboring sites must be empty. Proliferation leads to the propagation of cell population in space, even in the absence of mobility. It is thus not surprising that, using this approach for modeling cell proliferation, we can observe front propagation even when the cell hopping frequency is zero. Similar process-induced propagation has been observed in the case of nonlocal chemical reactions [37]. Following the framework used in the latter study, the minimal velocity of the traveling wave solutions of Eq. (6) can be obtained for E_i, X_i, and X_i^{ext} constant and reads as $v_{\min}^{\mathrm{NLP}} = 2\sqrt{\Theta \, (D + R)}$ where the superscript NLP stands for nonlocal proliferation.

The last term in the right-hand side of Eq. (6) does not appear in previous reaction-diffusion models for cell mobility where proliferation is seen as a local event. Those models differ from one another only in the form of the cell diffusivity (see Table I) and predict a traveling wave with constant velocity satisfying $v_{\min}^{\mathrm{LP}} = 2\sqrt{D\Theta}$ where the superscript LP stands for local proliferation. We can reproduce those results by integrating Eq. (6) without the last term. In that case, the front velocity goes to zero at high E-cadherin concentration ($D \to 0$) and proliferation-induced front propagation cannot be observed (Fig. 2).

TABLE I. Cell diffusivity in existing reaction-diffusion models for cell migration.

Fisher's equation	Reference
$\frac{\partial n}{\partial t} = -\frac{\partial}{\partial x}[-D(\cdot)\frac{\partial n}{\partial x}] + \Theta\, n\,(1-n)$	[21]
Constant diffusivity	
$D(\cdot) = D$	[22–25]
Contact inhibition models	
$D(\cdot) = D_0 A/(A + n)$	[26–28]
$D(\cdot) = (1-q)^{4n}[1 + 4n(1-n)\ln(1-q)]/4$	[29]
where A and q are positive constants	

IV. WOUND-HEALING CELL MIGRATION ASSAY

We focus now on the modeling of wound-healing assays [see Fig. 3(a)]. Considering the symmetry of such systems, we use a 1D lattice with no flux boundary conditions. The size of the system (150 sites) is sufficiently large to avoid boundary effect on wound-healing dynamics. We consider the distance between two neighboring sites to be $a = 20\ \mu$m, a typical diameter for an epithelial cell [38], and a wound of 16 sites (320 μm) at the beginning of the simulation. Initial conditions for cell occupation are taken

to be $n = 1$ outside the wound region and $n = 10^{-10}$ in the wound, a value sufficiently small not to affect the healing dynamics on the timescales studied. Equations (1)–(4) are integrated numerically using an explicit Euler method with time step $dt = 10^{-3}$ h. Initial concentration for E_i, X_i, and X_i^{ext} are the steady state values of the homogeneous system.

Comparing numerical results with experimental wound-healing data allows us to estimate the value of the proliferation rate Θ_0, the hopping rate of a cell with no E-cadherin Γ_0, and the adhesion parameter α. All other parameter values are known from previous studies (Table II). The value of the rate constant for E-cadherin degradation was found in the literature [39]. Its synthesis rate and the Michaelis-Menten constant for the activation of E-cadherin by the regulating factor were calibrated using CDH1 expression data in prostate cancer tissue in the OASIS genomics database [40]. We averaged CDH1 RNA-seq data for prostate adenocarcinoma tissues of >300 samples normalized with β-actin level. RNA/protein ratio provided by Schwanhäusser *et al.* was used to estimated E-cadherin concentration (~20 nM) and K was taken of the same order of magnitude as X concentration (1 nM) [39]. v_E was adjusted accordingly to fit the expression data. For the degradation constant and the synthesis rate of the regulating
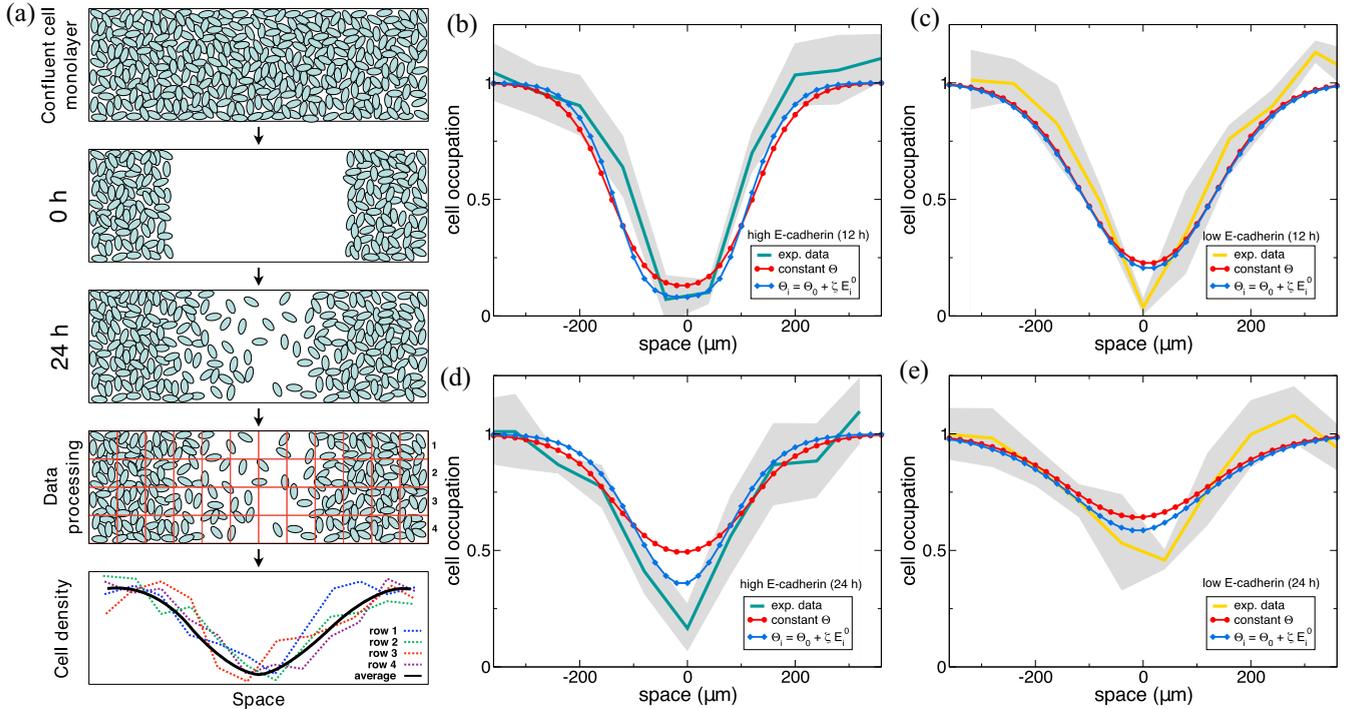

FIG. 3. Comparison between experimental and numerical profiles. (a) Schematic of the wound-healing cell migration assay and data processing. A section of a confluent cell monolayer is removed and the healing process due to cell diffusion and proliferation is monitored. Images of the system are captured at 0, 12, and 24 h. Cell counting is used to extract cell density profiles from the images. On each image of Fig. 2 from Fan *et al.* [18], we drew a square grid (80 μm × 80 μm) and cells were counted in each square. Averaging over the six rows of the grid and normalizing with respect to the cell occupation at the boundaries, we obtained the 1D cell occupation profiles. (b)–(e) Experimental cell occupations are obtained after wound-healing cell migration assay (plain lines and 95% confidence interval reconstructed from Fan *et al.* [18]). The confidence intervals are due to the data processing procedure. Numerical cell occupation profiles are obtained with best fitting parameter values using the E-cadherin-dependent diffusivity and constant proliferation model (red circles) and E-cadherin-dependent proliferation rate model (blue squares). Fitted values of Γ_0, α, Θ_0, and ζ are shown in Table IV and all other parameter values are shown in Table II. Apart from the synthesis rate of E-cadherin varied to mimic the experiments, all other parameters are the same in the four panels.

TABLE II. Parameter symbols, definitions, units, and default values used in the simulations.

a	Distance between first neighbors	20	μm	[38]
Γ_0	Hopping frequency of a cell with no E-cadherin	2.5	h^{-1}	[26]
α	Adhesion parameter	0.03	nM^{-1}	Estimated
Θ_0	Proliferation rate	0.1	h^{-1}	[26]
v_E	Synthesis rate of E-cadherin	7	$nM\ h^{-1}$	[40]
K	Michaelis-Menten constant for E-cadherin synthesis activation	1	nM	[40]
k_E	Rate constant for E-cadherin degradation rate	0.14	h^{-1}	[39]
v_X	Synthesis rate of the regulating factor X	0.04	$nM\ h^{-1}$	[41]
k_X	Rate constant for the regulating factor degradation	0.02	h^{-1}	[41]
k_{out}	Rate constant for the secretion of the regulating factor	0.02	h^{-1}	Estimated
k_{in}	Rate constant for the uptake of the extracellular regulating factor X^{ext}	0.1	h^{-1}	Estimated
Γ_X	Hopping frequency of the extracellular regulating factor X^{ext}	1	h^{-1}	[43]

factor, we took those of abundant circulating microRNAs, such as miR-21 [41]. Typical values of degradation rates for miRNAs are between 0.01 and 0.04 h^{-1} and we chose here $k_X = 0.02\ h^{-1}$. v_X was estimated to obtain 2 nM for the basal concentration of X, corresponding to ~2000 copies per cell [42]. The value of the diffusion coefficient of the extracellular factor is that of exosomes in the extracellular matrix [43].

Fan *et al.* performed wound-healing assays to quantify the influence of E-cadherin knockdown on cell migration [18]. Small interfering RNAs were used to downregulate E-cadherin expression. Images of wound-healing assays for untransfected cells (high E-cadherin expression) and transfected cells (low E-cadherin expression) demonstrated that E-cadherin silencing causes an increase of cell mobility. Their experimental images taken at 12 and 24 h were processed in this study for comparison with our numerical results [see Fig. 3(a)].

Transfected cells show an expression of E-cadherin three times smaller than untransfected ones. To mimic this, we set the synthesis rate of E-cadherin for transfected cells at the third of the basal synthesis rate. We then simulate the evolution of the cell occupation for high and low E-cadherin expression and compare the numerical profiles obtained at 12 and 24 h with the corresponding experimental profiles. The goodness of fit between experimental and numerical profiles is measured by

$$\chi^2 = \sum_{j=1}^{4} \sum_{l=1}^{12} \frac{(n_{l,\text{expt}} - n_{l,\text{num}})^2}{\sigma_{l,\text{expt}}^2}, \qquad (7)$$

where l refers to the data points along the x axis, j to the experimental parameters (12 or 24 h, high or low E-cadherin expression), and $\sigma_{l,\text{expt}}$ is the standard deviation associated to each experimental data point. Minimizing χ^2 gives the best fitting parameter set for the four experimental conditions simultaneously [Figs. 3(b)–3(e)]. We use a stochastic optimization method to obtain the best parameter set [44]. It consists of a large number of trials using randomly generated parameter vectors ($\Gamma_0, \alpha, \Theta_0$). Each parameter value is picked randomly with uniform probability distribution, between 0 and a maximum value above realistic values (see Table III). The spatial profiles of cell occupation obtained with the best parameter set (see Table IV) together with the experimental

profiles obtained after processing Fan *et al.*'s images are shown in Figs. 3(b)–3(e) (see red curves).

The numerical simulations both at high and low E-cadherin expression capture the main features of the experimental data. In particular, the model is capable of reproducing the effect of E-cadherin knockdown on cell mobility. Indeed, the four numerical profiles are obtained with the same values of Γ_0, α, and Θ_0. The dynamical changes are only due to the variation of the synthesis rate of E-cadherin, which corresponds to the procedure of the wound-healing assay.

In order to reproduce such wound-healing data with the phenomenological models currently available, in which cell diffusivity does not depend on an adhesion marker, one needs different values of *ad hoc* cell transport coefficient and proliferation rate. On the contrary, the approach proposed here relies on a unique set of parameter values at both high and low E-cadherin expression. Since wound-healing data are scarce and the variability between different tissues can be important, our model can be used to predict the influence of each parameter on wound-healing dynamics.

A. Controlling wound-healing dynamics

For the parameter values listed in Table II, partial wound-healing is observed after 24 h [see Fig. 4(a)]. The area beneath the cell occupation profile in the initial wound region is defined as the healed area [see gray area in Fig. 4(a)]. We now look at the time evolution of this quantity for various values of the major parameters of the model. Among those parameters, the proliferation rate is essential since it induces front propagation and thus wound healing. In the absence

TABLE III. Estimation of the parameter values of the model comparing experimental and numerical cell occupation profiles: range of variation of the parameter values for the stochastic optimization method. β quantifies the strength of the interaction between two molecules of E-cadherin and ζ, the effect of E-cadherin on cell proliferation. Those parameters will be introduced later.

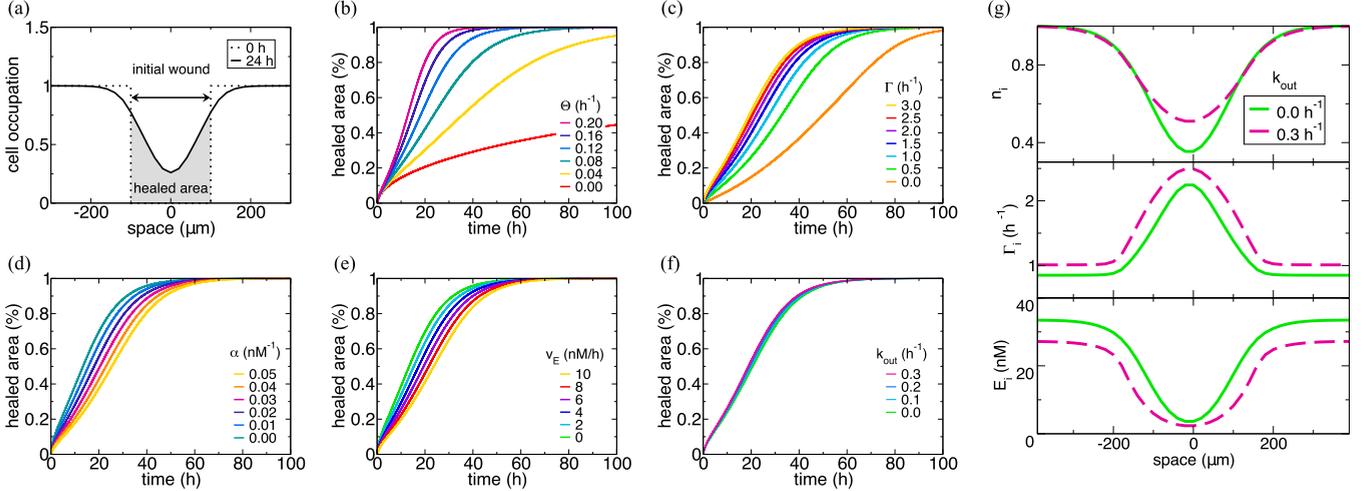|		
Γ_0	0–4	h^{-1}
α	0–0.2	nM^{-1}
β	0–0.2	nM^{-2}
Θ_0	0–0.15	h^{-1}
ζ	−0.01–0.01	nM^{-1}

FIG. 4. Controlling wound-healing dynamics. (a) Cell occupation at 0 and 24 h for the values of the parameters listed in Table II. The healed area is defined as the area beneath the cell occupation profile in the initial wound region (gray area). Time evolution of the healed area for increasing values of (b) the cell proliferation rate Θ_0 (bottom to top), (c) the hopping frequency of a cell with no E-cadherin Γ_0 (right to left), (d) the adhesion parameter α (left to right), (e) the synthesis rate of E-cadherin v_E (left to right), and (f) the rate constant for the secretion of the regulating factor k_{out} (right to left). (g) Cell occupation, hopping frequency, and cellular E-cadherin concentration profiles after 30 h for two different values of the rate of secretion k_{out}. When not varied, the parameter values are as shown in Table II with a constant proliferation rate.

of cell proliferation, healing is slow and cannot be achieved totally [see red line in Fig. 4(b)]. Indeed, cell occupation is progressively homogenized by diffusive processes while the total cell occupation remains constant. Varying the proliferation rate in a physiologically relevant range of values [23,24,26,45] shows that the healing rate increases sharply when increasing this parameter. This result is in agreement with existing experiments showing the role of cell proliferation in the restoration of the initial cell density in wound-healing assays [46].

Two parameters are involved in the expression of the local hopping frequency Γ_i: the hopping frequency of a cell with no E-cadherin (Γ_0) and an adhesion parameter (α). Healing is induced by nonlocal proliferation in absence of diffusion ($\Gamma_0 = 0$) and is accelerated when the hopping frequency of an isolated cell increases [Fig. 4(c)]. On the other hand, the rate of healing is notably faster in the absence of interaction

between E-cadherin molecules ($\alpha = 0$) and slows down with increasing α [Fig. 4(d)].

Wound-healing assay is often used to understand the influence of a given gene on cell mobility by modifying its expression [18,45,47–49]. In the same spirit, we have checked how modifications of the synthesis rate of E-cadherin affect the healing rate [Fig. 4(e)]. The absence of E-cadherin synthesis in the cell ($v_E = 0$) causes complete loss of cell-cell contact, the hopping frequency is maximal ($\Gamma_i = \Gamma_0 \; \forall \; i$) and the healing is fast.

As mentioned earlier, the expression, degradation, and regulating function of the chemical factor X take place in the cell. The effect of its synthesis rate on healing dynamics is very similar to that of E-cadherin. This molecule is also an intercellular communication agent since it is secreted and captured by cells and diffuses in the extracellular matrix. Although there is a direct connection between the concentration of the regulating factor and E-cadherin expression, changing the secretion constant (k_{out}) has no significant effect on the rate of change of the healed area [Fig. 4(f)] but affects the shape of the cell occupation profile [see Fig. 4(g)]. Indeed, in the absence of secretion ($k_{out} = 0$), the concentration of the regulating factor and E-cadherin concentration are both high in all cells and thus Γ_i is low [see the green plain lines in Fig. 4(g)]. When the regulating factor is secreted ($k_{out} > 0$), E-cadherin concentration decreases and Γ_i increases everywhere [see pink dashed lines in Fig. 4(g)]. The modification of the spatial profile of the hopping frequency affects cell occupation profile, which shows the role of exosome secretion in modifying the microenvironment of the cells [50]. We note that the capture parameter k_{in} has the exact opposite effect on wound healing. The hopping frequency of the extracellular factor Γ_X was not varied because it represents the mobility of exosomes-secreted species, which is not expected to vary significantly [43].

TABLE IV. Best fitting parameter set for the constant proliferation rate model, the neighbor-dependent hopping frequency model, and the E-cadherin-dependent proliferation rate model obtained with the stochastic optimization method.

Constant proliferation rate	
Γ_0	2.60 h^{-1}
α	0.0294 nM^{-1}
Θ_0	0.893 h^{-1}
χ^2	0.385
E-cadherin-dependent proliferation rate	
Γ_0	3.42 h^{-1}
α	0.0400 nM^{-1}
Θ_0	0.0631 h^{-1}
ζ	0.00180 nM^{-1}
χ^2	0.361

B. Neighbors-dependent hopping process

In Sec. II, we discussed the form of the hopping frequency Γ_i and its dependence on protein expression. Choosing Γ_i to depend on each cell's E-cadherin content only is sufficient to reproduce the wound-healing dynamics observed experimentally. However, an E-cadherin protein from a given cell mediates cell contact via homophilic interactions with another E-cadherin protein from an adjacent cell [6]. Cell mobility could thus also depend on the E-cadherin level of adjacent cells. The presented model can be extended to include interactions between E-cadherin proteins of neighboring cells in the hopping frequency. The probability of a cell hopping to the left is affected by the presence of a cell containing E-cadherin on its right and, equivalently, a hopping process to the right is affected by a cell on the left. We thus now write the hopping frequency to the left and to the right as

$$\Gamma_i^{\text{left}} = \Gamma_0 \, e^{-\alpha E_i^0}\big[1 + n_{i+1}(e^{-\beta E_i^0 E_{i+1}^0} - 1)\big] \tag{8}$$

and

$$\Gamma_i^{\text{right}} = \Gamma_0 \, e^{-\alpha E_i^0}\big[1 + n_{i-1}(e^{-\beta E_i^0 E_{i-1}^0} - 1)\big], \tag{9}$$

where α has the same definition as before and reflects the action of E-cadherin on the mobility of an isolated cell and β reflects the influence of the interaction between two molecules of E-cadherin. The details of the derivation leading to these expressions and the corresponding evolution equations are shown in Appendix C. The stochastic optimization method is used to estimate the values of Γ_0, α, β, and Θ_0 from experimental data. The range of values in which each parameter was varied is shown in Table III. We obtain that none of the parameter sets tested can give a better fit than the parameter values shown in Table IV (with constant proliferation rate) with $\beta = 0$. In summary, having a hopping frequency depending on each cell's content in E-cadherin is sufficient to reproduce the main dynamical features of the wound-healing experiments, and a more complex expression for the hopping frequency based on neighbor interactions cannot improve the fitting for the available experimental data. Such a model might, however, be crucial for explaining other experimental data. This result supports the fact that E-cadherin, in addition to its mechanical effect on cell adhesion, can have a signaling effect on cell polarity and contact inhibition [51].

C. E-cadherin and cell proliferation

While it is established that E-cadherin has a negative effect on cell mobility, its influence on cell proliferation is less clear. Indeed, studies have shown that E-cadherin has no effect on cell proliferation in metastatic prostate cancer cells [7], decreases cell proliferation by blocking G1/S phase in mammary epithelial cells [52], or increases cell proliferation in gastric cancer cells [53]. The modeling framework presented here can be used to investigate this unanswered question. Indeed, we can make the proliferation rate depend on the E-cadherin level and analyze how this change influences the best fitting values for wound-healing experiments. Using a linear dependence for simplicity, we write the proliferation rate as

$$\Theta_i = \Theta_0 + \zeta E_i^0, \tag{10}$$

where ζ reflects the effect of E-cadherin on cell proliferation and can be positive, negative, or equal to zero. The stochastic optimization method is used again to estimate the sign and value of ζ (see the range of variation in Table III). The best fitting parameter set is shown in Table IV and the corresponding numerical profiles of cell occupation are shown in Figs. 3(b)–3(e) (see blue curves). The agreement between the experimental and simulated profiles is qualitatively and quantitatively better with an E-cadherin-dependent cell proliferation than with a constant proliferation rate. The most striking difference is observed at high E-cadherin level after 24 h. The optimal value of ζ is positive, suggesting that E-cadherin increases cell proliferation in the system studied by Fan *et al.*

V. DISCUSSION

Most existing models describing collective cell spreading do not account for protein expression affecting cell mobility and cell proliferation. The models mentioned above that reproduce wound-healing assays are single-variable models for the cell density based on Fisher's equation [21]. With the appropriate set of parameter values, such models can reproduce the results of experiments performed on a given cell line. Nevertheless, in order to reproduce the results of a similar experiment in which the genetic expression of this cell line has been modified, parameter estimation must be performed again, and the best fitting values of the parameters will be different. Here, we have provided a mathematical model that overcomes the need of *ad hoc* parameters by combining the descriptions of cell density evolution, gene expression dynamics of epithelial-to-mesenchymal transition (EMT) markers, and extracellular regulating factor secretion. We have linked cell mobility and EMT in a discrete model through an E-cadherin-dependent hopping frequency of cells. Our model can reproduce the results of experiments performed for cells with different levels of E-cadherin expression, by changing the synthesis rate of E-cadherin only. With this approach, we obtain a single set of optimal values for adhesion, hopping, and proliferation parameters that reproduces experiments done under various genetic conditions.

We included in the model an exosomelike secretion of a regulating species involved in the EMT network. Since secretion rates are highly phenotype dependent [54], we have varied this parameter and showed that exchanges between cells and the extracellular matrix can locally modify gene expression dynamics, thereby affecting cell diffusivity and the spatial distribution of cells. This emphasizes the role of exosomes in rearranging cell microenvironment.

Our model describes cell proliferation as a nonlocal process and we thus observe proliferation-induced front propagation. This suggests that the role played by cell division as a transport process cannot be neglected. Moreover, we propose a model with an E-cadherin-dependent proliferation rate. This model captures more accurately the dynamics of the wound-healing experiments than with a constant proliferation rate and predicts that cell proliferation increases with E-cadherin concentration. This prediction of the model should be tested against dedicated experiments to be validated.

Standard methods for quantifying wound-healing assays involve measuring the distance between the two edges of the

wound every 12 and 24 hours. This method does not give information on the local evolution of the cell density and therefore does not allow estimating the contributions of the different mechanisms governing cell spreading. Cell counting techniques [25] would help establish whether a gene affects cell migration, proliferation, or both processes at the same time. They could also provide the conditions under which proliferation can be described with purely local models or if the nonlocal approach is needed.

Although we remained general and chose a simple gene regulatory network, many possibilities for future investigations arise from this work. An extension of the proposed model could account for the dependence of the proliferation rate on the level of a cell cycle marker. A first step in this direction was made in a recent model that accounts for the dynamics of the cell cycle by differentiating cells with respect to the mitosis phase in which they are [55]. Combining the approach presented here with the latter model would help refine the link between cell cycle and the proliferation-induced cell transport phenomenon. Moreover, specific models can be built upon the one presented here to include species of interest involved upstream of E-cadherin in the complex regulatory network orchestrating the EMT [10,16,47,56,57]. Finally, immune and cancer tissues can show significant differences in cell density and exosomal secretion [58]. The present model can be extended to account for such heterogeneities and understand mechanisms connecting tumor microenvironment and cancer progression.

ACKNOWLEDGMENTS

APPENDIX A: FROM STOCHASTIC PROCESSES TO A MEAN-FIELD MULTISCALE MODEL

In this Appendix, we provide details on how the mean-field evolution equations (1)–(4) can be derived from a discrete stochastic level of description. We restrict ourselves here to one-dimensional systems, for the sake of clarity. Extension to higher dimensional systems is straightforward.

The instantaneous spatial configuration of the system $\boldsymbol{\omega} = \{\omega_i\}$ is defined as the set of local stochastic processes ω_i where $i = 1, 2 \ldots L$ denotes the position in space, with

$$\omega_i = \left(n_i, \tilde{E}_i, \tilde{X}_i, \tilde{X}_i^{\text{ext}} \right). \tag{A1}$$

In the above definition, n_i is a Boolean variable accounting for the occupation of site i by a cell: $n_i = 1$ if there is a cell on this site and is equal to 0 otherwise. $\tilde{E}_i$ is the number of particles of E-cadherin, $\tilde{X}_i$ the number of particles of intracellular X factor, and $\tilde{X}_i^{\text{ext}}$ the number of particles of extracellular factor. We also associate to each site a volume V, which is assumed to remain constant. The temporal evolution of the probability $P(\boldsymbol{\omega}, t)$ to find the system in a configuration $\boldsymbol{\omega}$ at time t is

TABLE V. Propensity functions.

Process	Transition probability
Cell hopping to the left	$\Gamma_i / 2 \times n_i (1 - n_{i-1})$
Cell hopping to the right	$\Gamma_i / 2 \times n_i (1 - n_{i+1})$
(with $\Gamma_i = \Gamma_0 \exp\left[-\alpha \, \tilde{E}_i / V\right]$)	
Cell proliferation to the left	$\Theta_i / 2 \times n_i (1 - n_{i-1})$
Cell proliferation to the right	$\Theta_i / 2 \times n_i (1 - n_{i+1})$
E-cadherin synthesis	$V \, v_e \left(\frac{\tilde{X}_i}{V K n_i + \tilde{X}_i} \right) n_i$
E-cadherin degradation	$k_e \, \tilde{E}_i$
Chemical factor synthesis	$V \, v_X \, n_i$
Chemical factor degradation	$k_X \, \tilde{X}_i$
Chemical factor secretion	$k_{\text{out}} \, \tilde{X}_i$
Chemical factor uptake	$k_{in} \, \tilde{X}_i^{\text{ext}}$
Chemical factor hopping	$\Gamma_X / 2 \times \tilde{X}_i^{\text{ext}}$

assumed to obey a master equation

$$\frac{dP(\boldsymbol{\omega}, t)}{dt} = \sum_{i, \rho} \left[W_\rho^i(\boldsymbol{\omega}|\boldsymbol{\omega}') P(\boldsymbol{\omega}', t) - W_\rho^i(\boldsymbol{\omega}'|\boldsymbol{\omega}) P(\boldsymbol{\omega}, t) \right], \tag{A2}$$

in which the $\{W_\rho^i\}$ stand for the local transition probabilities per unit time of the various processes ρ. This choice amounts to assuming that the stochastic processes are Markovian.

Table V summarizes our choices for the local transition probabilities of the different events. Using these expressions, evolution equations for the means of n_i, $\tilde{E}_i$, $\tilde{X}_i$, and $\tilde{X}_i^{\text{ext}}$ can be obtained by multiplying both sides of (A2) with the appropriate variable and by summing over the configurations $\boldsymbol{\omega}$. This gives, for the mean occupation number $\langle n_i \rangle$,

$$\frac{d\langle n_i \rangle}{dt} = \left\langle \frac{\Gamma_{i+1}}{2} n_{i+1} (1 - n_i) \right\rangle + \left\langle \frac{\Gamma_{i-1}}{2} n_{i-1} (1 - n_i) \right\rangle$$
$$- \left\langle \frac{\Gamma_i}{2} n_i (2 - n_{i+1} - n_{i-1}) \right\rangle$$
$$+ \left\langle \frac{\Theta_{i+1}}{2} n_{i+1} (1 - n_i) \right\rangle + \left\langle \frac{\Theta_{i-1}}{2} n_{i-1} (1 - n_i) \right\rangle. \tag{A3}$$

Note that angular brackets represent ensemble averages. A similar procedure can be used to extract the evolution equation for the mean density of E-cadherin, $\langle E_i \rangle \equiv \langle \tilde{E}_i / V \rangle$:

$$\frac{d\langle E_i \rangle}{dt} = v_e \left\langle \frac{X_i}{K n_i + X_i} n_i \right\rangle - k_e \langle E_i \rangle$$
$$+ \left\langle \frac{\Gamma_{i+1}}{2} n_{i+1} (1 - n_i) E_{i+1} \right\rangle$$
$$+ \left\langle \frac{\Gamma_{i-1}}{2} n_{i-1} (1 - n_i) E_{i-1} \right\rangle$$
$$- \left\langle \frac{\Gamma_i}{2} n_i (2 - n_{i+1} - n_{i-1}) E_i \right\rangle$$
$$+ \left\langle \frac{\Theta_{i+1}}{2} n_{i+1} (1 - n_i) E_{i+1} \right\rangle$$
$$+ \left\langle \frac{\Theta_{i-1}}{2} n_{i-1} (1 - n_i) E_{i-1} \right\rangle, \tag{A4}$$

in which $X_i = \tilde{X}_i/V$ is the concentration of factor X in site i. The evolution equation for the mean of this local concentration reads as

$$
\frac{d\langle X_i \rangle}{dt} = v_X \langle n_i \rangle - k_X \langle X_i \rangle + k_{in} \langle X_i^{\text{ext}} \rangle - k_{\text{out}} \langle X_i \rangle
$$
$$
+ \left\langle \frac{\Gamma_{i+1}}{2} n_{i+1} (1 - n_i) X_{i+1} \right\rangle
$$
$$
+ \left\langle \frac{\Gamma_{i-1}}{2} n_{i-1} (1 - n_i) X_{i-1} \right\rangle
$$
$$
- \left\langle \frac{\Gamma_i}{2} n_i (2 - n_{i+1} - n_{i-1}) X_i \right\rangle
$$
$$
+ \left\langle \frac{\Theta_{i+1}}{2} n_{i+1} (1 - n_i) X_{i+1} \right\rangle
$$
$$
+ \left\langle \frac{\Theta_{i-1}}{2} n_{i-1} (1 - n_i) X_{i-1} \right\rangle. \tag{A5}
$$

Finally, we obtain for $\langle X_i^{\text{ext}} \rangle = \langle \tilde{X}_i^{\text{ext}}/V \rangle$,

$$
\frac{d\langle X_i^{\text{ext}} \rangle}{dt} = k_{\text{out}} \langle X_i \rangle - k_{in} \langle X_i^{\text{ext}} \rangle
$$
$$
+ \frac{\Gamma_X}{2} \langle X_{i+1}^{\text{ext}} + X_{i-1}^{\text{ext}} - 2 X_i^{\text{ext}} \rangle. \tag{A6}
$$

To obtain the evolution equations (1)–(4) presented in the main text, we first note that the variables entering the local state vector ω_i are expected to be strongly correlated to each other. For example, $\tilde{E}_i$ must always be zero whenever n_i is zero; a similar remark holds for $\tilde{X}_i$. We thus switch to new variables by introducing the amount of E-cadherin and of factor X per cell, $\tilde{E}_i^0 \equiv \tilde{E}_i/n_i$ and $\tilde{X}_i^0 \equiv \tilde{X}_i/n_i$, respectively. This defines a new local configuration $\boldsymbol{\omega}^* = (n_i, \tilde{E}_i^0, \tilde{X}_i^0, \tilde{X}_i^{\text{ext}})$, and a probability distribution $P(\boldsymbol{\omega}^*, t)$. We then make the assumption that this probability can be factorized as

$$
P(\boldsymbol{\omega}^*, t) = \prod_i P(\boldsymbol{\omega}_i^*, t) \tag{A7}
$$

and that, in turn,

$$
P(\omega_i^*, t) = P(n_i, t) P(\tilde{E}_i^0, t) P(\tilde{X}_i^0, t) P(\tilde{X}_i^{\text{ext}}, t). \tag{A8}
$$

Using this *mean-field* assumption, we find for the evolution equation for the cell occupation probability

$$
\frac{dn_i}{dt} = \frac{1}{2} [\Gamma_{i+1} n_{i+1} (1 - n_i) + \Gamma_{i-1} n_{i-1} (1 - n_i)
$$
$$
- \Gamma_i n_i (1 - n_{i+1}) - \Gamma_i n_i (1 - n_{i-1})
$$
$$
+ \Theta_{i+1} n_{i+1} (1 - n_i) + \Theta_{i-1} n_{i-1} (1 - n_i)]. \tag{A9}
$$

Note that we do not write the brackets representing the means anymore, for clarity. We obtain for the other previously introduced concentrations

$$
\frac{dE_i}{dt} = n_i v_E \frac{X_i^0}{K + X_i^0} - k_E E_i
$$
$$
+ \frac{1}{2} [\Gamma_{i+1} n_{i+1} (1 - n_i) E_{i+1}
$$
$$
+ \Gamma_{i-1} n_{i-1} (1 - n_i) E_{i-1}
$$

$$
- \Gamma_i n_i (1 - n_{i+1}) E_i - \Gamma_i n_i (1 - n_{i-1}) E_i
$$
$$
+ \Theta_{i+1} n_{i+1} (1 - n_i) E_{i+1}
$$
$$
+ \Theta_{i-1} n_{i-1} (1 - n_i) E_{i-1}], \tag{A10}
$$

$$
\frac{dX_i}{dt} = v_X n_i - k_X X_i - k_{\text{out}} X_i + k_{in} X_i^{\text{ext}} n_i
$$
$$
+ \frac{1}{2} [\Gamma_{i+1} n_{i+1} (1 - n_i) X_{i+1}
$$
$$
+ \Gamma_{i-1} n_{i-1} (1 - n_i) X_{i-1}
$$
$$
- \Gamma_i n_i (1 - n_{i+1}) X_i - \Gamma_i n_i (1 - n_{i-1}) X_i
$$
$$
+ \Theta_{i+1} n_{i+1} (1 - n_i) X_{i+1}
$$
$$
+ \Theta_{i-1} n_{i-1} (1 - n_i) X_{i-1}], \tag{A11}
$$

and

$$
\frac{dX_i^{\text{ext}}}{dt} = k_{\text{out}} X_i - k_{in} X_i^{\text{ext}} n_i
$$
$$
+ \Gamma_X \left(X_{i+1}^{\text{ext}} + X_{i-1}^{\text{ext}} - 2 X_i^{\text{ext}} \right). \tag{A12}
$$

APPENDIX B: CONTINUOUS-SPACE LIMIT

To obtain an evolution equation for the cell occupation in the continuous-space limit, we introduce the space coordinate $x = ia$ where a is the distance between two first neighbors. Equation (1) becomes

$$
\frac{\partial n(x, t)}{\partial t} = \frac{1}{2} \{ \Gamma(x + a, t) n(x + a, t) [1 - n(x, t)]
$$
$$
+ \Gamma(x - a, t) n(x - a, t) [1 - n(x, t)]
$$
$$
- \Gamma(x, t) n(x, t) [1 - n(x + a, t)]
$$
$$
- \Gamma(x, t) n(x, t) [1 - n(x - a, t)]
$$
$$
+ \Theta n(x + a, t) [1 - n(x, t)]
$$
$$
+ \Theta n(x - a, t) [1 - n(x, t)] \}. \tag{B1}
$$

In the limit where a is small compared to the characteristic length scale of spatial profiles, one can expand the cell occupation as

$$
n(x \pm a, t) \approx n(x, t) \pm a \frac{\partial n(x, t)}{\partial x}
$$
$$
+ \frac{a^2}{2} \frac{\partial^2 n(x, t)}{\partial x^2} + \cdots . \tag{B2}
$$

Keeping terms up to the second order in a, the rate of change of the cell occupation due to diffusive processes reads as

$$
\left. \frac{\partial n}{\partial t} \right|_{\text{diffusion}} = \left[\frac{\Gamma a^2}{2} \frac{\partial^2 n}{\partial x^2} + n \frac{\partial^2}{\partial x^2} \left(\frac{\Gamma a^2}{2} \right) - n^2 \frac{\partial^2}{\partial x^2} \left(\frac{\Gamma a^2}{2} \right) \right.
$$
$$
\left. + \frac{\partial n}{\partial x} \frac{\partial}{\partial x} (\Gamma a^2) + n \frac{\partial n}{\partial x} \frac{\partial}{\partial x} (\Gamma a^2) \right]
$$
$$
= - \frac{\partial}{\partial x} \left[-D \frac{\partial n}{\partial x} - n(1 - n) \frac{\partial D}{\partial x} \right], \tag{B3}
$$

where $D = \frac{\Gamma a^2}{2}$ is the diffusion coefficient of the cell. Similarly, the proliferation terms in the evolution equation for cell occupation can be written as

$$\left.\frac{\partial n}{\partial t}\right|_{\text{proliferation}} = \frac{\Theta}{2}\left[\left(n + a\frac{\partial n}{\partial x} + \frac{a^2}{2}\frac{\partial^2 n}{\partial x^2}\right)(1-n)\right.$$
$$\left. + \left(n - a\frac{\partial n}{\partial x} + \frac{a^2}{2}\frac{\partial^2 n}{\partial x^2}\right)(1-n)\right]$$
$$= \Theta\, n\,(1-n) + R(1-n)\frac{\partial^2 n}{\partial x^2}\,, \qquad \text{(B4)}$$

where $R = \frac{\Theta a^2}{2}$ is a proliferation-induced transport coefficient. Summing Eqs. (B3) and (B4), we obtain the evolution equation for the cell occupation in the continuous-space limit [Eq. (6)]:

$$\frac{\partial n}{\partial t} = \left.\frac{\partial n}{\partial t}\right|_{\text{diffusion}} + \left.\frac{\partial n}{\partial t}\right|_{\text{proliferation}}$$
$$= -\frac{\partial}{\partial x}\left[-D\frac{\partial n}{\partial x} - n(1-n)\frac{\partial D}{\partial x}\right]$$
$$+ \Theta\, n\,(1-n) + R\,(1-n)\frac{\partial^2 n}{\partial x^2}. \qquad \text{(B5)}$$

APPENDIX C: NEIGHBOR-DEPENDENT HOPPING FREQUENCY

We first write the stochastic hopping frequencies to the left and to the right as

$$\Gamma_i^{\text{left}} = \Gamma_0\, e^{-\alpha\tilde{E}_i - \beta\tilde{E}_i\tilde{E}_{i+1}} \qquad \text{(C1)}$$

and

$$\Gamma_i^{\text{right}} = \Gamma_0\, e^{-\alpha\tilde{E}_i - \beta\tilde{E}_i\tilde{E}_{i-1}}, \qquad \text{(C2)}$$

where α reflects the action of E-cadherin on the mobility of a simple cell and β reflects the strength of the interaction between two molecules of E-cadherin belonging to two neighboring cells. A hopping process to the left is affected by the presence of a cell on the right, and a hopping process to the right is affected by the presence of a cell on the left.

The mean hopping frequency to the left to insert in the evolution equation of the mean cell occupation can be obtained by considering the average probability of having a cell at site

i hopping to the left:

$$\left\langle \Gamma_i^{\text{left}} n_i\,(1-n_{i-1}) \right\rangle = \left\langle \Gamma_0\, e^{-\alpha\tilde{E}_i - \beta\tilde{E}_i\tilde{E}_{i+1}} n_i\,(1-n_{i-1}) \right\rangle$$
$$= \sum_{\boldsymbol{\omega}} \Gamma_0\, e^{-\alpha\tilde{E}_i - \beta\tilde{E}_i\tilde{E}_{i+1}} n_i\,(1-n_{i-1})\, P(\boldsymbol{\omega}). \qquad \text{(C3)}$$

Using the mean-field assumptions (A7) and (A8), this expression becomes

$$\left\langle \Gamma_i^{\text{left}} n_i\,(1-n_{i-1}) \right\rangle$$
$$= \Gamma_0\, e^{-\alpha E_i^0}\left[1 + n_{i+1}\left(e^{-\beta E_i^0 E_{i+1}^0} - 1\right)\right]n_i\,(1-n_{i-1}). \quad \text{(C4)}$$

The same derivation can be applied to the hopping frequency to the right and one obtains

$$\left\langle \Gamma_i^{\text{right}} n_i\,(1-n_{i+1}) \right\rangle$$
$$= \Gamma_0\, e^{-\alpha E_i^0}\left[1 + n_{i-1}\left(e^{-\beta E_i^0 E_{i-1}^0} - 1\right)\right]n_i\,(1-n_{i+1}). \quad \text{(C5)}$$

Again, we do not write the brackets representing the means anymore, for clarity. The mean-field evolution equations for n_i, E_i, and X_i become

$$\frac{dn_i}{dt} = \frac{1}{2}\left[\Gamma_{i+1}^{\text{left}} n_{i+1}\,(1-n_i) + \Gamma_{i-1}^{\text{right}} n_{i-1}\,(1-n_i)\right.$$
$$- \Gamma_i^{\text{right}} n_i\,(1-n_{i+1}) - \Gamma_i^{\text{left}} n_i\,(1-n_{i-1})$$
$$\left. + \Theta_{i+1}\, n_{i+1}\,(1-n_i) + \Theta_{i-1}\, n_{i-1}\,(1-n_i)\right], \quad \text{(C6)}$$

$$\frac{dE_i}{dt} = n_i\, v_E\,\frac{X_i^0}{K + X_i^0} - k_E\, E_i$$
$$+ \frac{1}{2}\left[\Gamma_{i+1}^{\text{left}} n_{i+1}(1-n_i)E_{i+1} + \Gamma_{i-1}^{\text{right}} n_{i-1}(1-n_i)E_{i-1}\right.$$
$$- \Gamma_i^{\text{right}} n_i\,(1-n_{i+1})E_i - \Gamma_i^{\text{left}} n_i\,(1-n_{i-1})E_i$$
$$\left. + \Theta_{i+1}\, n_{i+1}\,(1-n_i)E_{i+1} + \Theta_{i-1}\, n_{i-1}\,(1-n_i)E_{i-1}\right], \quad \text{(C7)}$$

and

$$\frac{dX_i}{dt} = v_X\, n_i - k_X\, X_i - k_{\text{out}}\, X_i + k_{in}\, X_i^{\text{ext}}\, n_i$$
$$+ \frac{1}{2}\left[\Gamma_{i+1}^{\text{left}} n_{i+1}(1-n_i)X_{i+1} + \Gamma_{i-1}^{\text{right}} n_{i-1}(1-n_i)X_{i-1}\right.$$
$$- \Gamma_i^{\text{right}} n_i\,(1-n_{i+1})X_i - \Gamma_i^{\text{left}} n_i\,(1-n_{i-1})X_i$$
$$\left. + \Theta_{i+1}\, n_{i+1}\,(1-n_i)X_{i+1} + \Theta_{i-1}\, n_{i-1}\,(1-n_i)X_{i-1}\right]. \quad \text{(C8)}$$

[1] A. W. Lambert, D. R. Pattabiraman, and R. A. Weinberg, Cell **168**, 670 (2017).

[2] I. Pastushenko, A. Brisebarre, A. Sifrim, M. Fioramonti, T. Revenco, S. Boumahdi, A. V. Keymeulen, D. Brown, V. Moers, S. Lemaire, S. De Clercq, E. Minguijón, C. Balsat, Y. Sokolow, C. Dubois, F. De Cock, S. Scozzaro, F. Sopena, A. Lanas, N. D'Haene *et al.*, Nature (London) **556**, 463 (2018).

[3] B. De Craene and G. Berx, Nat. Rev. Cancer **13**, 97 (2013).

[4] N. Skrypek, S. Goossens, E. De Smedt, N. Vandamme, and G. Berx, Trends Genet. **33**, 943 (2017).

[5] R. Kalluri and R. A. Weinberg, J. Clin. Invest. **119**, 1420 (2009).

[6] L. Shapiro and W. I. Weis, Cold Spring Harb. Perspect. Biol. **1**, a003053 (2009).

[7] A. P. Putzke, Am. J. Pathol. **179**, 400 (2011).

[8] M. Takeichi, Nat. Rev. Mol. Cell Biol. **15**, 397 (2014).

[9] J.-P. Thiery and J.-P. Seelman, Nat. Rev. Mol. Cell Biol. **7**, 131 (2006).

[10] S. Lamouille, J. Xu, and R. Derynck, Nat. Rev. Mol. Cell Biol. **15**, 178 (2014).

[11] P. A. Gregory, Mol. Biol. Cell. **22**, 1686 (2011).

[12] Y. Yuan, B. Liua, P. Xiea, M. Q. Zhanga, Y. Lia, Z. Xiea, and X. Wang, Int. J. Oncol. **40**, 1171 (2011).

[13] P. A. Gregory, C. P. Bracken, A. G. Bert, and G. J. Goodall, Cell Cycle **7**, 3112 (2008).

[14] H. Siemens, R. Jackstadt, S. Hunten, M. Kaller, A. Menssen, U. Gotz, and H. Hermeking, Cell Cycle. **10**, 4256 (2011).

[15] D. Iliopoulos, S. A. Jaeger, H. A. Hirsch, M. L. Bulyk, and K. Struhl, Mol. Cell. **39**, 493 (2010).

[16] N. Pencheva and S. F. Tavazoie, Nat. Cell Biol. **15**, 546 (2013).

[17] C.-C. Liang, A. Y. Park, and J.-L. Guan, Nat. Protoc. **2**, 329 (2007).

[18] L. Fan, H. Wang, X. Xia, Y. Rao, X. Ma, D. Ma, P. Wu, and G. Chen, Oncol. Lett. **4**, 1225 (2012).

[19] P. Kotaa, E. M. Terrella, D. A. Ritta, C. Insinnaa, C. J. Westlakea, and D. K. Morrisona, Proc. Natl. Acad. Sci. USA **116**, 3536 (2019).

[20] E. R. Shamir, E. Pappalardo, D. M. Jorgens, K. Coutinho, W.-T. Tsai, K. Aziz, M. Auer, P. T. Tran, J. S. Bader, and A. J. Ewald, J. Cell Biol. **204**, 839 (2014).

[21] R. A. Fisher, Ann. Eugenics **7**, 355 (1937).

[22] P. K. Maini, D. S. McElwain, and D. I. Leavesley, Tissue Eng. **10**, 475 (2004).

[23] M. J. Simpson, K. K. Treloar, B. J. Binder, P. Haridas, K. J. Manton, D. I. Leavesley, D. L. S. McElwain, and R. E. Baker, J. R. Soc. Interface **10**, 20130007 (2013).

[24] K. K. Treloar, M. J. Simpson, D. S. McElwain, and R. E. Baker, J. Theor. Biol. **356**, 71 (2014).

[25] S. T. Johnston, E. T. Shah, L. K. Chopin, D. L. S. McElwain, and M. J. Simpson, BMC Syst. Biol. **9**, 38 (2015).

[26] A. Q. Cai, K. A. Landman, and B. D. Hughes, J. Theor. Biol. **245**, 576 (2007).

[27] A. Tremel, A. Cai, N. Tirtaatmadja, B. Hughes, G. Stevens, K. Landman, and A. O'Connor, Chem. Eng. Sci. **64**, 247 (2009).

[28] Y. Lou, J. Xia, W. Tang, and Y. Chen, Phys. Rev. E **96**, 062418 (2017).

[29] E. Khain, M. Katakowski, N. Charteris, F. Jiang, and M. Chopp, Phys. Rev. E **86**, 011904 (2012).

[30] J. A. Sherratt and J. D. Murray, Proc. R. Soc. London, Ser. B **241**, 29 (1990).

[31] P. D. Dale, P. K. Maini, and J. A. Sherratt, Math. Biosci. **124**, 127 (1994).

[32] N. Kosaka, H. Iguchi, Y. Yoshioka, F. Takeshita, Y. Matsuki, and T. Ochiya, J. Biol. Chem. **285**, 17442 (2010).

[33] A. Grau Ribes, Y. De Decker, C. Gérard, and L. Rongy, Mol. BioSyst. **13**, 2379 (2017).

[34] J.-H. Fang, Z.-J. Zhang, L.-R. Shang, Y.-W. Luo, Y. Lin, Y. Yuan, and S.-M. Zhuang, Hepatology **68**, 1459 (2018).

[35] H.-W. Kang, M. Crawford, M. Fabbri, G. Nuovo, M. Garofalo, S. P. Nana-Sinkam, and A. Friedman, PLoS ONE **8**, e53663 (2013).

[36] E. E. Creemers, A. J. Tijsen, and Y. M. Pinto, Circ. Res. **110**, 483 (2012).

[37] Y. De Decker, G. A. Tsekouras, A. Provata, T. Erneux, and G. Nicolis, Phys. Rev. E **69**, 036203 (2004).

[38] A. C. Romano, E. M. Espana, S. H. Yoo, M. T. Budak, J. M. Wolosin, and S. C. G. Tseng, Invest. Ophthalmol. Vis. Sci. **44**, 5125 (2003).

[39] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, Nature (London) **473**, 337 (2011).

[40] J. Fernandez-Banet, A. Esposito, S. Coffin, I. B. Horvath, H. Estrella, S. Schefzick, S. Deng, K. Wang, K. AChing, Y. Ding, P. Roberts, P. A. Rejto, and Z. Kan, Nat. Methods **13**, 9 (2016).

[41] M. J. Marzi, F. Ghini, B. Cerruti, S. de Pretis, P. Bonetti, C. Giacomelli, M. M. Gorski, T. Kress, M. Pelizzola, H. Muller, B. Amati, and F. Nicassio, Genome Res. **26**, 554 (2016).

[42] U. Bissels, S. Wild, S. Tomiuk, A. Holste, M. Hafner, T. Tuschl, and A. Bosio, RNA **15**, 2375 (2009).

[43] T. Tian, Y.-L. Zhu, F.-H. Hu, Y.-Y. Wang, N.-P. Huang, and Z.-D. Xiao, J. Cell. Physiol. **228**, 1487 (2013).

[44] J. Li and R. R. Rhinehart, Comput. Chem. Eng. **22**, 427 (1998).

[45] M. Han, M. Liu, Y. Wang, Z. Mo, X. Bi, Z. Liu, Y. Fan, X. Chen, and C. Wu, Molec. Cell. Biochem. **363**, 427 (2011).

[46] R. Farooqui and G. Fenteany, J. Cell Sci. **118**, 51 (2005).

[47] T. R. Graham, H. E. Zhau, V. A. Odero-Marah, A. O. Osunkoya, K. S. Kimbro, M. Tighiouart, T. Liu, J. W. Simons, and R. M. O'Regan, Cancer Res. **68**, 2479 (2008).

[48] M. Kujawski, M. Kortylewski, H. Lee, A. Herrmann, H. Kay, and H. Yu, J. Clin. Invest. **118**, 3367 (2008).

[49] T. S. Teng, B. Lin, E. Manser, D. C. H. Ng, and X. Cao, J. Cell Sci. **122**, 4150 (2009).

[50] C. Kahlert and R. Kalluri, J. Molec. Mec. **91**, 431 (2013).

[51] E. Scarpa, A. Szabo, A. Bibonne, E. Theveneau, M. Parsons, and R. Mayor, Dev. Cell. **34**, 421 (2015).

[52] A. Stockinger, A. Eger, J. Wolf, H. Beug, and R. Foisner, J. Cell Biol. **154**, 1185 (2001).

[53] S. Y. Park, J.-H. Shin, and S.-H. Kee, Cancer Sci. **108**, 1769 (2017).

[54] J. Kowal, G. Arras, M. Colombo, M. Jouve, J. P. Morath, B. Primdal-Bengtson, F. Dingli, D. Loew, M. Tkach, and C. Thery, Proc. Natl. Acad. Sci. USA **113**, E968 (2016).

[55] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, and M. J. Simpson, Biophys. J. **114**, 1241 (2018).

[56] L. Guo, C. Chen, M. Shi, F. Wang, X. Chen, D. Diao, M. Hu, M. Yu, L. Qian, and N. Guo, Oncogene **32**, 5272 (2013).

[57] C. Ma, K. Nong, H. Zhu, W. Wang, X. Huang, Z. Yuan, and K. Ai, Tumor Biol. **35**, 9163 (2014).

[58] V. Luga, L. Zhang, A. M. Viloria-Petit, A. A. Ogunjimi, M. R. Inanlou, E. Chiu, M. Buchanan, A. N. Hosein, M. Basik, and J. L. Wrana, Cell **151**, 1542 (2012).