Supplemental Text

In this Supplemental Theory Note, we provide additional details on the modelling approach, analytical calculations underlying the results of the main text, as well as controls and sensitivity analyses on how different parameter values affect the theoretical branching structure in relation to our dataset. We also consider alternative models from our branching and annihilating random walk framework, to compare and contrast them quantitatively with our data.

1 Models of the branching and annihilating random walk

As detailed in the main text, our data indicate that ductal tips, both in mammary gland and, by analogy, in kidney, elongate and branch stochastically as a default state. However, tips can also terminate at any stage during morphogenesis, and thereby become irreversibly removed from the active pool of tissue precursors. In a first attempt to model the dynamics, in the following, we will consider the case in which tip termination is mediated by contact between two actively proliferating tips. Interestingly, such a model maps directly onto the problem of branching and annihilating random walks, which have been extensively studied by the physics community.

1.1 Description of the model

Classical branching and annihilating random walks (BARWs) are defined by a set of n walkers undergoing random walks in d spatial dimensions with a diffusion constant D. We denote as $\mathbf{r}_i(t)_{i=1..n}$ the position vector of each walker in space at a time t. In the general case, walkers A can either branch into $m+1$ new walkers:

$$A \to (m+1)A \text{ , at rate } r_{\text{b}}, \tag{1}$$

or they may annihilate into an inert state I when two walkers meet locally in space,

$$2A \to I \text{ , at a rate that scales with volume as } r_{\text{a}}. \tag{2}$$
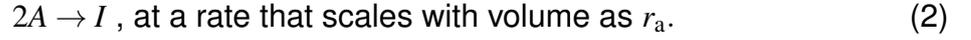
Here, the branch multiplicity m is a strictly positive integer. Walkers that fall into the inactive state do so irreversibly and do not diffuse, nor interact with other walkers. The total number of walkers, n, active or inactive, therefore evolves in time. We denote as $s_i(t)_{i=1..n}$ the state of each walker (with active denoted by 1 and inactive as 0).

Defined in this form, this classical model has been studied extensively in the literature (Cardy and Tauber, 1996). Insight into the behaviour of the model can be gained from the corresponding mean-field rate equation for the local density of walkers, $n(t)$, which takes the form:

$$\partial_t n(t) = r_{\text{b}} m n(t) - r_{\text{a}} n(t)^2 \tag{3}$$

This equation predicts a pair of stationary solutions corresponding to an unstable inactive state ($n = 0$) and a stable active state ($n_s = m r_{\text{b}}/r_{\text{a}}$), for any positive value of m. However, when fluctuations around the mean-field approximation are taken into account, the situation is revealed as more complex: In particular, it has been shown that the system can belong to two distinct universality classes for either odd or even values of m. For even values of m, the parity of the number of active particles is conserved

upon branching, falling into the parity-conserving class, while the case of odd values of m falls into the directed percolation universality class (Cardy and Tauber, 1996). Theoretical analysis of these two classes has been facilitated by the fact that they can be mapped on an Ising model, where branching and annihilation events can be mapped onto Glauber and spin-exchange Kawasaki dynamics (Ódor, 2004).

Turning to the biological data, virtually all biological trees examined (kidney, salivary gland, mammary gland) are driven by tips, often termed proliferative terminal end-buds (TEBs), which either advance forward, producing in their wake the ductal structure, or branch into two tips during a bifurcation event. In the case of mammary gland and kidney morphogenesis, the focus of the current work, the ductal region deposited by tips is inert, i.e. it does not proliferate further nor move. Moreover, most branching events are observed to involve tip bifurcation in which a single active tip divides into precisely two active tips.

Translated into the formalism described above, we therefore have $m = 1$, while the rate of bifurcation, r_{b}, and diffusion coefficient, D, can be fit from the kinetics of branching observed *in vivo*. Importantly, since the ductal trees constitute the past trajectories of each individual active walker (or tip), we can therefore reconstruct quantitatively the entire time course of the process from the analysis of the final branching structure in space.

Looking at the actual trajectories *in vivo* (Figure 1A and Figure S1), one detail must be amended straightaway. It is clear that the trajectories are not temporally uncorrelated random walks but, instead, display a typical persistence length that can again be read from the biological data. We must therefore amend the previous model by introducing a unit polarity vector $\mathbf{p}_i$, which specifies the direction of movement of a

given walker (we take the local velocity v of each walker to be the same). This polarity vector undergoes a persistent random walk with characteristic time τ_{p}. Translated spatially, this means that the persistence length of ductal branches should be $v\tau_{\mathrm{p}}$ and the long-term (i.e. at time scales larger than τ_{p}) diffusion constant of a walker is $v^2\tau_{\mathrm{p}}$.

1.2 Mean-field Fisher-KPP dynamics of front invasion

To account for the spatial dynamics of branching morphogenesis, we must amend the mean-field theory above to take into account of the evolving spatial structure. In line with the biological system, we take as an initial condition for branching morphogenesis a single active walker positioned at one side of a finite domain (Figure S1 and Movie S1). Branching morphogenesis then proceeds by invading into this domain. Following classical results of simple branching processes (Haskovec et al., 2016; Brunet et al., 2006), the spatial dynamics in d-dimensions can be written as

$$
\begin{cases}
\partial_t a = D\nabla^2 a + r_{\mathrm{b}} a(1 - \frac{a}{a_0}) \\[2ex]
\partial_t i = r_{\mathrm{e}} a + \frac{r_{\mathrm{b}}}{a_0} a^2
\end{cases}
\tag{4}
$$

where $a(\mathbf{r},t)$ and $i(\mathbf{r},t)$ represent, respectively, the local concentration of active and inactive particles and ∇^2 denotes the Laplace operator. Related to the description above, active particles diffuse with a diffusion constant D, branch at a rate r_{b} and annihilate when they meet, giving rise to a logistic growth term saturating at a density a_0. Inactive particles are constantly produced either when active particles move, or when two active particles meet. Their coefficient of diffusion is zero, expressing the fact that they remain immobile and inactive.

The first of these equations (4) translates to the well-known Fisher-KPP equation, after Kolmogorov-Petrovsky-Piscounov, which has been widely used in the theory of

stochastic processes as a simple approximation to describe front propagation (e.g. in the study of advantageous genetic mutations) (Fisher, 1937). Note that the second equation, describing the concentration of inactive particles, is slave to the dynamics of the active particles, and does not provide any feedback. According to this dynamics, the theory predicts that active walkers will form a sharp front (of characteristic length $\sqrt{D/r_b}$), which will invade the tissue as a solitary wave, with velocity $V^* = 2\sqrt{Dr_b}$. The associated wave profile links the unstable ($a = 0$) solution at the growing front of the ductal network to the stable ($a = a_0$) solution at the back (heteroclinic orbit).

Active tips thus form a non-equilibrium stable steady-state (i.e. a constant density profile in time, but which is driven by constant and compensatory creation/annihilation events), which invades the unstable inactive state, $a = 0$. However, this solution poses the problem that inactive particles keep being produced at the back of the wave, driving a divergence in their concentration over time as $i \propto a_0(r_e + r_b)t$. As the wave propagates linearly in time, this translates into a linear concentration gradient towards the back of the waves, i.e. the concentration becomes higher and higher as $x \to \infty$ with slope $a_0(r_e + r_b)/V^*$). Such behaviour is clearly untenable and unrealistic as a model of branching morphogenesis, and argues for an alternative mechanism of active tip termination.

1.3   A revised model of tip termination

From the biological perspective, the regulation of tip termination could constitute an incredibly complicated process. For instance, one could envisage that each walker/tip is specified early in development with disparate properties, and is programmed to terminate at a given time. At each bifurcation event, a given tip would segregate into

two tips which both inherit a defined differentiation program. At the other end of the spectrum, one could envisage an extrinsic regulatory mechanism based, for example, on hormonal levels (Sternlicht, 2005), where tips would be allowed to grow and branch during a period T while they invade the full domain, and would be stopped and made inactive by a global signal at the end of this period. One should note that this type of dynamics would still leave a signature in the ductal structure such that the concentration of ducts close to the initial tip location should be higher than at the more distant parts recently invaded.

Therefore, to define an alternative mechanism for tip termination, we returned to the morphogenesis data of the mammary gland to explore whether there was evidence of a simpler mode of regulation. A first obvious feature was that the density of ducts was surprisingly homogeneous spatially, arguing against the simple model described above (Figure 1B). A second conspicuous feature of the branching pattern was the rarity of ductal crossovers (quantified in Figure 1C as the ratio between ductal crossovers and the total number of branches in a gland, $n = 14$ glands in 7 mice). Moreover, we could often observe inactive tips in the vicinity of existing ducts (Figure 1A).

We therefore considered the alternative hypothesis that active tips do not only only annihilate when they come in proximity of one another, but also that they terminate when they come into proximity of an inactive particle (i.e. the trailing ducts):

$$A + I \rightarrow 2I \text{ , at rate } r_{\mathrm{a}} \tag{5}$$

Such a process drastically complicates the classical problem of BARWs, as it makes it fully non-Markovian: Understanding the dynamics of active walkers A requires a knowledge of the full history of the random walk, and not just the current spatial configuration. (Alternatively, one can hide the non-Markovian character of the problem

in the second species *I*, which previously was simply slave to the active particles.) This bears some resemblance to generalizations of directed percolation models; in particular, those which couple a directed-percolation process to a frozen field (Jensen and Dickman, 1993). A key difference then becomes that there are infinite numbers of absorbing states (as the configurations of the frozen field are taken into account), as opposed to simply $A = 0$. However, in our model, the frozen particles *I* actively feedback on the dynamics of *A*, a situation reminiscent of two-species epidemic type models (Van Wijland et al., 1998; Rossi et al., 2000). In the following subsection, we thus only concentrate on developing a mean-field theory for such modified two-species BARW.

1.3.1 Mean-field Fisher-KPP pulse invasion

In the present of process (5), the continuous mean-field approximation to the dynamics then modifies Eq. (4) to the form:

$$\begin{cases} \partial_t a = D\nabla^2 a + r_\mathrm{b} a (1 - \frac{a+i}{n_0}) \\ \\ \partial_t i = r_\mathrm{e} a + \frac{r_\mathrm{b}}{n_0} a(a+i) \end{cases} \tag{6}$$

where we renamed $a_0 \to n_0$ for the sake of clarity. To proceed further, it is useful to non-dimensionalize this set of equations. Rescaling all times by $1/r_\mathrm{e}$, all concentrations by n_0, and all distances by $\sqrt{D/r_\mathrm{e}}$, and introducing the dimensionless ratio $\bar{r} = r_\mathrm{b}/r_\mathrm{e}$, the system of equations in one-dimension (coordinate x) can be written as:

$$\begin{cases} \partial_t a = \partial_x^2 a + \bar{r} a (1 - a - i) \\ \\ \partial_t i = a + \bar{r} a(a+i) \end{cases} . \tag{7}$$

Indeed, the parameter $\bar{r}$ can also be removed by introducing two different scales for the concentrations of a and i. However, we choose not to do so in order to be able to

compare directly the magnitude of the two in a transparent manner, for reasons that will become clear later on.

To seek a traveling wave solution for a and i, it is convenient to change coordinates to the co-moving frame by defining a variable $z = x - Vt$, where V is the wave velocity. The system of equations then becomes

$$\begin{cases} -V\partial_x a = \partial_x^2 a + \bar{r}a(1 - a - i) \\ \\ -V\partial_x i = a + \bar{r}a(a + i) \end{cases}. \tag{8}$$

Non-integrability of these equations prevents an analytical solution. Instead, we will therefore first perform two approximations setting $a \ll i$ in the first equation, and $\bar{r}(a + i) \ll 1$ in the second. This allows us to gain significant analytical insight into the dynamics, and we will justify the two assumptions *a posteriori*. With these approximations, the coupled equations take the form:

$$\begin{cases} -V\partial_x a = \partial_x^2 a + \bar{r}a(1 - i) \\ \\ -V\partial_x i = a \end{cases} \tag{9}$$

Compared to the classical Fisher-KPP equation, an important difference is that the negative feedback does not arise from a, but instead from the quantity i, which is dynamically produced by a. This is known to produce so-called KPP pulses (Haskovec et al., 2016) instead of KPP waves. Indeed, a now has a single stable state $a = 0$, and the pulse joins an $a = 0$ front with an $a = 0$ back (homoclinic orbit).

However, the concentration of the "inhibitor", i, adopts a different profile. The steady-state concentration for i is harder to read from Eq. (9), but becomes obvious after substituting the second equation for a into the first, and integrating once, leading to:

$$-Va = \partial_x a + \bar{r}i\left(1 - \frac{i}{2}\right). \tag{10}$$

Note that the constant of integration vanishes since all concentrations and their derivatives must vanish at $x \to \infty$. It then becomes more apparent that $i = 2$, $a = 0$ is a stable solution at the back at the wave, together with $i = 0$, $a = 0$ at the front. In contrast to the pulse shape of the a profile, the concentration of inactive ducts i thus adopts a traveling wave solution. From a physical perspective, this corresponds to the fact that active walkers can freely diffuse at the front of the wave, in the absence of inhibitors, but that they are trapped by inactive walkers at the back of the wave, and become inactivated. These processes specify a ductal network of well-defined, homogeneous, density $i = 2$ at which the gland is patterned.

Using usual methods for the analysis of such travelling pulses, we assume that the profile at the front decays exponentially: $a = Ae^{\lambda_+ z}$. Inserting this ansatz into the linearised system of equations above, we obtain:

$$\lambda_+ = \frac{-V \pm \sqrt{V^2 - 4\bar{r}}}{2}$$

Similarly, at the back of the wave, with $a = Ae^{\lambda_- z}$, we obtain (noting that λ_- must be strictly positive),

$$\lambda_- = \frac{-V + \sqrt{V^2 + 4\bar{r}}}{2}$$

As $\lambda_\pm$ must be purely real, velocities below $V^* = 2\sqrt{\bar{r}}$ cannot propagate. Moreover, a classical result of such KPP systems is that the velocity selected is exactly V^*, as verified through numerical integration of these equations (Figure S4D-G). Therefore, the decay length of the font and back of the pulse are given by

$$\begin{cases} \lambda_+ = -\sqrt{\bar{r}} \\ \\ \lambda_- = (\sqrt{2} - 1)\sqrt{\bar{r}} \end{cases} \tag{11}$$

which predicts that the back of the pulse decays slower than the front.

We now return to consider the validity of the approximations made above. At the back of the pulse, $i \to 2$ and $a \to 0$, so that the first approximation $a \ll i$ is trivially satisfied. At the front of the pulse, it follows from the expressions above that $a/i \approx 2\bar{r}$. However, $\bar{r}$ is the ratio of the time scale of tip elongation compared to tip branching. On a discrete lattice, this would be the inverse of the average number of lattice steps which the walker explores before a branching event. Therefore, having a sparsely branched structure like the mammary gland imposes the condition $\bar{r} \ll 1$, as opposed to $\bar{r} \approx 1$, which would give rise to a dense structure, and the assumption $a/i \ll 1$ can be made with high precision. Similarly, $\bar{r} \ll 1$ allows us to also validate *a posteriori* the second approximation $\bar{r}(a+i) \ll 1$ (which, in fact, is exactly the same level of approximation as the first one, comparing $\bar{r}$ and $\mathcal{O}(1)$). However, for the sake of completeness, we also performed numerical integrations of the full equations for $\bar{r} = 0.1$ without approximation (Figure S4E-F), as well as numerical integrations for $\bar{r} = 1$ (Figure S4G), when the approximation becomes invalid, and verified that neither the phenomenology of the solution, nor the pulse asymmetry, are qualitatively affected.

1.3.2   Numerical simulations

We then returned to the full simulation of the BARW model, and checked whether the results matched the simple analytical theory described above. From previous work on classical BARWs, we expect the mean-field theory to give a good insight into the behavior for small fluctuations (i.e. small annihilation probability), while the active state can be destroyed by these fluctuations when they become larger. We simulated a quasi one-dimensional front invasion by considering an elongated two-dimensional rectangular domain Ω of dimensions L_x and $L_y < L_x$, so that $\Omega = \{(x,y) \in \mathbb{R}^2 : \{x \in$

$[0,L_x],y \in [0,L_y]\}\}$. The initial condition is a single active walker located at the left side of the domain $(0,L_y/2)$, with polarity facing the right/posterior side $\mathbf{p} = (1,0)$. We fixed the length unit of the simulation by the elementary length of the random walk at each time point l. Crucially, we made the most simple assumption for annihilation that an active particle (tip) terminates deterministically when entering into an annihilation radius R_a of any active or inactive particle belonging to a different duct. Therefore, there is no need for implementing a rate of annihilation r_a as it emerges from the spatial interactions. We set the timescale of the simulation as the time τ_s for each active particle to make an elementary step of length $l = 1$ in the direction $\mathbf{p}$. Therefore, the only truly free parameter of the numerical simulations is the branching rate r_b, which controls the final density of the ductal network and the average branch length l_d. More precisely, the key parameter in this problem is the dimension of the fat pad, rescaled by the average branch length l_d. For the default set of parameters used to describe mammary gland morphogenesis, we thus use the experimentally measured rescaled long and short dimensions of the mammary gland (resp. L_x/l_d and L_z/l_d, Figure S1D). We ran 2000 full stochastic simulations, and averaged the results, which are thus parameter-free predictions, i.e. do not involve the fitting of a free parameter.

Interestingly, we observed that the full simulation agrees qualitatively with the KPP pulse theory described above. The active walkers invade the domain as a pulse with a well-defined constant velocity, leaving behind a ductal structure of constant density. Quantitative comparisons between the profiles of both a and i in the simulation also agree well with the analytical theory (Figure 3 and Figure S4).

1.3.3 Details of the statistical approach

To challenge quantitatively the goodness of the fits in the main text, as well as its explanatory power, we verified the model predictions by calculating in both cases the coefficient of determination R^2, which is the simplest indicator of the goodness of a fit. In addition, we calculated the standard error of the fit, S, which is an absolute measure of the residuals of the fit, and has been shown to be more adapted for non-linear fitting procedures (Brown, 2001). For a good fit, R^2 should be as close to 1 as possible, whereas S should be as close to 0 as possible. One should note that, for the sake of completeness, we also calculated values of R^2 by taking first the logarithm of the data and predictions (which we denote as R_{log}^2 in the main text), given the power laws seen in several quantities of the dataset. This measure is complementary, as it gives comparatively more weight to the agreement between theory and experiments for small parameter values. Importantly, we verified that we get excellent goodness of fits measurements in each case.

Specifically, for the definition used in the main text, defining n as the number of the points being fitted, $\bar{y} = \frac{1}{n}\sum_i^n y_{\text{obs}}^i$ as the average of the observable, $S_{\text{tot}} = \sum_i^n (y_{\text{obs}}^i - \bar{y})^2$, and $S_{\text{res}} = \sum_i^n (y_{\text{obs}}^i - y_{\text{model}}^i)^2$, the coefficient of determination is defined as $R^2 = 1 - \frac{S_{\text{res}}}{S_{\text{tot}}}$, while the standard error of the fit is given by $S = \sqrt{\frac{S_{\text{res}}}{n}}$.

As described below, for the shaded areas of Figure 2E-F, we performed numerical integrations for the experimentally observed average value of the branch lengths, as well as for the values $\pm$ one standard deviation. For all three values of resulting r_{b} estimate, we then performed at least 2000 numerical simulations and computed curves for each case. The shaded area thus represent the sensitivity of the model prediction with respect to one standard deviation variation of the key parameter r_{b}. Moreover, to

build the error bars shown in Figure 2E-F and 7C-D for the experimental values of the cumulative subtree size distribution and subtree persistence, given the small number of subtrees in each mammary gland, we used a bootstrapping method. This method involves calculating the cumulative distribution function for a large number of artificial datasets, which are samples with replacement of the original dataset (i.e. sizes or maximal generation, respectively, of subtrees), and calculating error bars, defined here as confidence intervals of one standard deviation, from the resulting cumulative distributions of each artificial dataset.

For the shaded areas of Figure 7A-B, we performed at least 1000 numerical simulations using the default parameters (see below) and calculated in each case respectively the branch number per generation and inactive versus total tip number. We then computed averages and 95% confidence intervals from these simulations, which therefore express the stochastic fluctuations of the model for a given and constant value of the annihilation radius R_a.

2   Sensitivity analysis - effect of parameter variations on the predicted structure of the mammary gland

In this section, we explore further the overall robustness of your model with respect with parameter variations, and show in particular that its predictions are insensitive to details of the simulation scheme used.

2.1 Analysis of final mammary gland morphology

2.1.1 2D branching morphogenesis in an unbound domain

As noted in the main text, we first explored the case of branching morphogenesis in an unbound two-dimensional domain, in order to understand better the dynamics of branching and annihilation in the absence of annihilating outer boundary. As predicted by the KPP mean-field theory, we observed a characteristic pattern of radial expansion at a constant velocity (Figure S2A), with the active tips residing at the front of the invading wave. Importantly, we found a robust and characteristic convergence to a non-equilibrium steady-state network configuration (Figure S2B), on the timescale of a few generations, characterised by near-perfect balance between the rates of tip branching and termination. This confirms the analytical arguments developed above. In the following discussion, we now include annihilating walls, which represent the boundary of the mammary fat pad. This is motivated by the fact that tips cannot cross the boundary of the fat pad *in vivo*, and that numerous tip termination events can be observed at these boundaries (Figure 1A and Figure S1A).

In the following, we successively study the sensitivity of the results presented in the main text against different details of the numerical simulation scheme for the branching and annihilating random walks presented in the main text. We show in particular that the only key parameter influencing the results is the branching rate, for which we use the experimentally measured value in the simulations.

2.1.2 Key parameters for the mammary gland simulations

Before performing a sensitivity analysis as a function of parameter values, we list the default parameter set used in the main text for the numerical simulation of mammary

gland morphology. Based on experimental measures (Figure 1), we set a rate of branching of $r_b = 0.1$, together with box sizes of $L_x = 280$ and $L_y = 150$. At each time step (with step time $\tau = 1$ used as the unit time), tips move forward by a length $V\tau = 1$ (used as the unit length), along a polarity vector $\mathbf{p}_i$ (specified in 2D by an angle θ), which diffuses by a random angle of amplitude $\delta\theta = \pi/10$. We set the annihilation radius to the characteristic width of a mammary duct, estimated at $R_a = 3$ in simulation length units. Tips then terminate if they pass the bounds of the simulation box. With these parameters, we find that the model can predict with high precision the distribution of subtree size (defined as number of branches in a subtree, starting at branch level 6 from the embryonic rudimentary structure), as well as the subtree persistence (defined as the fraction of subtrees present at level 6 which are still active at a given level, or generation number). One should note that the broad nature of these distributions is a strong indication of the fact that active tips (and by extension different subtrees) compete neutrally: an ever-diminishing fraction of subtrees survives and colonizes an ever-expanding part of the fat pad. This is also a signature of the predominance of symmetric tip-bifurcations. Indeed, in a model where tips would divide asymmetrically between an active tip and an inactive tip (in a stereotypical fashion at the tip level), the aforementioned distributions would be sharply peaked around an average value, since there would be little subtree extinction.

2.1.3 Effect of the branching rate

First, we show in the main text as confidence intervals the effect of variations in the branch length, i.e. changes in the branching rate r_b. We use the experimentally measured standard deviation of the mean branch length among different mammary glands,

and thus simulate branching morphogenesis for $r_{\mathrm{b}} = 0.085$ and $r_{\mathrm{b}} = 0.12$, to build confidence intervals of one standard deviation with respect to the branching rate (shaded area in Figure 2D-F). As expected, lowering (respectively increasing) the branching rate increases (respectively decreases) the average number of branches in a mammary gland, and thus shifts significantly both the subtree size distribution and persistence (Figure 2E-F, shaded area). Interestingly, this has relatively little impact on the convergence towards balance between termination and branching (Figure 2D). Moreover, lowering (respectively increasing) the branching rate modifies the exponent of the giant number fluctuations (Figure S5C and main text): larger values of the branching rate tile space more effectively, so that the exponent is reduced, indicative of lower fluctuations. Conversely, smaller values of the branching rate approach the critical point, enhancing density fluctuations.

2.1.4 Effect of the annihilation radius

Second, we investigated the sensitivity of our results with respect to the value of the annihilation radius. As discussed in the main text, we do not expect this parameter to be crucial in two-dimensions for persistent random walks, as the probability for two lines to intersect is unity. Nevertheless, we wished to confirm this intuition quantitatively, and therefore ran our numerical simulations for different values of the annihilation radius R_{a} (a smaller value of $R_{\mathrm{a}} = 1.5$ is shown in Figure S2D and a larger value of $R_{\mathrm{a}} = 3.75$ is shown in Figure S2E). As expected, lowering the annihilation radius slightly delays the convergence towards balance between tip branching and termination (Figure S2F), but this effect was very small, even for more than two-fold variations in the annihilation radius R_{a}. Similarly, we observed that the cumulative subtree size probability is largely

unaffected by large differences in R_a (Figure S2G).

2.1.5   Effect of the persistence length

Third, we investigated the sensitivity of our results with respect to the value of the persistence length for the random walk. As stated above, we used a control value of $\delta\theta = \pi/10$. This was based on previous measurements of duct "tortuosity", i.e. the average real path length between consecutive duct branch points, L_d, was 6% larger than its straight branchpoint-to-branchpoint value l_d (Paine et al., 2016). For small path fluctuations, $h(x)$, with respect to the branchpoint-to-branchpoint axis x, $(L_d - l_d)/l_d \approx h'(x)^2/2 \approx \delta\theta^2/2$; hence the value chosen. Again, we ran our numerical simulations for different values of the persistence length, by varying the magnitude of the angle variation at each step of the random walk, $\delta\theta$. We considered both the limit of an infinity persistent walk ($\delta\theta = 0$, Figure S2H), and a lowly persistent random walk, with an angle $\delta\theta$ doubled compared to its reference value ($\delta\theta = 2\delta\theta_\mathrm{ref}$, Figure S2I). Although the microscopic morphology of the simulated glands was, as expected, slightly different, we found again that these parameter changes do not significantly affect either the evolution of the tip termination probability as a function of generation (Figure S2J), nor the cumulative subtree size probability (Figure S2K). Again, this argues for the robustness of our results with respect to even large perturbations of this parameter.

2.1.6   Effect of probabilistic annihilation

Fourth, although the mammary gland reconstructions indicate that the crossover between ducts is a rare event, it is not non-existent, as assumed in our simulations. Moreover, although we assume that termination occurs deterministically at a fixed ra-

dius between tip and duct, one expects the actual phenomenon to occur in a noisier and more stochastic manner. To remedy both of this limitations, we modified our simulations to allow for the possibility of crossovers, and implemented probabilistic annihilation. In particular, we assumed termination events occur stochastically

$$A + I \rightarrow 2I \text{ , at rate } r_{\mathrm{a}} \tag{12}$$

if an active tip A enters the vicinity of a duct, i.e. within the annihilation radius, R_a. One should note that the reference simulation is simply the deterministic case of $r_{\mathrm{a}} = 1$. In Figure S2L, we take $r_{\mathrm{a}} = 0.2$ and $R_a = 5$ (i.e. a larger annihilation radius than the reference case, to compensate for the less likely annihilation, and obtained simulated glands of comparable density).

As expected, although simulations now present a low but non-zero fractions of crossovers (Figure S2L), we still observed in this case the same convergence towards a balance between termination and branching (Figure S2N), with similar kinetics. Interestingly, the functional shape of the cumulative subtree size distribution was characterised by decreased heterogeneity, with less very large subtrees compared to the reference distribution (Figure S2O). This resembles the distribution that we previously found using a mean-field limit of this problem, i.e. exponential tails, arising from the stochastic choices of equipotent tips in a critical Galton-Watson birth-death type process (Scheele et al., 2017). This is not surprising, as in the limit of very large annihilation radii, the spatial advantage for a tip to be in a crowded region becomes vanishingly small, converging towards the mean-field solution.

2.1.7 Effect of the branching angle distribution

Fifth, we investigated the importance of the branching angle distribution in dictating the final morphology of the mammary gland. In the reference simulations presented in the main text, we used as an input for the branching angle distribution its experimentally observed counterpart. Indeed, as measured from our full gland reconstructions, we found that the branching angles of the two offsprings relative to the direction of the ancestor branch do not show any significant correlation, although the relative angle between the two offsprings are always above a minimum bound of roughly $\pi/16$. In the simulations, we thus stochastically and independently chose both from a Poisson distribution of parameter $\pi/6$ (together with a minimal bound of $\pi/16$).

However, we also wished to assess whether these choices of the branch angle distribution modified significantly our results. We thus modified our simulations by considering the branching angle between ancestor and offspring to be a given constant of $\alpha = 50$ degrees. As expected, we found a slightly more ordered gland morphology (Figure S2M). Importantly, however, we found that the convergence towards a balance between tip termination and branching (Figure S2N), as well as the cumulative distribution of subtree sizes, were only very weakly affected (Figure S2O).

2.1.8 Effect of additional self-avoidance

Finally, we wished to understand how our results would be modified if tips had additional self-avoiding properties, in addition to their previously-described branching and annihilating properties. Self-avoidance has been proposed in the past, both in the context of mammary and kidney development (Sternlicht, 2005, Davies et al., 2014), and can be modelled microscopically in a number of ways. Here, we chose the simplest

local implementation of such a self-avoiding rule. We assumed that a particle, located at a position $\mathbf{r}_i$, can sense an average density vector $\mathbf{p}_i^r$ arising from ducts and tips in a repulsion radius R_r (sketched on Figure S5E) such that,

$$\mathbf{p}_i^\mathrm{r} = \frac{\displaystyle\sum_{\substack{j \\ |\mathbf{r_i}-\mathbf{r_j}|<R_\mathrm{r}}} (\mathbf{r_i}-\mathbf{r_j})}{\left| \displaystyle\sum_{\substack{j \\ |\mathbf{r_i}-\mathbf{r_j}|<R_\mathrm{r}}} (\mathbf{r_i}-\mathbf{r_j}) \right|}. \tag{13}$$

$\mathbf{p}_i^r$ is thus a unit vector that weights equally every particle in a radius of R_r.

In order to implement repulsion, we thus considered that, at each random step of the random walk with $l = 1$ in the direction $\mathbf{p}_i$ (as described above), the particle performs an additional displacement of $-f_\mathrm{r}\,\mathbf{p}_i^\mathrm{r}$ (with its polarity vector $\mathbf{p}_i$ being updated accordingly). Thus, positive values of f_r correspond to self-avoiding random walks (as the particle will tend to move away from denser regions), while negative values of f_r correspond to self-attracting random walks. f_r thus represents the strength of the self-avoidance bias on the random walk. We investigated moderate ($f_r = 0.2$, Figure S5E-J) biases in our updated simulations in order to understand how this affects our results. One should note that boundaries can also be represented formally as ducts (with the same particle density), and included in Eq. (13), if one wants to assume that the boundaries of the mammary fat pad repel active tips ($f_\mathrm{r} = 0.2$, Figure S5G). One should also note that one must have $R_\mathrm{r} > R_\mathrm{a}$ in the simulations, i.e. a larger self-avoidance radius than the self-annihilation radius, otherwise tips terminate before they get a chance to sense their neighbours, and adapt their trajectories accordingly. In the simulations of Figure S5F-J, we have used $R_\mathrm{r} = 2R_\mathrm{a}$.

When running these simulations, we found that simulated glands with higher degrees of repulsion grow to larger subtree sizes and density, as expected from their repulsive properties, which allowed them to explore space more efficiently before ter-

minating. However, the kinetics of convergence towards balance of tip termination and branching is only very slightly delayed (one generation, Figure S5H). This is in large part because, in two-dimensions, particles still annihilate at comparable rates given the high probability of crossings.

However, one quantitative change arising from repulsion is the enhanced emergent anisotropy of the branching random walk. To quantify this, we performed the same quantifications of branch angle relative to the horizontal, as in the reference simulation (Figure 4C), and compared the reference angle distribution with the distributions in the different repulsive conditions. For small f_r, we found that the bias towards distal orientations (angles close to 0) increased with increasing repulsion f_r. Moreover, including the repulsive effect of the boundary increased the bias even further (Figure 4D). This is expected, as boundaries are now expected to "guide" via repulsion the tips in the proximal-distal direction. Interestingly, this thus predicts anisotropies which are larger than the average experimentally observed value, arguing that this might not be a key feature of the data. Moreover, the reference simulation already achieves a relatively large value of anisotropy in the absence of any repulsion. This indicates that self-avoidance could reinforce the emergent effects that we observe from branching and annihilating random walks, but are not necessary qualitatively and quantitatively to explain them. Of note, we did observe in our dataset some inter-gland heterogeneity in the branch angle distribution anisotropy, with some glands displaying much more directional invasion than others. This could potentially be explained by various strengths of the self-avoidance properties in different glands. Finally, we noted that self-repulsion decreases density fluctuations (Figure S5F,G), which can be observed quantitatively by a reduced exponent of giant number fluctuations (Figure S5J), as expected from an

ordering repulsive mechanism.

2.2   Analysis of kinetics of mammary gland invasion

Here, we expand on the theoretical results shown in Figure 3 of the main text. As described above, in addition to making predictions on the structure of the final mammary glands, our theory makes key predictions on the kinetics of its spatial patterning, with in particular a self-organisation of tips in a KPP pulse of asymmetric shape, travelling at constant speed. We therefore sought to verify them by performing EdU pulses in 5w-old mice, and sacrificing the mice 4h later, to catch all of the cells which were proliferative (or rather in S-phase) at that time. As shown previously (Scheele et al., 2017), the overwhelming bulk of proliferative cells were localized in tips, with very few in mature ducts. Moreover, we examined both fourth mammary glands ($n = 2$) and third mammary glands, ($n = 2$), and found consistent results, and in particular that active tips localized as a pulse at the edge of the invasion front, as predicted theoretically.

One should note that although the bulk of proliferative cells are localized in this pulse, one could also observe a fraction of tips which still contained a few proliferative cells, and were localized further away from the invasion front. However, these tips were always less proliferative than at the front, and were small, lacking the characteristic "balloon shape" of front tips (Figure S4I), i.e. the so-called terminal end buds in the literature. We also noted that these small "half-active" tips were typically localized at the edge of the fat pad. It is therefore tempting to speculate that they represent formerly active tips during their transition to termination. One can add them to the description (which is formally equivalent to describing the localization of newly formed nephrons in Kidney, as shown in Figure 6D and discussed below) by writing an additional equation

for their concentration h as:

$$\begin{cases} \partial_t a = \partial_x^2 a + \bar{r}a(1 - a - i - h) \\ \partial_t h = a + \bar{r}a(a + i) - r_h h \\ \partial_t i = r_h h \end{cases} \tag{14}$$

where r_h describes the rate of transition between half-active and inactive particles. Within this framework, h adopts a spatial shape very similar to the one of active tips (as its equation is slave to a as soon as $h \ll i$, which is typically the case, in common with the limit described in the section above).

Although r_h could be fit from the spatial distribution of half-active tips, its inclusion here would unnecessarily over-complicate the analysis. Instead, in Figure 3D, we limit attention to the concentration of active tips, i.e. tips which consist of over 50% of proliferative cells (as assessed by EdU expression). Importantly, even with this conservative definition of active tips, one still recapitulates a key prediction of the model regarding the asymmetry of the pulse: the back decay is much slower than the front decay. Moreover, although the model slightly underestimates the density of ducts and active tips (Figure 3D), it still provides a very good prediction for the detailed shape of both, with exponential decays on both sides of the pulse. In the front in particular, we found that both tips and ducts decay with the same slope, as predicted. One should note that, for the simulations underlying the theoretical curves of Figure 3D, we fit the relative sizes of the fat pad by measuring it once again for the $n = 4$ glands used in the EdU experiment, as these were larger than the previous dataset (by around 60%), all other parameters being kept constant. Again, however, this did not change the conclusion nor the simulated profile shapes of the glands.

Altogether, our modelling, combined with whole-gland reconstruction and EdU as-

says suggest a detailed mechanism through which mammary tips are able to self-organise into an invasion front and pattern the gland at constant density. In particular, this does not need any encoded gradients to serve as a guide, as the anisotropy of the invasion simply derives from the initial condition, i.e. the position of the first mammary tip in the fat pad.

3 Alternative models of two-dimensional branching morphogenesis

To underline the singular predictive nature of the BARW model, in this section we consider the behavior of different models of branching morphogenesis inspired by past proposals in the literature. In particular, we find that these alternative models are inconsistent with various aspects of the data, both qualitatively and quantitatively.

3.1 Fractal patterning

Fractal patterns have been proposed in the past to explain the branching morphogenesis of the lung (Iber and Menshykau, 2013) and human mammary lobule (Honeth et al., 2015). The emergence of fractality finds its origin in the aforementioned problem of space filling for exponentially increasing tip numbers (if they simply divide symmetrically). In our framework, the problem is abolished by termination, which regulates the epithelial density. In the fractal branching concept, branch length and width diminishes geometrically as a function of generation, so that an exponentially increasing number of branches can fit in a finite space (it is not yet clear how such rules would emerge and be implemented biologically). Importantly, key lines of evidence argue against such a

model in the mammary gland (as well as the kidney):

- The branching pattern is not observed to be deterministic (viz. subtrees do not show homogeneity in size) in these organs.

- The branching pattern cannot be explained by purely symmetrically dividing tips (Figure 2B-D and Figure S1), i.e. the number of tips does not increase exponentially with generation number.

- Average branch lengths and widths are seen to be independent of generation number in the mammary gland (Scheele et al., 2017) (while it reaches a plateaus in the kidney).

3.2   Branching and self-avoiding random walks

For kidney branching morphogenesis, a model of branching tips with self-avoiding properties was proposed and compared to the first branching events in *in vitro* experiments (Davies et al., 2014). An advantage of such a model to understand mammary morphogenesis is that it avoids crossovers between ducts, thus reproducing a key aspect of the experimental data. However, a major contrast with our framework is that the number of tips would still increase exponentially with generation number, as there is no termination. Therefore, this model cannot explain the data, as the branching pattern cannot be explained by purely symmetrically dividing tips (Figure 2D). Moreover, as mentioned above, a large degree of self-avoidance increase significantly the anisotropy of the branching pattern to values which exceed the experimentally measured mammary anisotropy (Figure S5F,G,I).

This points to the fact that tip annihilation must be a key feature of any viable model

of branching morphogenesis. However, we note that alternative sources of regulation/termination could be conceived over that considered in the main text.

3.3   Regulation of branching and/or tip-tip interactions

Although we have so far placed emphasis on tip termination, the balance between proliferation and termination could be enforced in a density-dependent manner. From a mean-field perspective, these two options are formally identical and, therefore, cannot be distinguished. However, in full spatial simulations, the two hypotheses yield crucially different outcomes: Although a density-dependent regulation of branching can reproduce the constant average density of ducts (Figure S4A) as well as the pulse kinetics of invasion, it yields numerous crossovers (on the order of the number of branches, thus significantly overestimating the experimentally observed value of Figure 1C by an order of magnitude).

Alternatively, we have also already explored above the the canonical form of BARWs discussed in the physics literature, which posit that termination occurs only when tips meet (instead of a tip and a duct). However, we show that this fails to reproduce even qualitatively both the density profile of tips and ducts (tips form a front instead of a pulse, and ducts are patterned according to a density gradient) and the absence of crossovers between ducts (Figure S4C,D). Therefore, this model does not reproduce the most elementary qualitative features of the data.

3.4   Stochastic termination, independent of spatial cues

Next, we explored the assumption that the frequencies of tip termination and branching are encoded intrinsically, independent of any spatial information in the system. We

thus parametrize termination based on the experimentally measured data (Figure 2D), and implement it as a stochastic process, similar to the spatially-averaged model from (Scheele et al., 2017).

Again, although the model faithfully "reproduces" the convergence towards balance by construction, it also fails to predict key qualitative aspects of the data. In particular, because it does not take into account spatial cues, it fails once again to reproduce the absence of ductal crossovers. Moreover, because of the absence of self-organizing properties arising from tip-duct interaction, active tips fail to self-organize into a sharp pulse of active tips at the edge of the tissue (see Figure S4B for a typical numerical integration), and are instead uniformly dispersed in the fat pad. Interestingly, because invasion is not directional anymore, the trees fail to fill the entire fat pad, instead exploring the proximal part before stochastically terminating. Therefore, we get a very poor fit to the angle probability distribution ($R^2 = 0.24$).

The models above fail in particular because they do not take into account properly tip-duct interaction, strengthening our case that these interactions must form the regulatory basis of any viable model of mammary branching morphogenesis. In the following, we will explore models that keep tip-ducts interaction in one form or another, to ensure that the ductal trees do not display crossovers, i.e. models that reproduce, at least qualitatively, some basic aspects of the data. We will see, however, that this still leaves room for widely different models, that we compare quantitatively to our data to show whether and how they can be excluded.

3.5 Branching and self-avoiding random walks with side-branching

One possible way to address the problem of tip termination above would be consider a model based on a self-avoiding random walk with side-branching. This would correspond to introducing two "classes" of tips and branches: the main branches/tips are the same as that described above, and perform a branching and self-avoiding random walk, maintaining a sufficient spacing between them to avoid termination. However, we further assume that upon each branching event, there is a probability r_s to initiate side-branches, i.e. tips that are unable to branch again, and terminate after a short typical length, regardless of the external environment.

This would thus be an alternative framework where termination is implemented "intrinsically" upon asymmetric branching, in opposition to the BARW framework above, where termination is an extrinsic event based on local spatial rules and independent of the branching events themselves. Indeed, the potential contribution of side-branching has been hotly debated in the mammary field in particular (Sternlicht, 2005), as well as its timing during pubertal development and pregnancy. Because of the self-avoiding nature of the dynamics, crossovers are avoided and, because of intrinsic termination, for $r_s = 1$, the number of tips does not increase exponentially, and is intrinsically balanced.

We therefore explored the possibility of mammary morphogenesis occurring sequentially, in the spirit of proposals such as in (Huebner and Ewald, 2014): An early phase of largely symmetric branching serves to increase the net number of tips, as observed experimentally, followed by a phase of purely asymmetric side-branching, with main branches avoiding each other. For robustness, we tested two options for the average threshold generation number to move to a phase of side-branching: generation

$n = 5$ (Figure S3G and blue curves in Figure S3I-K) and $n = 6$ (Figure S3H and orange curves in Figure S3I-K), and averaged in each case 1000 full stochastic simulations. We took a large radius of repulsion $R_a = L_x/20$ and repulsion strength $f_r = 0.6$ (as otherwise ductal crossovers would occur with high probability). With these parameters, one can obtain a satisfactory fit to the subtree persistence (Figure S3K).

Importantly however, this model missed key aspects of the experimental data:

- At these levels of repulsion (again, which are necessary to avoid crossovers), the simulated ductal trees were highly directional, with an angle distribution that fits very poorly with the data (Figure S3I, $R^2 = -4.1$ and $R^2 = -2.8$ resp.), as it overestimates drastically the degree of directionality (as is obvious from the simulation snapshots).

- Moreover, because of the absence of strong competition between subtrees, the predicted subtree distribution fit poorly with the experimental data (Figure S3J, $R^2 = 0.4$ and $R^2 = 0.7$, respectively), as it underestimates drastically the subtree size heterogeneity. Note that calculating R^2_{log} would provide an even worse correlation, as the tails of model and data distribution diverge strongly.

3.6 Branching and annihilating random walks with external guidance

Next, we considered a key alternative to our model, which involves the guidance of tips by external chemical gradients/cytokines. Although such gradients have not been found in the mammary gland, *in vitro* experiments have shown that mammary organoids are seemingly attracted by beads soaked in FGF10 (Zhang et al., 2014),

raising the possibility that such guidance cues may be operative *in vivo*. This question is particularly topical, as such collective migration via external guidance is a key feature of many morphogenetic events, reviewed for instance in (Huebner and Ewald, 2014).

More generally, whether morphogenesis proceeds via self-organizing features, or via decoding positional information has been a fundamental debate over the past decades in a variety of system, and is still largely unresolved (Kerszberg and Wolpert, 2007). We thus explored the possibility of external gradients guiding mammary morphogenesis. Specifically, we retained the BARW framework defined above, but imposed an additional external guidance, i.e. a biased instead of isotropic random walk. More precisely, at each step of a tip i in a direction $\mathbf{p}_i$, we imposed an external field in the distal direction $g_x \mathbf{x}$, with $g_x > 0$ (and calculated the updated unit vector $\mathbf{p}_i$ in the presence of this field). Thus, $g_x = 0$ is our classical BARW framework developed above, while increasing values of g_x lead to more and more directional migration towards the distal direction. In particular, we found that the effect of the perturbation became noticeable when g_x is of a comparable order of magnitude to the rotational diffusivity of the tips (Figure S3A-C). We thus tested various values of $g_x = 0.05, 0.1, 0.2$ (respectively blue, orange and yellow lines in Figure S3D-F) and averaged in each case 1000 full stochastic simulations.

Importantly, this model also missed key aspects of the data:

- The main departure between simulation and experiment was, predictably, on the anisotropy of the branching pattern. We computed for each value of g_x, as for the BARW simulations in the main text, the probability distribution of having a branch at an angle $\theta \in [0, \pi]$ from the horizontal $\mathbf{x}$. Crucially, this systematically

81

overestimated the anisotropy, with increasing values of g_x causing an increasing anisotropy, and providing a very poor fit for the data (Figure S3D, $R^2 = -2.8$ for the best fitting case of $g_x = 0.05$, thus a much worse fit than the control case of $g_x = 0$).

- Moreover, we also found that the predicted subtree size distribution (Figure S3E) and persistence (Figure S3F) fitted increasingly poorly the experimental data (respectively $R^2 = 0.82$ to $R^2 = 0.56$ and $R^2 = 0.93$ to $R^2 = 0.83$, for $g_x = 0.05$ and $g_x = 0.2$). Note that, calculating R^2_{log} would indicate an even worse fit, as the tails of model and data distribution diverge strongly.

Finally, we sought to test more directly the existence of guiding chemical gradients by performing a quantitative analysis of transplant experiments, as detailed in the main text. Crucially, this showed that mammary stem cells transplanted in the center of a fat pad could form branched structures which could propagate isotropically in the fat pad, without large distal-proximal bias. This validates directly and quantitatively an absence of a pre-determined chemical gradient, and lends strong support to our model of emergent directionality arising from isotropic and local interactions.

3.7   Branching and annihilating random walks with side-branching

Next, even though this possibility is very close to the one considered in the main text, we examined whether the data could accommodate a BARW framework coupled with significant side-branching (defined as above as the generation with probability r_s of tips that are unable to undergo further branching). We therefore explored various values of r_s, all of the other parameters being the same as the default simulations from the main text, and averaged in each case 1000 full stochastic simulations (Figure S3L-N).

Importantly however, large values of r_s again missed key aspects of the data. In particular, increasing values of r_s caused a higher and higher value of the subtree persistence, as the perturbation decreased inter-tip competition. Therefore, the simulations with side branching systematically overestimated subtree persistence (Figure S3O). At the same time, they also systematically underestimated subtree size heterogeneity, for the same reason as above (Figure S3P). Thus, including medium to large values of r_s ($r_s = 0.2$ in blue, $r_s = 0.5$ in orange and $r_s = 0.75$ in yellow) systematically worsened the fit of the model to the data ($R^2 = 0.91$ for persistence and $R^2 = 0.85$ for size distribution for $r_s = 0.5$), allowing us to rule out the possibility of a large fraction of side-branching during 3-8 weeks pubertal dynamics. We note, however, that it has been reported that side-branching could play a more dominant role later on after the fat pad has been filled (Sternlicht, 2005).

3.8   Branching and annihilating random walks with large side-branching and repulsion

Finally, we considered whether a large amount of "secondary phase" side-branching could still accomodate the data if we added repulsion (again, in the presence of annihilation as in the control simulation). We thus implemented the same two-phase model of Section 3.5 (symmetric branching followed by a phase of side-branching with probability r_s after an average generation of $n = 6$), again with repulsion. We used here a repulsion radius of $R_a = L_x/60$ and a repulsion strength $f_r = 0.6$ (although it should be noted that these are not as crucial as in Section 3.5, as crossover here is avoided in any case by the presence of annihilation).

Again, to explore the phase diagram, we tested various values of $r_s = 0.75, 0.9, 1$

(respectively blue, orange and yellow lines in Figure S3R) and averaged in each case 1000 full stochastic simulations. We show in Figure S3Q an example of a typical final configuration for $r_s = 0.9$. Here, although the model reproduces rather well the orientation of the tree, as well as its constant spatial density and absence of crossovers, the key issue is again the low competition between subtrees, which causes a overestimate of subtree persistence, and an underestimate of the heterogeneity of the subtree size distribution (Figure S3R, $R^2 = 0.82$ for $r_s = 0.75$ to $R^2 = 0.82$ for $r_s = 0.9$ and $R^2 = 0.36$ for $r_s = 1$). Note that, once again, calculating R^2_{log} would indicate an even worse fit, as the tails of model and data distribution diverge strongly.

Therefore, while one can conceive of various alternative kinds of model to the BARW with varying degrees of complexity, none of them appear to reproduce the full dataset, reinforcing the seemingly unique predictive power of the BARW framework proposed in the main text. In particular, these analyses emphasize that tip termination/annihilation is a key feature that any viable model must possess to understand the data. More surprisingly, we note that more complex models (for instance involving time-varying rules, side-branching and repulsion) do a worse job at predicting the data that the simpler BARW model we propose.

4 Molecular basis for branching and termination *in vivo*

In this section, we discuss the rationale behind the perturbation experiments that we performed to challenge our model, in particular the implantation of soaked beads during pubertal branching morphogenesis to provide ectopic supply of various positive or negative regulators of mammary branching. These experiments, shown in Figure 5 and Figure S5L-P, provides some hints on the underlying molecular regulation of the

events of branching and termination, that we have so far addressed from a mesoscopic perspective.

4.1 Details on the experimental approach

As discussed in the methods, soaked beads were surgically implanted on top of the fat pad of third and fourth mammary glands, so that they did not physically prevent the invasion of the mammary epithelium, but still delivered proteins in its vicinity. This technique was used in particular with success in the context of chick limb development (Fallon et al., 1994; Raspopovic et al., 2014). As a control that beads did not physically prevent invasion, we implanted control beads soaked in PBS, and verified on whole mounts that branches could grow below them (i.e. were overlayed in the maximum intensity projections shown in main and supplemental Figures). To investigate both positive and negative regulators of branching morphogenesis, we soaked beads in either:

- TGF (Transforming Growth Factor)$-\beta$, which has been widely studied in the context of mammary development as the key negative regulator of mammary growth, both *in vitro* (Nelson et al., 2006) and *in vivo* (Silberstein, 2001), since the seminar experiments of (Silberstein and Daniel, 1987) of implanting a large TGF-β releasing pellet to show that it inhibits mammary invasion. However, it is still not fully clear from a quantitative perspective if TGF-β acts mainly as a repulsive cue, a negative regulator on branching, or a positive regulator on termination (Daniel et al., 1989; Kahata et al., 2017), and what is the characteristic length scale of this regulation.

- FGF10 (Fibrobastic Growth Factor 10), which has been identified *in vivo* as a key

regulator of branching outgrowth in lung, prostate, salivary gland and pancreas (Iber and Menshykau, 2013). FGF10 is the predominant stroma FGF ligand expressed during pubertal mammary morphogenesis, and directs mammary cell migration *in vitro* (Zhang et al., 2014), raising the hypothesis that it can be involved in branching initiation and guidance *in vivo*. However, its effect has not yet fully been elucidated in the mammary gland *in vivo*, and in particular, a quantitative understanding of its contribution in terms of elongation, guidance and/or branching remains in question.

Therefore, in addition to providing molecular mechanisms, such experiments have the potential to clarify aforementioned issues of the relative contributions of guidance and/or repulsion compared to branching and annihilation.

4.2 Details on the theoretical approach

In order to theoretically simulate the effect of the beads with respect to our BARW framework, we first considered three possibilities: the beads would enhance locally termination (in that case, we use the same termination radius for beads as ducts for the sake of simplicity), that beads would inhibit locally termination (in that case, we implemented that within the same radius, no termination could occur), and that beads would enhance branching (in that case, we implemented that within a critical radius, the branching rate was enhanced by a given factor h_b, with $h_b = 2$ in Figure 5C, corresponding to the observed increase in branch rate, and $h_b = 3$ in Figure S5L for comparison).

We randomly positioned in simulations four clusters of 15 beads, to mimic the experimental configuration. A first thing we asked was the probability of bead-duct cross-

over in all cases, to give a better perspective of how much cross-over is expected in the control case. As shown in Figure S5L,O, in the control case of beads devoid of any effect, the probability for cross-over was slightly below 40%, whereas by construction, the case of beads inducing termination displayed a zero-probability for bead-duct cross-over.

4.3 Results

Importantly, when examining the mammary gland morphology in the presence of TGF-β soaked beads, a first and key finding was the absence of any overlay between a bead and a duct (Figure 5A and Figure S5M,O, $n = 4$ glands, around 50 beads in each). In particular, in regions of high bead density, the fat pad remained entirely devoid of any epithelial invasion, while in regions of lower bead density, one could observe tips having stopped just before encountering the beads (with a characteristic distance of $100 - 200\mu$m, see blue asterisks on Figure 5A). The effect was particularly sticking and significant when compared to gland morphology in the presence of control, PBS-soaked beads, in which numerous duct-bead overlays were observed (Figure 5B and Figure S5N), with a probability around 35% very close to the predicted value from the model. This validate the assumption that control beads have no effect of the trajectory of the BARW, and are thus an accurate and non-perturbative control. Importantly, quantifying the average branchpoint to next branchpoint distance far from TGF-β1 beads demonstrated that there was no statistical different on the branching far from the beads (Figure S5P). Comparison with the TGF-β1 beads thus demonstrates the locality and effect of TGF-β1 in tip termination.

Next, we assessed the influence of ectopic delivery of FGF10 through the same

soaked beads approach. In proximity to the beads, numerous events of side-branching could be observed (Figure 5C), as well as much shorter branchpoint to branchpoint distance. To quantify this effect, as well as its locality, we measured the average branchpoint to branchpoint distance in four settings: in regions close and far from the beads (we calculated "close" to the bead region as regions which are closer than 500μm to a bead), both for control beads and FGF10 soaked beads. Importantly, when comparing glands with control beads vs glands with FGF10 beads ($n = 3$ glands for FGF10 beads, $n = 3$ glands for control beads), we could find no statistically significant difference between the branch length far from control beads, in proximity to control beads, or far from FGF10 beads (adjusted $P > 0.99$ in all cases, from a Kruskal-Wallis test with multiple comparisons and correction). However, the branches close to FGF10 beads were statistically much shorter (approximately twice shorter, $P < 0.001$ from a Kruskal-Wallis test, in all comparisons with the three controls). We thus compared these glands with FGF10 beads with numerical simulations locally enhancing branching rate by a factor $h_b = 2$ (Figure 5C), which reproduced well the enhanced branch density locally with minimal directionality.

5 Sensitivity analysis - effect of parameter variations on the predicted structure of the kidney

Having considered the application of the BARW model (and alternatives) to the mouse mammary gland epithelium we now turn to consider in detail the modelling approach used to simulate the three-dimension branching morphogenesis of the mouse kidney. In particular, we consider how the three-dimensionality of the kidney influences the

analysis of the BARW model dynamics.

5.1   Details on the 3D numerical simulations of the BARW model

To address the dynamics of the BARW in 3D, we parametrise the model in spherical coordinates (r, θ, ϕ), with the initial condition of a single active tip positioned at coordinate $r = 0$ with orientation $\theta = \pi/2$ and $\phi = 0$. Tips branch at a rate r_b and terminate deterministically if they enter into a radius of R_a of another duct. Upon bifurcation of a tip i, two offspring tips are produced, at an angle α_i from their ancestor duct. Based on previously published data on kidney branching, α_i is chosen randomly from a uniform interval $[\alpha_o - \delta\alpha, \alpha_o + \delta\alpha]$, with $\alpha_o = 50°$ and $\delta\alpha/\alpha_0 = 1/3$ (Short et al., 2014). One should note that the results depend only very weakly on these parameters. Moreover, the two offspring tips adopt experimentally a diametrically opposite position, and we therefore pick randomly the angle $\beta \in [0, 2\pi]$ for the plane of this bifurcation (see Figure S6A for a schematic).

The angles (θ_1, ϕ_1) of offspring 1 can thus be calculated from the angles (θ, ϕ) by solving the equations:

$$\begin{cases} \cos\theta_1 \sin\phi_1 \cos\theta \sin\phi + \sin\theta_1 \sin\phi_1 \sin\theta \sin\phi + \cos\phi_1 \cos\phi = \cos\alpha_i \\ \\ -\cos\theta_1 \sin\phi_1 \sin\theta + \sin\theta_1 \sin\phi_1 \cos\theta = \sin\beta \sin\alpha_i \end{cases} \tag{15}$$

and the angles (θ_2, ϕ_2) of offspring 2 can be calculated by the same system of equation (15) by substituting $\beta \to \beta + \pi$. This guarantees that there is both an angle of α_i between the directions of ancestor and offsprings, as well as that the two offsprings go in diametrically opposite directions relative to this angle.

Furthermore, as detailed in the main text, one must specify the anisotropy of kidney expansion in the simulations (Figure S6B-C), as this does not appear to arise from the

redirection or annihilation of tips going along the short axis. Indeed, if that were the case, one would see, by analogy to the mammary gland expansion, that only tips along the long axis are proliferative. In contrast, both proliferation and nephrogenesis in kidney are organised in continuous rims all across the surface of the developing organ (Costantini and Kopan, 2010). Anisotropy does not seem to be driven either by a much shorter branch length along the short axis, as we found that the average branch lengths depends very little on the direction of growth (Figure S6E), whereas the branch length distribution was consistently well-fitted by a single exponential with a threshold (which we fit from the experimental value, see theory versus experimental distributions on Figure S6F). Therefore, the simplest model of anisotropy is to assume that the time scale of the branching and elongation process depend on the angles (θ, ϕ) of a given branch relative to the centre of mass, being rescaled by a factor $\mathscr{F}(\theta, \phi)$. We denote L_z the longest axis ($\theta = 0$), L_x the shortest axis ($\theta = \pi/2, \phi = 0$), and L_y the intermediate axis ($\theta = \pi/2, \phi = \pi/2$), which we measure experimentally as $L_x/L_z = 0.5$ and $L_y/L_z = 0.4$. We thus parametrise kidneys as triaxial ellipsoids using the formula:

$$\mathscr{F}(\theta, \phi) = L_z \sqrt{\frac{\sin^2 \theta \sin^2 \phi}{L_x^2} + \frac{\sin^2 \theta \cos^2 \phi}{L_y^2} + \frac{\cos^2 \theta}{L_z^2}}. \qquad (16)$$

Finally, although the global distribution of branch lengths in kidney is well-fitted by a single exponential (Figure S6D,F), indicative of a stochastic branching with constant rate r_b, it has been shown that the first bifurcations are more regular and stereotypic (Sampogna et al., 2015; Short et al., 2014), intervening dominantly in a single plane and along the future long axis of the kidney. To be able to make accurate experimental predictions later, we thus grew a "seed" kidney for $n = 5$ generations of bifurcations, at a deterministic branch length of $1/r_b$, before using the stochastic branching dynamics

described above. (Note, however, that we still allow for annihilation in the early phase, with the same annihilation radius, for the sake of simplicity and consistency.) Such a modelling is clearly a simplification, as finer dependencies of the branch length and width with generation number have been reported in the past (Short et al., 2014). However, our goal here is to define basic design principles of kidney growth, and we thus neglect such "second-order" features.

5.2   Sensitivity analysis of the final kidney morphology

In this section, we turn to a sensitivity analysis of 3D BARW to understand how variations in the parameters of the model affect quantitatively and qualitatively the results that we observe for kidney morphogenesis.

5.2.1   Key parameters for the default kidney simulations

Once again, before performing a sensitivity analysis as a function of parameter values, we list the default parameter set used in the main text for the numerical simulation of kidney morphology. Numerical simulations were performed in three-dimensions. At each time step (with $\tau = 1$ again used as the unit time), tips move forward by a length $V\tau = 1$ (again used as the unit length), along a polarity vector $\mathbf{p}_i$, or angle θ, which diffuse by a random angle of amplitude $\delta\theta = \pi/10$ (as in the case of mammary gland simulations). We set the branching rate to $r_{\mathrm{b}} = 0.2$ and finding a best-fit value for Figure 6D for the annihilation radius of $R_{\mathrm{a}} = 1.3$ in simulation length units (so that the only key parameter controlling the kidney morphology is $R'_{\mathrm{a}} = r_{\mathrm{b}}R_{\mathrm{a}} = 0.26$). Note that the bias towards ductal termination at the largest generation numbers is partially an artifact of the analysis, as these represent tips at the periphery of the kidney (Figure 6E) that

remain proliferatively active but, without "progeny" yet, are counted as "terminated".

5.2.2   Mapping between simulated and real embryonic time

In contrast to the mammary gland, where we only investigated either the final state of the mature organ, or the steady state of the invasion front, here we wished to understand quantitatively the full developmental time course of embryonic kidney formation (as assayed experimentally by (Sampogna et al., 2015), the results of which we reanalyzed here). As shown previously (Sampogna et al., 2015), the speed of kidney branching morphogenesis (defined as the characteristic time to form a branch) decays in time during embryogenesis (even though the average branch size decays slowly in time). This means that simulated time will depend non-linearly on real embryonic time. For the sake of simplicity, we thus assumed constant time-invariant properties in the simulations, and then linked simulated time to embryonic time by fitting the average number of branches at that developmental time (from E11 to E19, see Figure S6G). All of the results presented in Figure 7A,C,D are thus obtained using the simulated embryonic time correspondence shown on Figure S6G.

5.2.3   Giant density fluctuations in kidney

In order to assess giant number fluctuations in the experimental data, without being corrupted by boundary effects, we only examined the detailed reconstructions of E17 to E19 kidneys ($n = 3$, average shown on Figure 7E; results for different time points were consistent, as indicated by the small error bars), performing the analysis in a rectangular three-dimensional box around the center of gravity of each kidney, that thus excluded boundary effects. Results consistently showed robust power law de-

pendences with exponents larger than 0.5, indicative of giant number fluctuations. This indicates the generality of the results uncovered in mammary gland morphogenesis. We then performed the same analysis for the density fluctuations of the simulated kidney, on E19-equivalent trees, again avoiding the edges of growth to prevent boundary effects from corrupting the data.

5.2.4 Effect of the annihilation radius

In contrast to the case of two-dimensional BARWs, the value of the annihilation radius R_{a} plays a crucial role in three dimensions. As mentioned above, this is because vanishingly small radii R_{a} give rise to vanishingly few crossovers/terminations in three dimensions. Thus, the radius R_{a} becomes a key parameter, and has to be fitted with respect to the kidney data in order to be able to make quantitative predictions, as discussed in the main text and in the sections above.

We therefore wished to understand how variations of this parameter translate to the morphology, heterogeneity and detailed topology of the kidney simulations. First, we suppressed all annihilation events ($R_{\mathrm{a}} = 0$) to check whether the heterogeneity of the branch level distribution could stem from purely stochastic branching and size anisotropy. Importantly, this provided a very poor fit to the data (first panel, Figure S6J), showing the importance of annihilation for heterogeneity.

Next, we performed the same three-dimensional simulation as in the control for large ($R_{\mathrm{a}} = 0.5R_{\mathrm{a}}^{\mathrm{ref}}$) or small ($R_{\mathrm{a}} = 2R_{\mathrm{a}}^{\mathrm{ref}}$) values of the annihilation radius (respectively centre left and centre right on Figure S6I, to be compared to the left panel for control). As expected, large radii enhance the heterogeneity of the simulated kidneys, as seen by a broadening of the segment distributions as a function of generation and at differ-

ent developmental timings (second panel on Figure S6J). This translates into a larger nephron to active tip ratio (Figure S6K), largely overestimating the experimentally observed values. Conversely, small radii decrease the heterogeneity of the simulated kidneys (third panel on Figure S6J), and underestimate the nephron to active tip ratio (Figure S6K).

Interestingly, one can observe that such changes in the radius R_a produces large variations in the predicted observables of kidney morphology, and both alternative values of R_a provide worse fits for the current data set. This confirms the importance of this parameter, and argues that it can be estimated rather precisely from our fitting procedure.

5.2.5 Effect of additional self-avoidance

Although the model performs well to reproduce key features of the kidney structure, topology and nephrogenesis pattern, as is obvious qualitatively from the three-dimensional reconstructions, and quantitatively from the giant density computation of Figure 7E, it overestimates the spatial density fluctuations, so that experimental reconstructions were consistently more "ordered" than their simulated counterpart. As mentioned above, it has been proposed that kidney has self-avoiding properties (Davies et al., 2014), in a Bmp7-dependent manner, as inhibition of Bmp7 function in cultured kidneys causes collisions between tips and ducts. However, another interpretation of this data, would be that collisions are avoided instead by termination, rather than repulsion. Notably, evidence in favour of this is obtained from culture studies of two kidney buds in close proximity (reproduced in Figure S6M from (Davies et al., 2014)). Indeed, when we simulated two trees growing in a similar geometry to the experiments, both for the

case of pure termination without repulsion (Figure S6N) and pure repulsion without termination (Figure S6O), we found that the simulations with termination reproduced better the presence of numerous (terminated) tips in the contact zone between the two kidneys, whereas the pure repulsion simulations displayed an absence of tips at the contact zone, as they avoided the zone by adoption of diverging flow motion. These observations lend further qualitative support for the regulatory basis of the current modeling scheme.

As mentioned in the sections above, tip repulsion without termination cannot explain the experimental data. However, to expand our analysis, we questioned whether, when applied to the *in vivo* kidney data, a degree of self-avoidance coupled to the BARW framework would improve the theoretical predictions. As the avoidance strength f_r is increased, the number of annihilation events diminishes, and the branching topology thus becomes more deterministic, and characterised by more peaked distributions. However, this could be counteracted by increasing in concert the annihilation radius R_a. In particular, we found that for an annihilation radius of $R_\mathrm{a} = 2R_\mathrm{a}^\mathrm{ref}$ and $f_\mathrm{r} = 0.3$ together with a repulsion radius of $R_\mathrm{r} = 6$, all other parameters being maintained the same (see Figure S6I for a typical simulation output showing higher order than control simulations), we could still obtain good fits for the nephron versus tip number as a function of time (Figure S6L), as well as for the number of branches per generation distributions at all time points (Figure S6J, bottom). Moreover, in these simulations, the spatial variations in density were markedly reduced (Figure 6E), so that the exponent of the giant number fluctuations observed experimentally could be predicted with much better accuracy compared to the non-repulsive case. This argues that some self-avoidance may cooperate with annihilation in kidney morphogenesis, and defines

quantitatively tools to assess both contributions from *in vivo* data.

5.2.6   Effect of the growth anisotropy

We then turned to explore the contribution of the anisotropy in promoting subtree heterogeneity in the kidney. We therefore performed numerical simulations with the same set of parameters as wild-type, only in isotropic growth conditions ($L_x = L_y = L_z$, Figure S7A). As expected, the key phenomenology of a self-organised pulse of active tips at the edges of the kidney was not affected by changes in isotropic conditions (Figure S7A). Moreover, we observed similar kinetics and scaling laws in the number of nephrons versus tips as a function of time. A key difference was observed in the number of branches per generation as a function of time. Indeed, we observed a marked reduction in the width of the distribution, as expected from reducing the anisotropy-induced growth advantage of subtrees in favorable directions (Figure S7B).

5.3   Proximity to the critical point

Finally, we wished to assess systematically how the stochasticity of branching morphogenesis could lead to a stochastic transition to a fully annihilated state. Indeed, within the mean-field theory, annihilation of the entire tree can never happen, as any non-vanishing value of the branching rate is sufficient to yield a steady state non-zero density of active tips. However, fluctuations, when taken into account through the full numerical simulations, are able to destroy this active steady state, implying a non-zero probability for full tree annihilation.

We thus varied the frequency of termination by performing a parameter sweep in the annihilation radius R_a. (Although the branching probability r_b could also have been

used as a converse variable, this is computationally more intensive as it requires simulating many particles for the low branching rate limit.) We otherwise used the same parameter set as in the control kidney simulations (non-repulsive), although we used three-dimensional isotropic simulations without a deterministic seed here, in order not to confound the analysis. For obvious technical reasons, we cannot simulate kidney for an infinite amount of time, so that finite-size effects are expected, i.e. full annihilations which might occur later than the threshold simulation time are discarded. However, the frequency of such events become vanishingly small in time, and only matter in the very close vicinity to the critical point. We used the criteria that trees reaching $100,000$ particles are considered non-annihilated/survivors, ran at least 1000 simulations per parameter value (see Figure S7E-G for representative examples), and calculated the tree survival probability as a function of the annihilation radius.

Importantly, and in qualitative agreement with the literature (Cardy and Tauber, 1996), we found that above a threshold in the annihilation radius R_a, the tree survival probability vanishes, while it continuously increases below this threshold, indicative of a continuous phase transition (Figure 7F). Below a secondary lower threshold, we also noticed that survival always occurs. Interestingly, both these thresholds are larger than the control value used to fit the *in vivo* kidney data (the same is true of mammary glands and two-dimensional simulations), by ratios of 2 and 3, respectively (Figure 7F, dashed vertical line). This is expected from a biological perspective, as one wishes to avoid tree extinction before it grows to a mature size. However, as discussed in the main text, in mutant conditions, the value of the annihilation radius could be closer to values allowing for tree extinction, a hypothesis to be confirmed in subsequent studies which could shed new light on branching pathologies.

6   Vitamin A deficient kidney

In this last kidney-related section, we now apply the analyses above to understand per-turbations to normal kidney development. In particular, we re-analyzed a data set from the segmentation of the E15 mouse kidney on animals maintained on a low vitamin A diet (Sampogna et al., 2015). This phenotype is particularly interesting because, although the branching rate and total number of branches are maintained in this con-dition, it was found that kidney became much more heterogeneous with, in particular, an increase of the maximal branch generation number attained at E15 (some 3 gener-ations more than wild-type, $P < 0.05$), and thus a correspondingly decreased number of mid-generation branches.

We therefore sought to test whether our model could be applied to understand this new phenotype. Given the sensitivity analysis developed above, we first tested whether a change of anisotropy in kidney shape could explain this behavior, but found no significant changes in the kidney aspect ratio compared to wild-type ($P > 0.1$ both for L_x/L_z and L_y/L_z). However, a key geometric change was that Vitamin A deficient kidney was much smaller, behaving as a scaled down version of their wild-type counterparts (occupying on average 35% of normal wild-type volume, $P < 0.05$). Therefore, as a consequence, the density of tips at the outer surface rim was nearly twice as large in the Vitamin A deficient kidneys as wild-type.

This means that, although the branching rate is identical, the elongation rate of tips must be reduced in Vitamin A deficient kidneys, and lead us to the hypothesis that this could explain, by itself, the enhanced heterogeneity, by increasing tip-tip competition. Indeed, for lower characteristic branch lengths, crowding-induced terminations are ex-pected to arise earlier. Guided by our previous results, we therefore sought to check

whether this single geometric change in elongation rate of kidney would reproduce quantitatively the phenotypic differences between the wild-type and Vitamin deficient kidney. We thus decreased the elongation rate by a factor two in the simulations, and kept all other parameters constant, which lead to the correct reduction in kidney volume at E15 (Figure S7H,I for comparison of wild-type and mutant simulations).

We then computed by the same methods as before the predicted average number of branches per generation, for the same total number of branches. Importantly, we found that the simulations for Vitamin A deficient kidneys displayed on average two more generations than simulations for wild-type kidneys (Figure S7J), closely mirroring the data. Moreover, the model provided an overall excellent prediction ($R^2 = 0.98$, $S = 6$) for the Vitamin deficient data. This validates our model of crowding-induced termination by showing that one can predict kidney structure from simple geometric properties such as average branch length. Interestingly, one should note that, in the Vitamin A deficient kidneys, nephrogenesis is also impaired (Sampogna et al., 2015). This displays an additional layer of complexity compared to our framework, as we have made the simplifying assumption to equate inactive tips with tips fusing to glomeruli to initiate nephron formation. This hints, in particular, to the fact that crowding tip termination could be a distinct, not-fully overlapping property of nephron initiation, something that would need to be tested, for instance, by combining large-scale reconstructions with EdU assays.

7 Branching and annihilating random walks arising from a generic reaction-diffusion mechanism

In this section we investigate whether and how BARWs can be implemented from canonical, "off-the-shelf", molecular regulatory programs of the type present in epithelial cells. To this end, we made use of the framework of reaction-diffusion systems, which have been widely and successfully studied since Turing's seminal work (Turing, 1990) to understand collective biological phenomena. In the original concept of Turing, and later used to explain neural impulses, a diffusible activator A both self-activates and activates another diffusible molecule I. This second molecule I is an inhibitor, negatively regulating A as well as itself. Under general conditions on the diffusion coefficients, this system was shown to generate arrested phase separation into complex motifs with a well-defined length scale (dots on an hexagonal lattice, stripes, reconnected labyrinths, etc.).

Later, through the seminal work of Meinhardt and collaborators (Meinhardt, 1982) this formalism was extended to show that the inclusion of a third molecular component, functioning as a substrate for the first two, allowed for the formation of branched structures. Recently, Guo and colleague have revisited this paradigm (Guo et al., 2014) using the following four-species system, including an activator A, an inhibitor I, differ-

entiated inactive cells Y, and a substrate S:

$$
\begin{cases}
\partial_t A = c\frac{A^2 S}{I} - \mu A + D_A \nabla^2 A + \rho_A Y \\[2em]
\partial_t I = cA^2 S - \nu I - D_I \nabla^2 I + \rho_I Y \\[2em]
\partial_t S = c_0 - \gamma S - \varepsilon Y S + D_S \nabla^2 S \\[2em]
\partial_t Y = dA - eY + \frac{Y}{1+fY^2}
\end{cases}
\tag{17}
$$

where c, μ, ρ_A, ν, ρ_I, c_0, γ, ε, d, e and f are coefficients of interactions, and D_A, D_I and D_S are coefficients of diffusions (for details see (Guo et al., 2014)). Although this model adds quite a bit of complexity to the original Turing system, numerical analysis shows that it is able to reproduce the different modes of branching morphogenesis observed *in vivo*. In particular, increasing the parameter ε, which quantifies the consumption of the substrate by the epithelium, switches from a mode of side-branching to a mode of exclusive tip-splitting morphogenesis.

Interestingly, although the authors concentrated in their study on the mode of branching, we performed numerical integrations of these equations for longer times and noticed that the branching pattern also displayed an annihilating property (Figure S7N). Indeed, whereas classical Turing patterns reconnect, giving rise to complex multigraphs, there is a substrate here, which is continuously consumed by differentiated cells in the ductal network. As the substrate mediates the activator self-reinforcing loop in the equations for A and I, when a growing tip approaches an existing duct, it feels a substrate depletion which stops it.

Thus, these results show that phenomenology of the BARW framework can be encapsulated from a more molecular perspective from a Turing-like model from reaction-diffusion systems.

References

Brown, A. M. (2001), 'A step-by-step guide to non-linear regression analysis of experimental data using a microsoft excel spreadsheet', *Computer methods and programs in biomedicine* **65**(3), 191–200.

Brunet, E., Derrida, B., Mueller, A. and Munier, S. (2006), 'Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts', *Physical Review E* **73**(5), 056126.

Daniel, C. W., Silberstein, G. B., Van Horn, K., Strickland, P. and Robinson, S. (1989), 'Tgf-β1-induced inhibition of mouse mammary ductal growth: developmental specificity and characterization', *Developmental biology* **135**(1), 20–30.

Fallon, J. F., Lopez, A., Ros, M. A., Savage, M. P., Olwin, B. B. and Simandl, B. K. (1994), 'Fgf-2: apical ectodermal ridge growth signal for chick limb development', *Science* **264**(5155), 104–108.

Haskovec, J., Markowich, P., Perthame, B. and Schlottbom, M. (2016), 'Notes on a pde system for biological network formation', *Nonlinear Analysis: Theory, Methods & Applications* **138**, 127–155.

Honeth, G., Schiavinotto, T., Vaggi, F., Marlow, R., Kanno, T., Shinomiya, I., Lombardi, S., Buchupalli, B., Graham, R., Gazinska, P. et al. (2015), 'Models of breast morphogenesis based on localization of stem cells in the developing mammary lobule', *Stem cell reports* **4**(4), 699–711.

Huebner, R. J. and Ewald, A. J. (2014), Cellular foundations of mammary tubulogenesis, *in* 'Seminars in cell & developmental biology', Vol. 31, Elsevier, pp. 124–131.

Jensen, I. and Dickman, R. (1993), 'Nonequilibrium phase transitions in systems with infinitely many absorbing states', *Physical Review E* **48**(3), 1710.

Kahata, K., Maturi, V. and Moustakas, A. (2017), 'Tgf-β family signaling in ductal differentiation and branching morphogenesis', *Cold Spring Harbor Perspectives in Biology* p. a031997.

Kerszberg, M. and Wolpert, L. (2007), 'Specifying positional information in the embryo: looking beyond morphogens', *Cell* **130**(2), 205–209.

Ódor, G. (2004), 'Universality classes in nonequilibrium lattice systems', *Reviews of modern physics* **76**(3), 663.

Raspopovic, J., Marcon, L., Russo, L. and Sharpe, J. (2014), 'Digit patterning is controlled by a bmp-sox9-wnt turing network modulated by morphogen gradients', *Science* **345**(6196), 566–570.

Rossi, M., Pastor-Satorras, R. and Vespignani, A. (2000), 'Universality class of absorbing phase transitions with a conserved field', *Physical review letters* **85**(9), 1803.

Turing, A. M. (1990), 'The chemical basis of morphogenesis', *Bulletin of mathematical biology* **52**(1-2), 153–197.

Van Wijland, F., Oerding, K. and Hilhorst, H. (1998), 'Wilson renormalization of a reaction–diffusion process', *Physica A: Statistical Mechanics and its Applications* **251**(1), 179–201.