bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis

Ben Lambert[1,*], Adam L. MacLean[2,3,*], Alexander G. Fletcher[2,4], Alexander N. Coombes[5,6], Melissa H. Little[6,7], Helen M. Byrne[2,†]

[1]Department of Zoology, University of Oxford, Oxford, UK

[2]Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, Oxford, UK

[3]Current address: Department of Mathematics, University of California, Irvine, Irvine CA, USA

[4]Current address: Bateson Centre, University of Sheffield, Sheffield, UK

[5]Department of Anatomy & Neuroscience, University of Melbourne, Melbourne 3010 VIC, Australia

[6]Murdoch Childrens Research Institute, Flemington Rd, Parkville, Melbourne, 3052 VIC, Australia

[7]Department of Paediatrics, The University of Melbourne, Melbourne, 3010 VIC, Australia

*These authors contributed equally.

†email: helen.byrne@maths.ox.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

The adult mammalian kidney has a complex, highly-branched collecting duct epithelium that arises as a ureteric bud sidebranch from an epithelial tube known as the nephric duct. Subsequent branching of the ureteric bud to form the collecting duct tree is regulated by subcellular interactions between the epithelium and a population of mesenchymal cells that surround the tips of outgrowing branches. The mesenchymal cells produce glial cell-line derived neurotrophic factor (GDNF), that binds with RET receptors on the surface of the epithelial cells to stimulate several subcellular pathways in the epithelium. Such interactions are known to be a prerequisite for normal branching development, although competing theories exist for their role in morphogenesis. Here we introduce the first agent-based model of *ex vivo* kidney uretic branching. Through comparison with experimental data, we show that growth factor-regulated growth mechanisms can explain early epithelial cell branching, but only if epithelial cell division depends in a switch-like way on the local growth factor concentration; cell division occurring only if the driving growth factor level exceeds a threshold. We also show how a recently-developed method, "Approximate Approximate Bayesian Computation", can be used to infer key model parameters, and reveal the dependency between the parameters controlling a growth factor-dependent growth switch. These results are consistent with a requirement for signals controlling proliferation and chemotaxis, both of which are previously identified roles for GDNF.

Keywords

Organogenesis, cellular automaton, developmental processes, mathematical modelling, morphology.

Author Summary

A number of important congenital disorders arise due to incomplete development of the mammalian kidney. Elucidating the cause of these conditions requires an understanding of the mechanisms that contribute to kidney morphogenesis. Whilst experimental work has suggested several candidate mechanisms, their importance is still not well understood. Here we develop a computational model of kidney morphogenesis at the individual cell level to compare these different hypotheses. We hypothesise that a generic growth factor, that we term "GDNF", produced from the mesenchyme surrounding the epithelium, can drive a number of cellular responses, based on prior evidence. Simulations of our agent-based model reveal that diffusion of GDNF, coupled with GDNF-stimulated epithelial cell division, can generate the branching patterns seen in *ex vivo* kidney explant experiments. We also find that branching depends on the sensitivity of cell proliferation to changes in GDNF levels. In particular our model only generates realistic branching when there is significant variation in GDNF levels along the boundary of the epithelium, and most cells divide only if the local concentration of GDNF exceeds a threshold value. We conclude that feedback between mesenchymal cells that produce GDNF, and epithelial cells that consume it, is vital for normal kidney organogenesis.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 Introduction

The kidney is a complex organ with a highly branched structure. Its primary function is to filter urea and other waste products from the blood and metabolic system. Kidney and urinary tract congenital disorders are amongst the most common birth defects, with many of these conditions being caused by incomplete branching during development (Airik & Kispert, 2007). To understand how such disorders arise, we must first address fundamental questions about the developing kidney, including how branching is initiated, how it is regulated, and how cessation of branching is controlled. These processes are complex and, as such, biological experiments that aim to answer these questions are technically difficult to conduct. Furthermore, each experiment can typically investigate only a single facet of kidney morphogenesis. Mathematical and computational models can assist in this endeavour by integrating multiple biological hypotheses via a well-defined set of assumptions. Simulations of the resulting models can be performed to test alternative hypotheses about the mechanisms that regulate kidney morphogenesis, and generate new predictions that can be tested experimentally (Cebrian *et al.*, 2014; Combes, 2015; Murray *et al.*, 1983; Packard *et al.*, 2013; Short *et al.*, 2014; Zubkov *et al.*, 2015).

Nephrons are the primary functional units of a kidney. A fully developed human kidney has between 0.2 and 1.8 million nephrons (Hughson *et al.*, 2003), that are connected by a system of collecting ducts to the ureter and bladder (Costantini, 2006). Most of the organ is formed during embryonic development, beginning five weeks post-gestation in humans and embryonic day 10.5 (E10.5) in mice (Carlson, 2013; Cebrian *et al.*, 2004). An outgrowth of epithelial cells from the nephric duct leads to the formation of the ureteric bud epithelium (Little, 2015). As the ureteric bud grows and branches, a cloud of mesenchymal progenitor cells form caps at the tips, and eventually differentiate into nephrons.

In addition to their role as nephron precursors, mesenchymal cells release cytokines that influence the growth and branching of epithelial cells as they form the collecting tree structure. Glial cell line-derived neurotrophic factor (GDNF) is one such growth factor. GDNF is expressed in the metanephric mesenchyme by E11.5, and is thought to diffuse across to the epithelial cell layer, where its binding to cell surface RET receptors transduces signals that are essential for morphogenesis (Durbec *et al.*, 1996). GDNF is known to be a chemoattractant, and also to stimulate the outgrowth of epithelial cells (Little & McMahon, 2012). However, there is no consensus about whether these mechanisms are sufficient to generate normal branching, or whether additional chemical or mechanical mechanisms are needed. For example, GDNF-independent signals involving members of the fibroblast growth factor family and Activin are known to affect branching (Maeshima *et al.*, 2007; Michos *et al.*, 2010; Miyazaki *et al.*, 2000; Qiao *et al.*, 2001; Tee *et al.*, 2010, 2013). As the roles of GDNF and other growth factors remain to be fully elucidated, in this paper we focus on a single generic growth factor which we term "GDNF" and attribute to it actions and effects that we acknowledge are likely caused by a combination of different factors. Although the mechanical forces between epithelium and mesenchyme may influence branching patterns, they are not a prerequisite for branching. Indeed experiments performed by Qiao *et al.* (1999) have revealed that cell–cell contact between mesenchyme and epithelium is not required for branching; exposure of epithelial cells to soluble factors derived from metanephric mesenchyme is sufficient. While it is possible that mechanical forces between the epithelial cells may be an important mechanism for branching, to our knowledge no

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

experiments have yet directly investigated this hypothesis.

Branching morphogenesis is a characteristic feature of many mammalian organs including the kidney, lung, vasculature, the saliva and mammary glands, and in limb development (Affolter *et al.*, 2009). In each of these tissues, branched structures arise due to repetition of three cellular motifs: *bud formation* followed by *bud extension*, then *bud splitting*, with these processes being facilitated by regulation at the tips and stalks of individual buds. Similar pathways and network topologies perform this regulation in vastly different systems; for example, fibroblast growth factor signalling in combination with Delta/Notch signals regulate branching of both trachea and vertebrate vasculature (Affolter *et al.*, 2009; Ochoa-Espinosa & Affolter, 2012). Several alternative theories have been proposed to explain how such signalling generates branched structures. For example, Turing's reaction-diffusion mechanism can give rise to branched structures (Menshykau & Iber, 2013), although mechanical theories have also been proposed (Varner & Nelson, 2014). To distinguish between competing theories, simple systems are needed, whose variability can be controlled. Explant models of branching have a valuable role to play here as they can recapitulate branching morphogenesis, and are amenable to image analyses that are difficult or impossible to perform *in vivo* (Basson *et al.*, 2006; Serls *et al.*, 2005; Watanabe & Costantini, 2004). Additionally, it is possible not only to monitor the structure of an explant as it evolves over time but also to determine for example how these dynamics change when the composition of the culture medium that bathes the explant is altered.

A variety of mathematical models have been developed to study different aspects of branching morphogenesis, including physical cellular processes, such as proliferation and migration (Hirashima *et al.*, 2009), the underlying molecular processes (Zubkov *et al.*, 2015), or a combination of the two (Adivarahan *et al.*, 2013; Clément & Mauroy, 2014; Menshykau & Iber, 2013). A recent study used ordinary differential equations to describe how the dynamics of the growing epithelial cell populations at the tips of branches are regulated by mesenchymal cells (Zubkov *et al.*, 2015). Comparison of simulation results and experimental data revealed that the mathemematical model could recapitulate the observed dynamics for the ratio of epithelial to mesenchymal cells at the branch tips. In other studies a Turing system (Turing, 1952) involving the diffusion of the secreted ligands GDNF (produced by mesenchyme) and Wnt11 (produced by the epithelium in response to the presence of GDNF), as well as RET receptors, is proposed to be the primary driver of branching (Clément & Mauroy, 2014; Menshykau & Iber, 2013). Predictions generated by these models include the need for cooperative receptor-ligand binding (Menshykau & Iber, 2013), and a mechanism for self avoidance (so that growing tips do not come into contact) (Clément & Mauroy, 2014), which recent experiments suggest may be orchestrated by signalling of BMP7 that is produced by neighbouring branches (Davies *et al.*, 2014).

Whilst a few models of developmental processes consider individual cell behaviour (Fletcher *et al.*, 2014; McLennan *et al.*, 2012, 2015), the majority treats cell populations as a continuum (Clément & Mauroy, 2014; Menshykau & Iber, 2013; Scialdone *et al.*, 2013; Zubkov *et al.*, 2015). While these models have generated valuable insight into different aspects of organ development, they are limited in their ability to investigate the influence of cellular and subcellular mechanisms. Agent-based frameworks that model individual cell behaviour can offer significant advantages here.

In a biological context, agent-based models treat cells as individual agents, whose behaviour is specified by a predefined set of rules that can be deterministic and/or stochastic. First

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

developed to study the dynamics of replication (von Neumann, 1966), simple agent-based models such as cellular automata (CA), and generalisations of these such as cellular Potts and hybrid models are now used widely to study a range of biological systems that include biochemical reaction networks, stem cell proliferation and differentiation, tumour angiogenesis and metastasis (Alarcón *et al.*, 2003; Gerlee & Anderson, 2015; Macklin *et al.*, 2012; Roeder *et al.*, 2006; Scott *et al.*, 2014). Such "lattice-based" models restrict agents to sites on a fixed lattice. As such they are typically simpler in design and faster to simulate than alternative off-lattice models, which place fewer restrictions on the movement of cells (or agents) (Bentley *et al.*, 2009; Pathmanathan *et al.*, 2009; Perfahl *et al.*, 2016).

In this paper we investigate a series of experiments performed by Watanabe & Costantini (2004) in which kidneys from mouse embryos were grown in culture. We develop an agent-based framework to model the growth of these explants, and focus on interactions between growth factors (referred to as "GDNF ") derived from the mesenchyme surrounding the epithelium and epithelial cell surface RET receptors. Since the experimental data we have available is limited in detail we formulate an idealised model that contains the minimal number of assumptions necessary to recapitulate key branching features. Whilst the model does not explicitly account for mesenchyme cells, local levels of the generic growth factor, which we term GDNF, serve as a proxy for their influence. We use a CA approach which does not require specific assumptions about the nature of cell-cell forces that cannot (as yet) be experimentally verified. In more detail, our agent-based epithelial cells reside on a regular, two-dimensional grid, their rates of migration and proliferation being regulated by (and, in turn, regulating) the local distribution of GDNF.

Typically the parameters of computational models cannot be directly measured; they must be inferred from experimental data. Inference can be comparatively expensive in agent-based models, with the stochastic nature of the models requiring multiple simulations for each choice of parameter values. This cost can preclude parameter estimation even with state-of-the-art techniques, such as approximate Bayesian computation (ABC) (Johnson *et al.*, 2014; Liepe *et al.*, 2014; Toni *et al.*, 2009). ABC compares simulations from a model with experimental data, and – using statistics that summarise the system behaviour – accepts simulations (and the parameters that generated them), if the statistics for model and data lie within an acceptable distance threshold (see for example, Johnston *et al.* (2014); Vo *et al.* (2015)). ABC relies on the ability to simulate a model relatively quickly (typically faster than one second per one stochastic realisation of the model). In situations where this condition is not met, new methods are required for parameter inference. One such method is "Approximate Approximate Bayesian Computation" (AABC) (Buzbas & Rosenberg, 2015). With AABC actual model simulations are used to generate pseudo-replicates of the model that can be compared with the data. Since these pseudo-data are generated in a fraction of the time needed to perform real simulations, this method can exhibit significant speed up compared to ABC. Our agent-based CA model takes approximately a minute to generate a single simulation on a desktop computer, and hence we chose to use AABC for parameter inference.

We analyse the model first by direct simulation across a range of parameter regimes, in each case comparing model outputs and experimental data. We next use AABC to quantitatively fit the model to data, and reveal its dependence on key parameters. In doing so, we show AABC can be used to integrate the agent-based model (ABM) and the experimental data to answer

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the following questions:

1. Does the mechanism of GDNF-dependent cellular growth (that is, where cell proliferation rates increase with increases in the local concentration of GDNF) lead kidney explants to develop branches as seen *ex vivo*?

2. Which characteristics of GDNF-mediated mechanisms are necessary to generate these branches?

3. How sensitive is the model to changes in those parameters that influence branching?

We now briefly outline the structure of the paper. In Section 2.1 we introduce the experimental data that we use in this study, as well as the details of how it is processed to yield summary statistics that represent key features of explant branching. In Section 2.2 we describe our CA model, and in doing so, explain how individual cell behaviour depends on a field of GDNF that diffuses through the simulation domain. In Section 2.3 we explain how we implemented the AABC method, and it is used to investigate the factors that affect explant branching. In Section 3.1 we demonstrate how our model is able to recapitulate the branching of kidney explants, and in Section 3.2 we use direct simulation to determine how branching of our model explants depends on specific biological processes. Finally in Section 3.3 we use AABC to quantitatively investigate how explant branching depends on three key model parameters.

2 Materials and Methods

2.1 Kidney explant data experimental data and image processing

Watanabe & Costantini (2004) have developed an *ex vivo* assay to study epithelial branching in murine embryonic development. Kidneys were dissected from E11.5 embryos from a Hoxb7/EGFP transgenic line expressing green fluorescent protein (GFP) throughout the nephric duct and the ureteric bud (Srinivas *et al.*, 1999). Kidneys were cultured in fetal bovine serum and imaged every 30 minutes over a period of 96 hours, generating movies of the developing explants.

We processed videos of the experimental data (three explants in total) by extracting the area of the epithelial mass and the number of branch points in each frame. All image processing and analyses were performed using Matlab (MathWorks, Natick, MA). We started by extracting the bulk epithelial mass and then used this to calculate the medial axis skeleton (see Figs. 1, 3): this can be derived as the loci of centres of bi-tangent circles that fit entirely within the epithelial region (Lee, 1982). We then counted the number of branch points on each skeleton and used this as a measure of the amount of branching that had occurred. Calculation of the medial axis skeleton and counting of branching points were carried out using Matlab's "bwmorph" function, together with a third-party Matlab package (available from "https://uk.mathworks.com/matlabcentral/fileexchange/11123-better-skeletonization"). While this summary statistic captures the aggregate number of branches, it does not explain other features such as whether the branches are primary or secondary, nor does it provide information

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

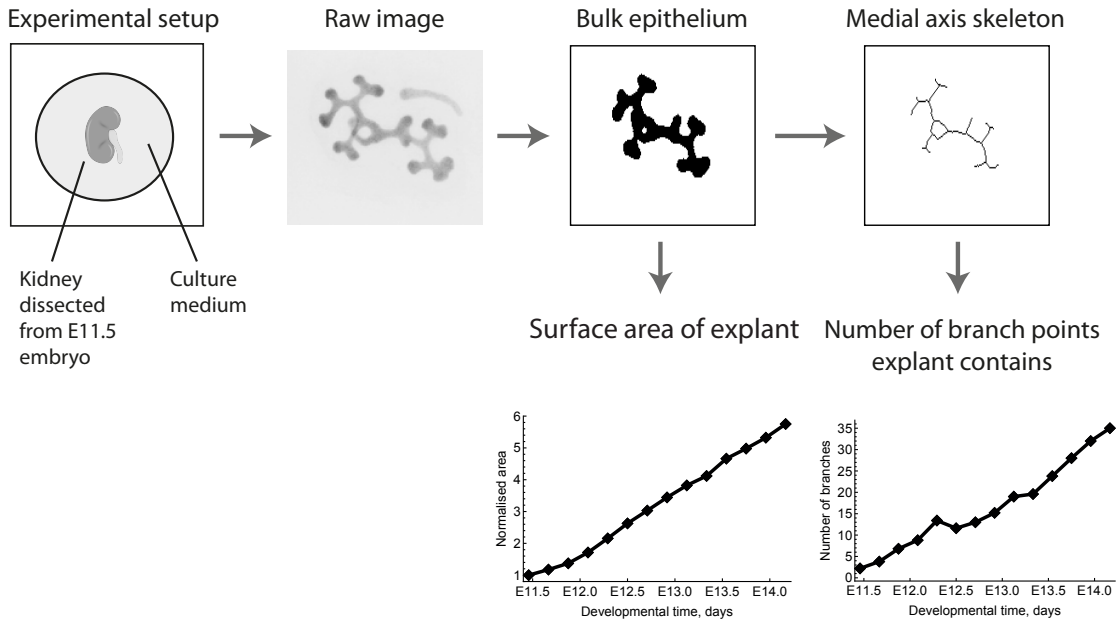

Figure 1. Workflow used to extract data from the explant images. The same process is applied to simulation results generated from the CA model. The raw image is reproduced with permission from Watanabe & Costantini (2004).

about the tip-to-stalk ratio (Bush *et al.*, 2014; Little, 2015); the collection and analysis of these more detailed statistics is postponed for future work.

For comparison with our model simulations we also monitored the area of the bulk epithelium as a measure of the change in the number of cells over time. To facilitate comparison of the data from each experiment we normalised the area of each epithelial mass at time t by its initial mass (at $t = 0$).

For each CA simulation (described below) we recorded the positions of all epithelial cells over time. We then applied the same image processing methods to the CA data as were applied to the experimental data, in order to generate comparable summary statistics for branching and cell numbers.

2.2 Cellular automaton model development

Our model comprises a two-dimensional cellular automaton (CA) model in which the behaviour of individual epithelial cells (i.e. their speed and direction, their rate of division and consumption of GDNF) depends on the locations of nearby cells and the local concentration of GDNF. Cell-cell interactions are governed by rules chosen to replicate observed behaviour. Our model is admittedly an idealisation of the biological processes that underpin kidney morphogenesis. In the absence of suitable experimental data, our model does not explicitly include mesenchyme cells even though they likely surround the explanted epithelium (they are invisible in the exper-

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

imental images). Similarly our model does not account for a GDNF-Wnt11 positive feedback loop: GDNF secreted by mesenchymal cells binds to RET receptors on epithelial cells and stimulates them to produce Wnt11 which binds to mesenchymal cells, stimulating further uptake of GDNF (Majumdar *et al.*, 2003). However since the available experimental data provides limited information we chose to formulate a simple model that could still reproduce key features of kidney morphogenesis.

The epithelial cells occupy a square domain which is discretised into $N \times N$ equally-spaced grid points. Each lattice site is occupied by either an epithelial cell or extracellular matrix (ECM). At $t = 0$, a mass of epithelial cells is introduced towards the centre of the domain. Its shape and size (i.e. the number of cells) are chosen to resemble those from the initial images of kidney explants from Watanabe & Costantini (2004). All other sites are occupied with ECM at $t = 0$.

GDNF field. We assume that GDNF binds to epithelial cells at rate K_G, and diffuses from the boundaries of the grid, where it is maintained at a constant GDNF concentration, G_∞. This implicitly represents a far-field approximation of the experimental conditions, where the ECM area is large, and the local concentration of GDNF is continuously replenished. These conditions are chosen to mimic the effects of GDNF produced by mesenchymal cells (that are likely present) as well as other growth factors that are present in the culture medium. Since the time scale for GDNF diffusion (seconds) is much shorter than the time scale for cell division (hours), we use a quasi-steady-state 2D reaction-diffusion-equation (with Dirichlet boundary conditions) to model the distribution of GDNF, $G(x, y, t)$,

$$\frac{\partial G}{\partial t} \approx 0 = D_G \nabla^2 G - \Phi_G \tag{1}$$

where D_G denotes the assumed-constant diffusion coefficient for GDNF, and Φ_G is the local rate of GDNF consumption,

$$\Phi_G = \begin{cases} K_G G(x, y, t), & \text{for an epithelial cell located at grid point } (x, y) \text{ at time } t, \\ 0, & \text{for ECM.} \end{cases} \tag{2}$$

where we assume for simplicity that the rate of GDNF consumption is linearly dependent on the concentration of substrate (with rate parameter K_G). In what follows, it is convenient to recast eqn. (1) in terms of a dimensionless GDNF concentration $g = \frac{G}{G_\infty}$, and non-dimensional spatial scale $\eta = \frac{x}{L}$, where $0 \leq g \leq 1$ is the GDNF concentration as a ratio of that at the boundaries, and $0 \leq \eta \leq 1$ is a spatial coordinate as a fraction of the simulation domain size. The reaction-diffusion equation can then be written as,

$$d_g \nabla_\eta^2 g = \phi_g, \tag{3}$$

where ∇_η^2 is the Laplacian with respect to the non-dimensional spatial coordinate, η; $d_g = \frac{D_G}{K_G L^2}$ is the non-dimensional diffusion coefficient; and $\phi_g = \frac{\Phi_G}{K_G G_\infty}$ is the non-dimensional GDNF uptake term. In these coordinates the boundary conditions are specified as,

$$\begin{aligned} g(0, y, t) &= 1 = g(L, y, t) \\ g(x, 0, t) &= 1 = g(x, L, t) \end{aligned} \tag{4}$$

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

where $0 \leq x, y \leq L$, (where $L = N - 1$) and the local rate of GDNF consumption is given by,

$$\phi_g = \begin{cases} g(x,y,t), & \text{for epithelium at grid point (x,y) at time t,} \\ 0, & \text{for ECM.} \end{cases} \tag{5}$$

In the absence of experimental estimates for the parameters D_G and K_G, preliminary numerical experiments revealed that fixing $d_g = 0.006$ lead to sufficient variation in GDNF across the spatial domain for branching to occur. (See Table S1 for a summary of the parameter values used in each simulation.)

At the end of each time step (see "Cell based rules" section) we solved eqns. (3), (4) and (5) using the method of explicit finite differences (implemented in Matlab). Both the CA model and the finite difference scheme are implemented on the same discrete grid, so the local GDNF level used in the update rules is its value at the grid point where the cell is located. The GDNF field affects the movement and behaviour of epithelial cells, while their location and rates of GDNF uptake affect the evolution of the GDNF field (Fig. 2).

Cell-based rules. Simple rules are used to determine whether individual cells move or divide (see Algorithm 1, and Table S1 for the parameter values used); cell death is assumed to be negligible in line with previous experimental results (Hartman *et al.*, 2007). The algorithm begins by solving for the steady state solution of the reaction-diffusion equation (as described above) for the current epithelial cell locations. Using a randomly-permuted list of cell indices, each cell is then visited in turn, and updated. Updates proceed as follows. If a cell has empty neighbouring sites then we determine whether to propose a movement or division event into one of the empty locations; the local GDNF concentration determines whether the proposed action occurs and, if so, the empty site at which the action is carried out.

During this process we establish whether any of the 4 adjacent sites in the von Neumann neighbourhood (up, down, left, right) are empty (no epithelium present). If this is the case then we select a move with probability p_{move}, or a cell division with probability $(1 - p_{move})$. The probability p_{move} is independent of GDNF, and is used simply to propose an action (that may not be undertaken). If an action is selected and feasible, our cell-based rules dictate whether it is executed. Moves are always carried out; whether cell division occurs depends on local levels of GDNF. We refer to this process as "GDNF-stimulated cell division". In particular, the probability p_{cd} that a cell at location (x, y) at time t divides is given by,

$$p_{cd} = \Phi(c_1 + c_2 g(x,y,t)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_1 + c_2 g(x,y,t)} e^{-\frac{\tau^2}{2}} \, \mathrm{d}\tau \tag{6}$$

where $\Phi(.)$ is the standard normal cumulative distribution function. This functional form allows a switch-like behaviour for cell division. In eqn. (6) the parameter c_1 controls the location of the switch, while the parameter c_2 determines its sensitivity to the local GDNF level, $g(x, y, t)$. In most simulations we fix $c_1 < 0$ and $c_2 > 0$, so that in the absence of GDNF ($g = 0$) by default epithelial cells do not divide.

If a move or cell division event is to occur, we must decide which empty neighbouring grid point will be occupied by the result of the action. This choice (if there is more than one empty

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

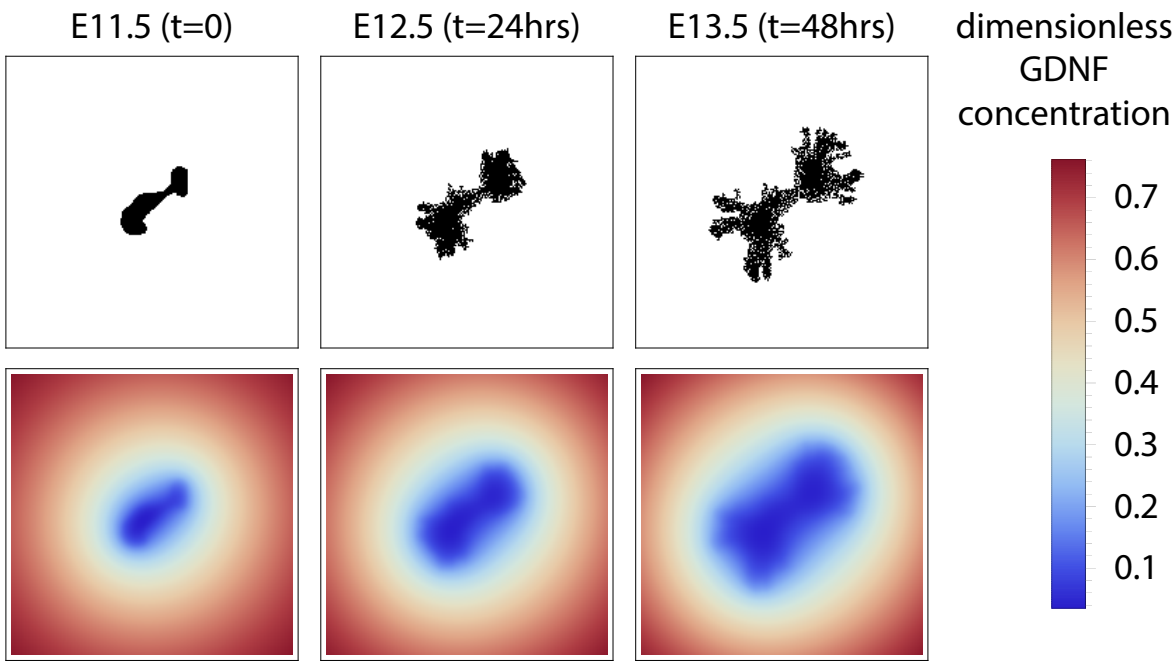

Figure 2. A series of plots from a typical CA simulation illustrating the interaction between epithelial cell positions (top) and GDNF level (bottom). Note: the GDNF level at the edges is not $g_\infty = 1$ because the panels only show the central 140×140 grid points of the domain (itself of size 400× 400). The parameter values used to generate these plots are shown in Table S1.

grid point) is biased by GDNF levels at the neighbouring sites. For moves this process represents "chemotaxis" and for cell divisions it represents "anisotropic cell division". In either case the probability p_i of selecting an empty neighbouring grid point $1 \leq i \leq K$ is given by,

$$p_i = \frac{exp(\beta_a g_i)}{\sum_{j=1}^{K} exp(\beta_a g_j)} \tag{7}$$

where $a \in \{\text{move}, \text{cell-division}\}$ specifies the action type, $1 \leq K \leq 4$ is the number of empty neighbours, and g_i is the GDNF concentration at site i. In eqn. (7) the non-negative parameter β_a controls the sensitivity of the selection to GDNF concentration, which can differ for chemotaxis and anisotropic cell division.

2.3 Parameter inference for agent-based models

Approximate Bayesian computation (ABC) allows parameter inference and model selection using distance-based criteria to compare simulations and data (Sunnåker *et al.*, 2013). A distance metric is introduced (e.g. Euclidean or Manhattan) and simulations are rejected if they yield values for a (set of) chosen statistic(s) that exceed a threshold value. ABC is appropriate when

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Algorithm 1 Pseudocode for updating the CA.

```
solve for steady state spatial distribution for g = "non-dimensional GDNF concentration"
create a randomly permuted list of cells, C
for each cell in C do
    if cell has empty von Neumann neighbours then
        //determine whether to propose a move or a cell division
        generate u_1 ~ U(0,1)
        if p_move > u_1 then
            proposed action = move
        else
            proposed action = cell division
        end if
        do GDNF-STIMULATED ACTION(proposed action)
    else
        proposed action = NULL
        consider next cell in C
    end if

end for

function GDNF-STIMULATED ACTION(proposed action)
    if proposed action = cell division then
        calculate p_cd = Φ(c_1 + c_2 g) = (1/√(2π)) ∫_{-∞}^{c_1+c_2 g} e^{-τ²/2} dτ
        generate u_2 ~ U(0,1)
        if p_cd > u_2 then
            do ANISOTROPIC CELL DIVISION
        end if
    else
        do CHEMOTAXIS
    end if
end function
function CHEMOTAXIS
    select grid point i out of K empty neighbouring sites with probability
```

$$p_i = \frac{exp(\beta_m g_i)}{\sum_{j=1}^{K} exp(\beta_m g_j)}$$

```
    move cell into site i
end function
function ANIOTROPHIC CELL DIVISION
    select grid point i out of K empty neighbouring sites with probability
```

$$p_i = \frac{exp(\beta_{cd} g_i)}{\sum_{j=1}^{K} exp(\beta_{cd} g_j)}$$

```
    place daughter cell in i
end function
```

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

a likelihood function is difficult or intractable to calculate. It has been applied to a wide range of problems in systems biology (Beaumont *et al.*, 2002, 2010; Csilléry *et al.*, 2010a). A prerequisite of ABC is that simulating a dataset from the model must be relatively inexpensive so that it is possible to adequately sample from the posterior parameter space. This is the case whether performing simple ABC rejection or more sophisticated procedures, such as sequential Monte Carlo ABC (Toni *et al.*, 2009). A simulation time for a single dataset of $\mathcal{O}(seconds)$ or greater is thus typically prohibitive for performing ABC.

A new algorithm named "approximate approximate Bayesian computation" (AABC) has been designed to resolve this problem (Buzbas & Rosenberg, 2015). By replacing true model simulations with realisations of a suitably-tuned statistical model, AABC enables parameter inference for a class of models that previously presented large (or insurmountable) computational challenges. Agent-based models, such as our CA model, represent such a class, that have with few notable exceptions (Johnston *et al.*, 2014; Jones *et al.*, 2015; Sottoriva & Tavare, 2010), evaded approximate Bayesian inference.

Here we use AABC to infer parameters of the CA model that yield behaviour consistent with the average summary statistics extracted from dynamic experimental data for kidney explants generated by Watanabe & Costantini (2004). These summary statistics (see section 2.1 above) measure two facets of the evolving explants: the number of branches, and the rate of cell proliferation. By comparing summary statistics from the experimental data with equivalent statistics from our simulations (as detailed below) we investigate the sensitivity of explant growth patterns to variation in three model parameters c_1, c_2, and p_{move}. We specify uniform priors for each of these parameters with the following bounds: $c_1 \in [-20, -40]$, $c_2 \in [40, 280]$ and $p_{move} \in [0, 1]$. All other parameters are held fixed (as per Table S1).

Specifically the algorithm implements the following steps:

1. Simulate the CA model, and accept a subset of particles with parameter sets $\theta_i = (c_{1i}, c_{2i}, p_{movei})$ and corresponding datasets $\mathbf{x}_i = (x_{1i}, x_{2i})$, $i \in (1, 2, ..., m)$, where m is the number of simulations (each with a unique set of parameters sampled from the priors) and $N = 2$ is the number of replicates per parameter set. For our application, each dataset x consists of data for the normalised area and the number of branch points at $t = 10, 20, 30$ hours after E11.5.

2. Sample a new set of parameter values, θ^*, from the prior.

3. Calculate the weights, ω_i, using an Epanechnikov kernel (Buzbas & Rosenberg, 2015):

$$\omega_i = \frac{3}{4}\frac{1}{(\theta^* - \theta_{(k+1)})}\left[1 - \left\|\frac{\theta^* - \theta_i}{\theta^* - \theta_{(k+1)}}\right\|^2\right]\mathbb{1}_{\{\|\theta^* - \theta_i\| < \|\theta^* - \theta_{(k+1)}\|\}},$$

   where $\mathbb{1}_{\{\|\theta^* - \theta_i\| < \|\theta^* - \theta_{(k+1)}\|\}}$ is an indicator function with value 1 for the k parameter values $\theta = \theta_1, \theta_2, ..., \theta_k$ with the shortest Euclidean distance from θ^*, and 0 otherwise (from θ_{k+1} onwards).

4. Select $(\theta_i, \mathbf{x}_i)$, $i \in (1, 2, ..., k)$, for which $\omega_i > 0$.

5. Draw a sample ϕ – used to specify data resampling probabilities – from a Dirichlet distribution parameterised by ω_i, $i \in (1, 2, ..., k)$.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

6. Simulate a new dataset $\mathbf{x}^*$ of area and branching time point data (of size $N = 2$ replicates) by: (i) resampling datapoints from $\mathbf{x}_i$ with probabilities set by ϕ, and (ii) assuming that each replicate is equally probable.

7. Calculate the Euclidean distance between the real and simulated datasets, and add θ^* to the posterior, iff $||\mathbf{x}_i - \mathbf{x}^*|| < \epsilon$. We choose ϵ such that 5% of simulations are accepted as posterior samples.

8. Repeat steps (2)-(7) until convergence in the approximate posterior distribution is reached.

We perform Step (1) of the algorithm in Matlab (MathWorks, Natick, MA), and steps (2)-(7) in Julia (v0.3.5, julialang.org).

3 Results

3.1 GDNF-directed cellular proliferation can explain the branching patterns observed when kidney explants are cultured *ex vivo*

Simulations of the CA model reveal that it can generate branching patterns similar to those identified from *ex vivo* kidney explant data of Watanabe & Costantini (2004). In Fig. 3 we compare results from a typical simulation with experimental data collected at four time points. The CA model recapitulates notable features of branching: branching at both ends of the buds, followed by secondary branching at the branch tips. Our simulated explants also often produce branching events in which three or more branches emerge from a single tip, events which are observed in the developing kidney (Menshykau & Iber, 2013).

figure	**GDNF-stimulated cell division**	**chemotaxis**	**anisotropic cell division (ACD)**	**branching**
4,5A	-	-	-	-
4,5B	+	-	-	+
4,5C	-	+	-	-
4,5D	+	+	-	+
4,5E	-	-	+	-
4,5F	+	-	+	++
4,5G	-	+	+	-
4,5H	+	+	+	++

Table 1. A summary of the effect of different GDNF signalling mechanisms on explant branching. Each row of the table corresponds to specific panels of Figs. 4 and 5, with corresponding parameter sets indicated in Table S1. In the middle three columns "-" indicates that the mechanism is inactive, "+" indicates that it is active. In the right-hand column "-" indicates no significant branching, "+" indicates modest branching and "++" indicates that branching that is consistent with experimental data.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

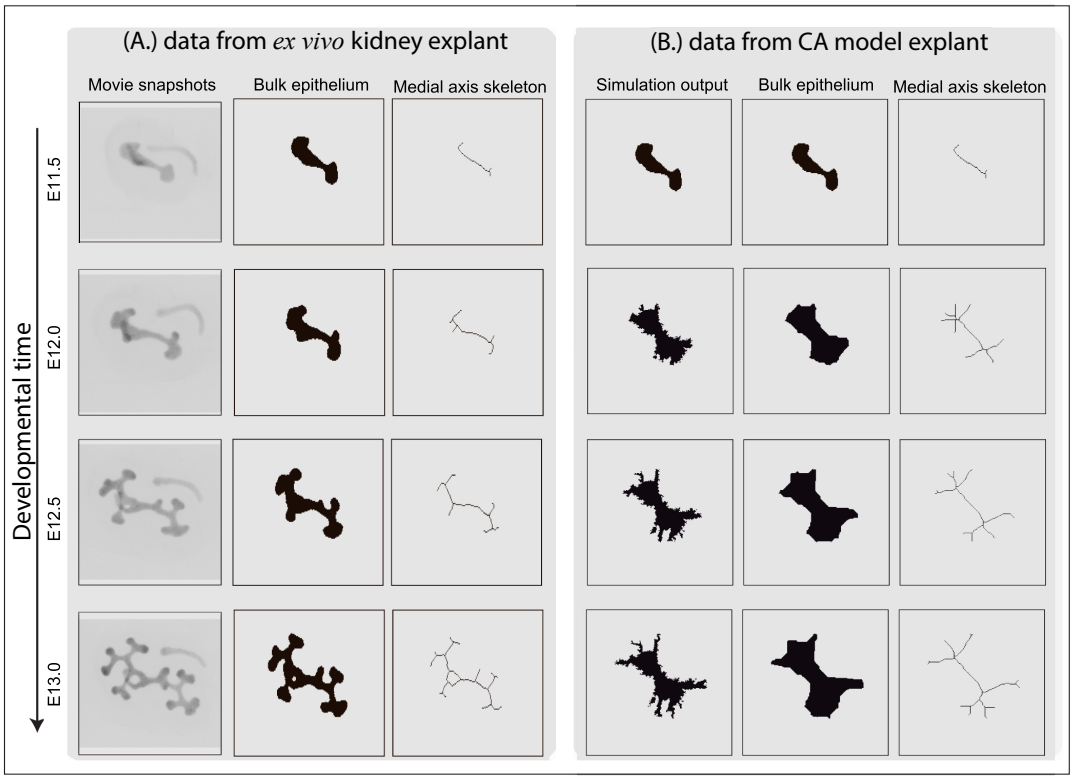

Figure 3. Comparison of branching patterns associated with (A.) *ex vivo* data and (B.) typical simulations of the CA model. Image processing techniques (see section 2.1) are used first to extract the bulk shape of the epithelial cells (middle panel) and then to generate the medial axis skeleton (right panel), which characterises the shape of the growing explant. The parameter values used to generate these simulation results are shown in Table S1.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

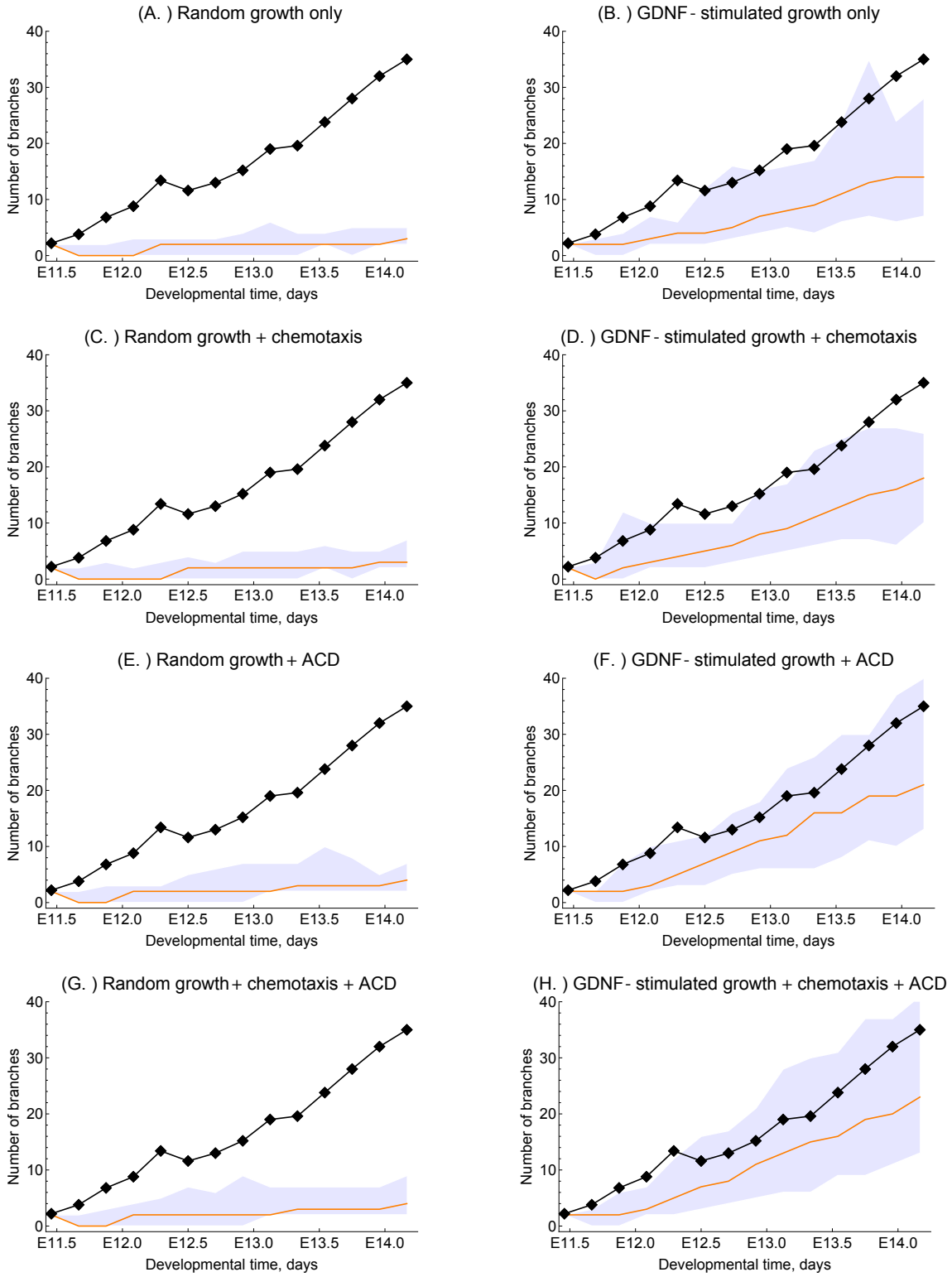

Figure 4. The effect of different GDNF signalling mechanisms on explant branching. (A.)-(H.) show simulation results with the indicated mechanisms implemented, and correspond to cell distributions shown in the panels of Fig. 5 with the same letters. In each panel the black line and points represent the evolution of branches from an explant experiment in Watanabe & Costantini (2004); the orange line represents the mean branching observed by model simulation (n= 200) and the shaded region indicates the 95% confidence interval. "ACD" indicates "anisotrophic cell division". The parameter values used in each case are in Table S1.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

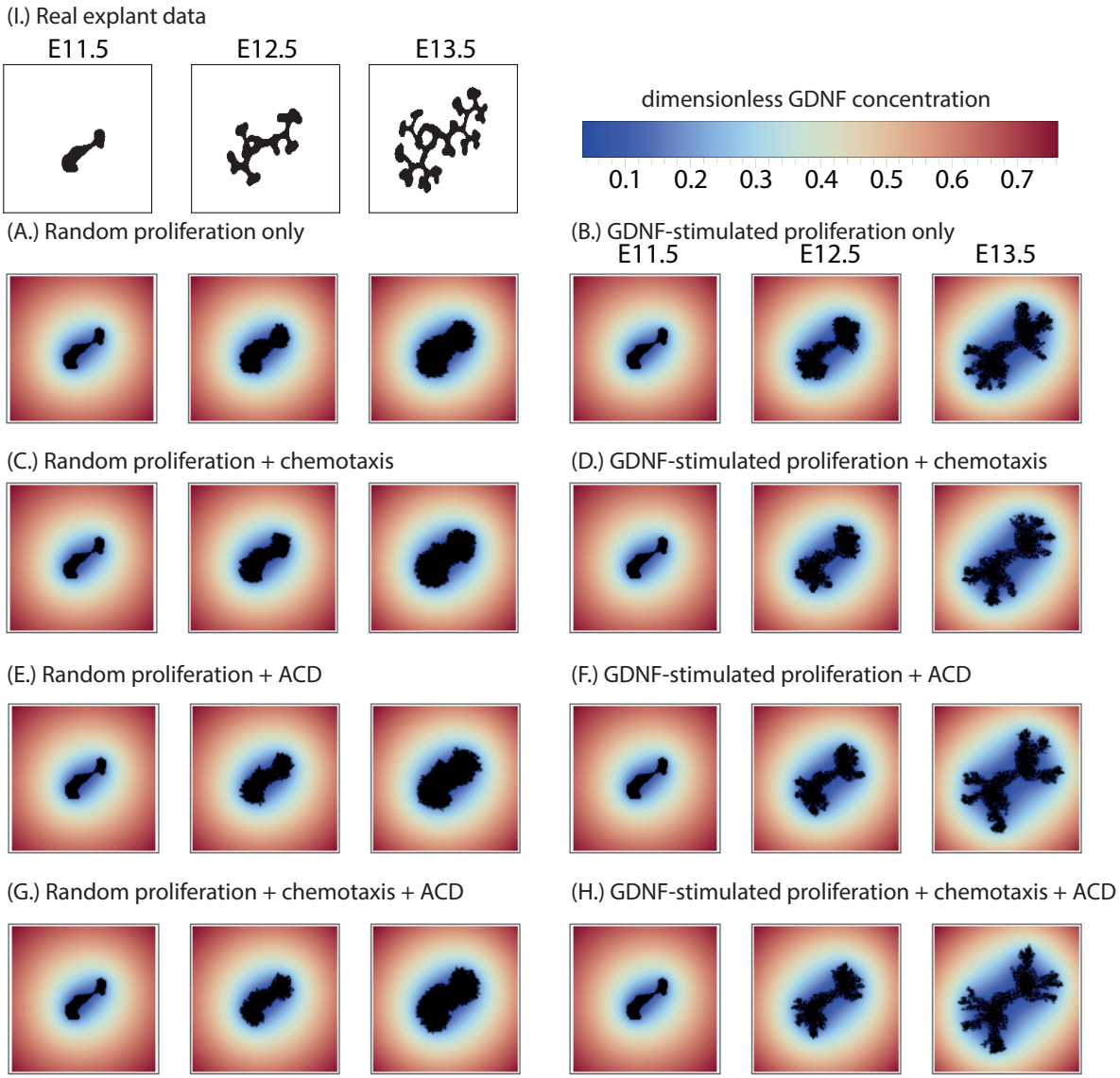

Figure 5. The effect of different GDNF signalling mechanisms on explant branching. (A.)-(H.) show simulation results with the indicated mechanisms implemented, and correspond to the panels in Fig. 4 with the same letters. The black mass is the epithelial cells, and the coloured shading shows the corresponding dimensionless distribution of GDNF. For comparison experimental from the first video in Watanabe & Costantini (2004) are presented in (I.). "ACD" indicates "anisotrophic cell division". Note: the GDNF level at the edges is not $g_\infty = 1$ because the panels only show the central 140×140 grid points of the domain (itself of size 400×400). The parameter values used in each case are in Table S1.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

3.2 Finding a minimal set of sub-cellular mechanisms necessary to generate branching

We also used our model to investigate the contribution of cell proliferation, anisotropic cell division and chemotaxis to branching (Table 1 and Figs. 4 and 5). Each mechanism is regulated by local levels of GDNF: the rate of cell division increases with local levels of GDNF in a sigmoidal manner (see Fig. 6); regarding chemotaxis, cells are more likely to migrate up local GDNF gradients; for anisotropic cell division (ACD), daughter cells are located preferentially in sites with higher levels of GDNF. For both chemotaxis and ACD we acknowledge that our model is an idealisation of the biological processes involved. For example RET-dependent movement (Riccio *et al.*, 2016) and luminal mitosis (Packard *et al.*, 2013) are not considered explicitly. Even so our model captures some of the features of these processes and, hence, can still be used to determine their relative contributions for branching.

Model simulations reveal that when cell division is independent of GDNF, no branching occurs; the explant grows as an approximately circular mass (see Figs. 4,5A). When only cell proliferation is regulated by GDNF, some branching occurs, although the number of branches is fewer than for explant growth (see Figs. 4,5B). When chemotaxis and/or anisotropic cell division depend on GDNF although the rate of proliferation is independent of GDNF, the model does not exhibit branching (Figs. 4,5 C,E,G). Additionally the cumulative effect of GDNF-stimulated proliferation and chemotaxis on branching is no greater than proliferation stimulated by GDNF alone (Figs. 4,5D). By contrast when proliferation and anisotropic cell division depend on GDNF, the number of branches observed along the branching trajectory increases such that simulation results are in good agreement with the experimentally observed branching patterns (Figs. 4,5F). The best agreement with the experimental data is obtained when all three processes (i.e. proliferation, chemotaxis and anisotropic cell division) depend on GDNF levels (Figs. 4,5H). These results do not appear to depend strongly on the initial shape of the epithelium used to in model simulations. In particular simulations initialised using different experimental explant data yielded similar results (see Figs. S1, S2).

These simulation studies reveal several key messages. First, GDNF-stimulated proliferation can generate biologically realistic branching patterns. Second, what we term anisotropic cell division ameliorates the rate of branching when it is coupled to GDNF-stimulated growth. Finally, within the regions of parameter space studied, GDNF-controlled chemotaxis does not strongly affect the branching of explants. In order to determine whether the weak dependence on chemotaxis was due to an overly simplistic treatment of cell-cell interactions, we revised our CA model ensure that epithelial cells remain attached to at least one neighbouring cell. However this change did not significantly alter the observed branching patterns (results not shown).

We also investigated how the shape of the GDNF-controlled switch that regulates cell proliferation affects branching (Fig. 6). When the slope of the switching function is gradual, the epithelium grows as a circular mass and no branching occurs (Fig. 6A). Similarly, when the threshold of the switch is high, the epithelial growth rate and branching rate are both too low (Fig. 6C). We therefore tune the threshold and slope of the GDNF-mediated switch (Fig. 6B) in order to maximise the amount of branching. Again we show in the Supplementary Materials how these results are robust to variation in the shape of the epithelial cell starting mass (see Figs. S3 and S4).

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

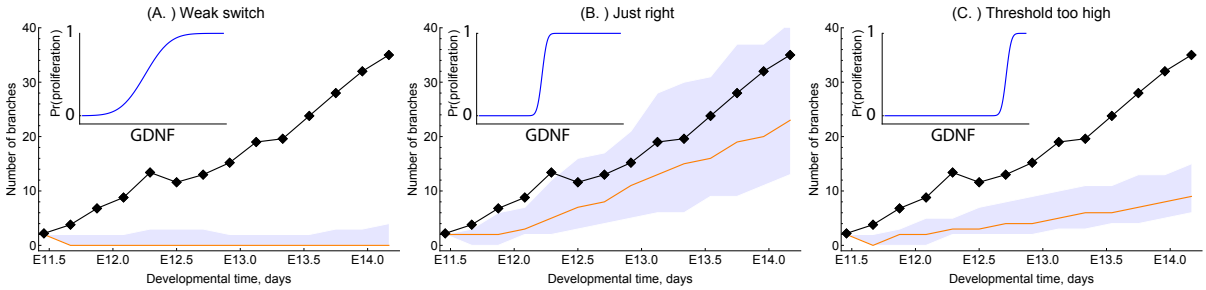

Figure 6. Series of simulation results showing how the shape of the GDNF-mediated proliferation switch influences the branching dynamics. For the results shown in each of the panels we ran simulations with $c_1 = -25$ so that in the absence of GDNF cells would not divide, and we vary c_2 across each of the panels: for (A.) $c_2 = 400$, (B.) $c_2 = 120$ and (C.) $c_2 = 20$. The inset panels show the location of the proliferation switch in GDNF-space (horizontal axis), against the probability of growth (vertical axis). In each panel the black line and points represent the evolution of branches from an explant experiment in Watanabe & Costantini (2004); the orange line represents the mean branching observed by model simulation (n= 200) and the shaded region indicates the 95% confidence interval. The parameter values used in each case are shown in Table S1.

3.3 Branching is sensitive to the form of the GDNF proliferation switch

Having shown that our CA model can reproduce qualitative features associated with early kidney morphogenesis, we now study the dependence of branching characteristics on the parameters c_1, c_2, and p_{move}; c_1 and c_2 jointly determine the sensitivity of cell proliferation to GDNF levels (see eqn. (6)), and p_{move} is the probability of cell migration.

We performed 50,000 AABC simulations and accepted the top 5% (the 5% with overall smallest Euclidean distances from the summary statistics associated with the experimental data) to compose approximate posterior samples, as outlined in section 2.3. We use two summary statistics to compare simulations with data: the normalised area and the total number of branch points as calculated from medial axis skeletons of the epithelial cell mass.

In Fig. 7A we plot those trajectories that have been accepted as posterior samples. The marginal posterior distributions associated with each parameter are shown on the diagonal of Fig. 7B alongside the two-dimensional posterior joint density distributions for each pair of parameters. This reveals that the marginal posterior distribution for c_2 deviates most from its prior, taking values between 40 and 280 with high probability; both c_1 and p_{move} deviate less from their prior, although p_{move} shows deviation at high values where little or no growth occurs.

The joint density plots in Fig. 7B reveal dependencies in parameter space. For the parameter pair (c_1, c_2), a negative correlation is observed; we note also that for low values of c_1 (i.e. $c_1 \in [-40, -25]$) c_2 is tightly constrained. This dependence is expected since the two parameters jointly determine the sensitivity of the switch to GDNF. In particular, under a local non-dimensional concentration of GDNF of $g = -\frac{c_1}{c_2}$ a cell with free neighbouring space divides with probability 0.5. This means that there are regions of (c_1, c_2) space (where $c_1 \sim -kc_2$)

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

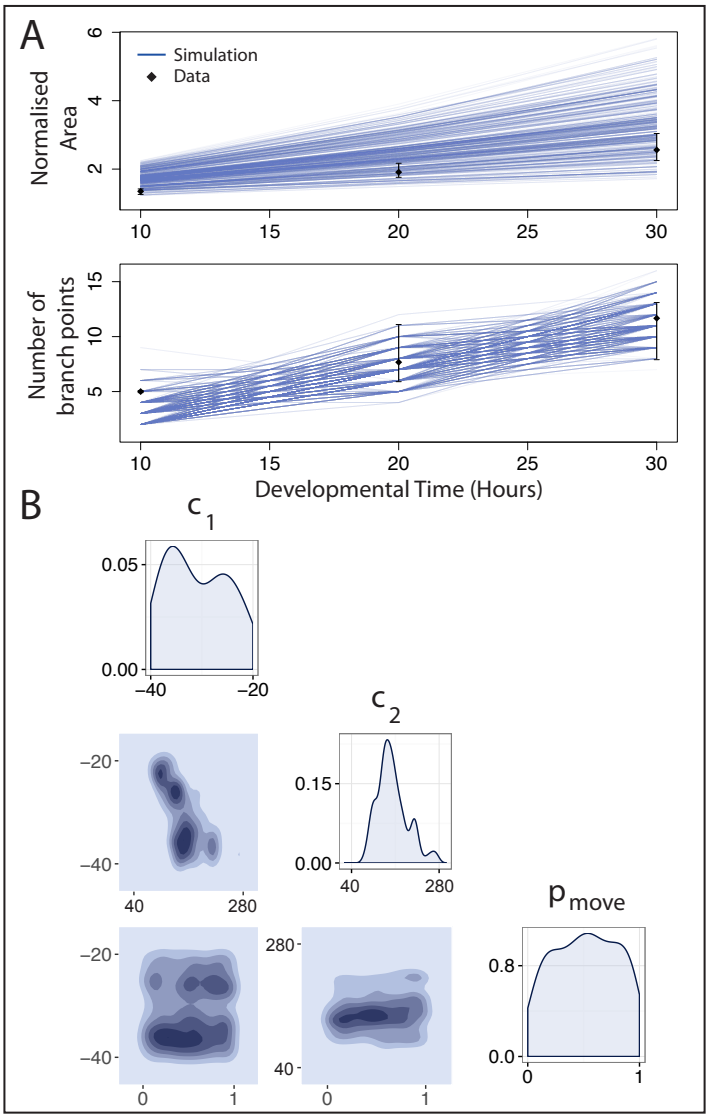

Figure 7. Approximate posteriors distributions for three model parameters as estimated by AABC. The parameters c_1 and c_2 affect the location and sensitivity of the GDNF-mediated growth switch, and p_{move} determines the relative likelihood of cell movement rather than cell division (see "Materials and Methods"). (A) Accepted trajectories in the posterior simulated from the model and their comparison with the *ex vivo* data (black dots indicate means, and error bars show the range of the data). The experimental data are the means of summary statistics extracted from three explant experiment videos in Watanabe & Costantini (2004). (B) Posterior parameter distributions for single (on the diagonal) and joint pairs of parameters. See Table S1 for information about the parameter values used in the simulations.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

which result in practically identical GDNF switches. For the other two parameter pairs, the joint density plots highlight what has been termed "sloppiness" (Gutenkunst *et al.*, 2007): the explant branching phenotype is robust to changes in the values of c_1 and p_{move}. Parameter sloppiness can be symptomatic of several problems, including insufficient data, or a model that is overly complex for the available data. Additionally we cannot discount the possibility that our summary statistics are suboptimal.

4 Discussion

We have developed a new model to describe branching morphogenesis in the developing kidney. While there are many agent-based models of other branching systems (e.g. Hartmann & Miura (2006); Iber & Menshykau (2013); Kondo & Miura (2010); Lindenmayer (1968); Mantzaris *et al.* (2004)) to our knowledge, this is the first agent-based model of kidney morphogenesis. A significant advantage of our approach is that unlike spatially-averaged compartment-based methods (e.g. Menshykau & Iber (2013); Zubkov *et al.* (2015)) it allows study of processes at the single-cell level. This spatial resolution may be particularly important in determining the thresholds at which branching occurs. The CA framework allows a cell-based description of tissue morphogenesis and facilitates the future addition of other biophysical mechanisms and the signalling pathways that regulate them, as well as extension to include other cell types (e.g. mesenchymal cap cells).

At present there is no consensus about whether diffusion-driven Turing patterns of GDNF coupled with GDNF-regulated proliferation mechanisms can explain branching, or whether other chemical and mechanical mechanisms are required. Our simulation results indicate that the GDNF-mediated proliferation may suffice to generate branching in this system. We find that to recapitulate the branching behaviour of the developing kidney, dynamic spatial patterning of epithelial cell proliferation had to be included in the model. This was the case across all the parameter sets that we considered, however there may be isolated regions in parameter space in which branching occurs via independent mechanisms not included in our model (e.g. due to mechanical cell-cell interactions). Simulation studies demonstrated that GDNF-stimulated tissue growth together with chemotaxis and anisotropic cell division provided the best fit to the explant branching data studied. In a recent study, Riccio *et al.* (2016) conclude, based on looking at the behaviour of tip cells across development, that GDNF signalling largely drives cell movement (rather than proliferation), and that this drives branching. While GDNF-dependent chemotaxis may play a role later in kidney development (we focus here on E11.5 to E14), our results suggest that it may not contribute significantly during early branching of the epithelium, and supports a greater role for proliferation than that study identified. However, it is also possible that there may be mesenchyme-derived factors other than GDNF that provide this signal. Inference of the CA model on a three-dimensional subset of parameters revealed dependency between the parameters controlling the GDNF-stimulated cell proliferation switch (c_1 and c_2). Specifically this result indicated how cell proliferation should depend strongly on GDNF levels. However it remains to identify a biological mechanism that fulfils this criteria.

Whilst our model can generate explant patterns that mimic some aspects of the experimental data, we recognise that our simulation results often differ in the finer details. At present it is

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

not clear whether these discrepancies are due to biological mechanisms that we have neglected, or due to the relatively simple nature of our model framework. For example these differences might be reduced if an off-lattice agent-based model, incorporating cell-cell contact forces, were used. We also note that the differences between our simulation results and the experimental data are highest for later developmental times, and we speculate that other mechanisms may be responsible for continued branching at these later time points. In particular our model does not explicitly include mesenchymal cells that produce growth factors which may influence branching at later developmental times. Another candidate mechanism is heterogeneity in RET expression amongst different epithelial cells (Shakya *et al.*, 2005). In particular, epithelial cells with differing levels of RET expression have been shown to compete for positions within branches (Riccio *et al.*, 2016), and contribute unequally to the tips of the developing kidney (Chi *et al.*, 2009). We also assumed that the time scale for GDNF diffusion is much slower than the time scale for its uptake/binding to RET receptors. However we cannot discount that future experiments may invalidate this assumption.

As important differences exist between the dynamics of branching as it occurs *ex vivo* and *in vivo* (Short *et al.*, 2014), caution should be exercised when extrapolating from the former to the latter. Here we focus on the *ex vivo* model, due to the availability of published image data. This system also lends itself more naturally to 2D modelling, which is less computationally expensive than 3D simulations. Scaling an ABM approach to 3D represents a significant computational hurdle, and we cannot discount the possibility that the different topology may require qualitatively different mechanisms for branching to occur. However our results are supported by the results of computational modelling by Clément & Mauroy (2014) who found that diffusion of growth factors was sufficient to generate realistic 3D branching patterns. Even so our 2D model permits investigation of the dominant mechanisms in a simpler geometry, and presents an opportunity for a first model validation step before investigating 3D dynamics. Similarly, we note that our model omits certain details of the biology, for example, the (likely) presence of a growth factor-producing mesenchyme. This choice was dictated partly by the sparsity of the available experimental data. However since our relatively simple model was able to reproduce key aspects of the explant branching we chose not to include further (uncertain) biological details.

The flexibility that CA models offer in their ability to describe spatiotemporal heterogeneities is not only an advantage but also a limitation, because the biological interpretation of each update rule is not always clear. Modelling cell proliferation and migration with biophysical and mechanical forces could improve the mechanistic understanding of the model, but comes at significant computational cost (see for example, Kim *et al.* (2007); Rejniak & Anderson (2011)).

To perform parameter estimation, we have implemented a version of AABC, an approximate Bayesian inference scheme that is ideally suited to models such as ours that are computationally expensive to simulate. With the ever-increasing resolution of spatiotemporal data and, concurrently, an increasing number of models developed to describe relevant biological phenomena, we propose that AABC may find useful applications to a range of models in systems biology, outside the more typical population genetics applications for which it was developed (Beaumont *et al.*, 2002; Buzbas & Rosenberg, 2015; Csilléry *et al.*, 2010b).

Spatially-averaged models of organ development have generated several mechanisms for branching kidney organogenesis. In Zubkov *et al.* (2015), branching occurs at a specific cell ratio of epithelial (tip) and mesenchymal (cap) cells. In both Clément & Mauroy (2014) and

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Menshykau & Iber (2013), a growth-promoting ligand mechanism is used; in Menshykau & Iber (2013), the authors show that this leads to a Turing-type mechanism through interaction of GDNF and the RET receptor. The model that we present is consistent with these results, but goes further by proposing cellular scale rules that, coupled to the influence of a ligand field, enable branching.

Branching morphogenesis is an old problem in mathematical biology (Murray *et al.*, 1983), yet many open questions remain to be addressed. Here, using an agent-based model that directly describes the cell–cell interactions that occur during organ development, we shed light on the processes involved in defining the structure of the kidney. In the future, we propose that more complex hybrid models that combine biophysical and experimentally-validated rules for the migration and proliferation of epithelial cells will lead to further advances in our understanding of kidney morphogenesis. Additionally we argue that future models of *in vivo* kidney development should include other cells types known to be involved in organogenesis, for example, mesenchymal cells.

Acknowledgments

We thank Joe Pitt-Francis for useful advice on the image-processing of the explant data. We also thank the EPSRC for partial funding.

References

Adivarahan, Srivathsan, Menshykau, Denis, Michos, Odyssé, & Iber, Dagmar. 2013. Dynamic Image-Based Modelling of Kidney Branching Morphogenesis. *Pages 106–119 of: Computational Methods in Systems Biology*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Affolter, Markus, Zeller, Rolf, & Caussinus, Emmanuel. 2009. Tissue remodelling through branching morphogenesis. *Nature Reviews Molecular Cell Biology*, **10**(12), 831–842.

Airik, R, & Kispert, A. 2007. Down the tube of obstructive nephropathies: the importance of tissue interactions during ureter development. *Kidney International*, **72**(12), 1459–1467.

Alarcón, Tomas, Byrne, Helen M, & Maini, Philip K. 2003. A cellular automaton model for tumour growth in inhomogeneous environment. *Journal of Theoretical Biology*, **225**(2), 257–274.

Basson, M Albert, Watson-Johnson, Judy, Shakya, Reena, Akbulut, Simge, Hyink, Deborah, Costantini, Frank D, Wilson, Patricia D, Mason, Ivor J, & Licht, Jonathan D. 2006. Branching morphogenesis of the ureteric epithelium during kidney development is coordinated by the opposing functions of GDNF and Sprouty1. *Developmental Biology*, **299**(2), 466–477.

Beaumont, Mark A, Zhang, Wenyang, & Balding, David J. 2002. Approximate Bayesian computation in population genetics. *Genetics*, **162**(4), 2025–2035.

Beaumont, Mark A, *et al.* 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution and Systematics*, **41**(379-406), 1.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Bentley, Katie, Mariggi, Giovanni, Gerhardt, Holger, & Bates, Paul A. 2009. Tipping the balance: robustness of tip cell selection, migration and fusion in angiogenesis. *PLoS Computational Biology*, **5**(10), e1000549.

Bush, Kevin T, Martovetsky, Gleb, & Nigam, Sanjay K. 2014. Relevance of ureteric bud development and branching to tissue engineering, regeneration and repair in acute and chronic kidney disease. *Current Opinion in Organ Transplantation*, **19**(2), 153–161.

Buzbas, Erkan O, & Rosenberg, Noah A. 2015. AABC: approximate approximate Bayesian computation for inference in population-genetic models. *Theoretical Population Biology*, **99**(Feb.), 31–42.

Carlson, Bruce M. 2013. *Human Embryology and Developmental Biology*. Fifth edn. Saunders.

Cebrian, Cristina, Borodo, Karolina, Charles, Nikki, & Herzlinger, Doris A. 2004. Morphometric index of the developing murine kidney. *Developmental Dynamics*, **231**(3), 601–608.

Cebrian, Cristina, Asai, Naoya, D'Agati, Vivette, & Costantini, Frank D. 2014. The number of fetal nephron progenitor cells limits ureteric branching and adult nephron endowment. *Cell Reports*, **7**(1), 127–137.

Chi, Xuan, Michos, Odyssé, Shakya, Reena, Riccio, Paul, Enomoto, Hideki, Licht, Jonathan D, Asai, Naoya, Takahashi, Masahide, Ohgami, Nobutaka, Kato, Masashi, Mendelsohn, Cathy, & Costantini, Frank D. 2009. Ret-dependent cell rearrangements in the Wolffian duct epithelium initiate ureteric bud morphogenesis. *Developmental Cell*, **17**(2), 199–209.

Clément, Raphaël, & Mauroy, Benjamin. 2014. An archetypal mechanism for branching organogenesis. *Physical Biology*, **11**(1), 016003.

Combes, Alexander N. 2015. Towards a quantitative model of kidney morphogenesis. *Nephrology*, **20**(5), 312–314.

Costantini, Frank D. 2006. Renal branching morphogenesis: concepts, questions, and recent advances. *Differentiation*, **74**(7), 402–421.

Csilléry, Katalin, Blum, Michael GB, Gaggiotti, Oscar E, & François, Olivier. 2010a. Approximate Bayesian computation (ABC) in practice. *Trends in ecology & evolution*, **25**(7), 410–418.

Csilléry, Katalin, Blum, Michael G B, Gaggiotti, Oscar E, & François, Olivier. 2010b. Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, **25**(7), 410–418.

Davies, Jamie A, Hohenstein, Peter, Chang, C-Hong, & Berry, Rachel. 2014. A self-avoidance mechanism in patterning of the urinary collecting duct tree. *BMC developmental biology*, **14**(1), 1.

Durbec, Pascale, Marcos-Gutierrez, Camelia V, Kilkenny, Carol, Grigoriou, Maria, Wartiowaara, Kirmo, Suvanto, Petro, Smith, Darrin, Ponder, Bruce, Costantini, Frank D, Saarma, Mart, Sariola, Hannu, & Pachnis, Vassilis. 1996. GDNF signalling through the Ret receptor tyrosine kinase. *Nature*, **381**(6585), 789–793.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fletcher, Alexander G, Osterfield, Miriam, Baker, Ruth E, & Shvartsman, Stanislav Y. 2014. Vertex models of epithelial morphogenesis. *Biophysical Journal*, **106**(11), 2291–2304.

Gerlee, Philip, & Anderson, Alexander R A. 2015. The evolution of carrying capacity in constrained and expanding tumour cell populations. *Physical Biology*, **12**(5), 056001.

Gutenkunst, Ryan N, Waterfall, Joshua J, Casey, Fergal P, Brown, Kevin S, Myers, Christopher R, & Sethna, James P. 2007. Universally sloppy parameter sensitivities in systems biology models. *PLoS Computational Biology*, **3**(10), 1871–1878.

Hartman, Heather A, Lai, Hsiao L, & Patterson, Larry T. 2007. Cessation of renal morphogenesis in mice. *Developmental biology*, **310**(2), 379–387.

Hartmann, Dirk, & Miura, Takashi. 2006. Modelling in vitro lung branching morphogenesis during development. *Journal of theoretical biology*, **242**(4), 862–872.

Hirashima, Tsuyoshi, Iwasa, Yoh, & Morishita, Yoshihiro. 2009. Dynamic modeling of branching morphogenesis of ureteric bud in early kidney development. *Journal of Theoretical Biology*, **259**(1), 58–66.

Hughson, Michael, Farris, Alton B, Douglas-Denton, Rebecca, Hoy, Wendy E, & Bertram, John F. 2003. Glomerular number and size in autopsy kidneys: the relationship to birth weight. *Kidney International*, **63**(6), 2113–2122.

Iber, Dagmar, & Menshykau, Denis. 2013. The control of branching morphogenesis. *Open Biology*, **3**(9), 130088.

Johnson, Robert, Kirk, Paul, & Stumpf, Michael P H. 2014. SYSBIONS: Nested sampling for systems biology. *Bioinformatics*.

Johnston, Stuart T, Simpson, Matthew J, McElwain, DL Sean, Binder, Benjamin J, & Ross, Joshua V. 2014. Interpreting scratch assays using pair density dynamics and approximate Bayesian computation. *Open biology*, **4**(9), 140097.

Jones, Phoebe J M, Sim, Aaron, Taylor, Harriet B, Bugeon, Laurence, Dallman, Magaret J, Pereira, Bernard, Stumpf, Michael P H, & Liepe, Juliane. 2015. Inference of random walk models to describe leukocyte migration. *Physical Biology*, **12**(6), 066001.

Kim, Yangjin, Stolarska, Magdalena A, & Othmer, Hans G. 2007. A hybrid model for tumor spheroid growth in vitro I: theoretical development and early results. *Mathematical Models and Methods in Applied Sciences*, **17**(supp01), 1773–1798.

Kondo, Shigeru, & Miura, Takashi. 2010. Reaction-diffusion model as a framework for understanding biological pattern formation. *science*, **329**(5999), 1616–1620.

Lee, D T. 1982. Medial axis transformation of a planar shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **4**(4), 363–369.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Liepe, Juliane, Kirk, Paul, Filippi, Sarah, Toni, Tina, Barnes, Chris P, & Stumpf, Michael P H. 2014. A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nature Protocols*, **9**(2), 439–456.

Lindenmayer, Aristid. 1968. Mathematical models for cellular interactions in development II. Simple and branching filaments with two-sided inputs. *Journal of theoretical biology*, **18**(3), 300–315.

Little, Melissa H. 2015. Improving our resolution of kidney morphogenesis across time and space. *Current Opinion in Genetics & Development*, **32**(Mar.), 135–143.

Little, Melissa H, & McMahon, Andrew P. 2012. Mammalian kidney development: principles, progress, and projections. *Cold Spring Harbor Perspectives in Biology*, **4**(5).

Macklin, Paul, Edgerton, Mary E, Thompson, Alastair M, & Cristini, Vittorio. 2012. Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression. *Journal of Theoretical Biology*, **301**(May), 122–140.

Maeshima, Akito, Sakurai, Hiroyuki, Choi, Yohan, Kitamura, Shinji, Vaughn, Duke A, Tee, James B, & Nigam, Sanjay K. 2007. Glial cell-derived neurotrophic factor independent ureteric bud outgrowth from the Wolffian duct. *Journal of the American Society of Nephrology*, **18**(12), 3147–3155.

Majumdar, Arindam, Vainio, Seppo, Kispert, Andreas, McMahon, Jill, & McMahon, Andrew P. 2003. Wnt11 and Ret/Gdnf pathways cooperate in regulating ureteric branching during metanephric kidney development. *Development*, **130**(14), 3175–3185.

Mantzaris, Nikos V, Webb, Steve, & Othmer, Hans G. 2004. Mathematical modeling of tumor-induced angiogenesis. *Journal of mathematical biology*, **49**(2), 111–187.

McLennan, Rebecca, Dyson, Louise, Prather, Katherine W, Morrison, Jason A, Baker, Ruth E, Maini, Philip K, & Kulesa, Paul M. 2012. Multiscale mechanisms of cell migration during development: theory and experiment. *Development*, **139**(16), 2935–2944.

McLennan, Rebecca, Schumacher, Linus J, Morrison, Jason A, Teddy, Jessica M, Ridenour, Dennis A, Box, Andrew C, Semerad, Craig L, Li, Hua, McDowell, William, Kay, David, Maini, Philip K, Baker, Ruth E, & Kulesa, Paul M. 2015. Neural crest migration is driven by a few trailblazer cells with a unique molecular signature narrowly confined to the invasive front. *Development*, **142**(11), 2014–2025.

Menshykau, Denis, & Iber, Dagmar. 2013. Kidney branching morphogenesis under the control of a ligand–receptor-based Turing mechanism. *Physical Biology*, **10**(4), 046003.

Michos, Odyssé, Cebrian, Cristina, Hyink, Deborah, Grieshammer, Uta, Williams, Linda, D'Agati, Vivette, Licht, Jonathan D, Martin, Gail R, & Costantini, Frank. 2010. Kidney development in the absence of Gdnf and Spry1 requires Fgf10. *PLoS Genet*, **6**(1), e1000809.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Miyazaki, Yoichi, Oshima, Keisuke, Fogo, Agnes, Hogan, Brigid L M, & Ichikawa, Iekuni. 2000. Bone morphogenetic protein 4 regulates the budding site and elongation of the mouse ureter. *Journal of Clinical Investigation*, **105**(7), 863–873.

Murray, J D, Oster, George F, & Harris, A K. 1983. A mechanical model for mesenchymal morphogenesis. *Journal of Mathematical Biology*, **17**(1), 125–129.

Ochoa-Espinosa, Amanda, & Affolter, Markus. 2012. Branching morphogenesis: from cells to organs and back. *Cold Spring Harbor Perspectives in Biology*, **4**(10).

Packard, Adam, Georgas, Kylie, Michos, Odyssé, Riccio, Paul, Cebrian, Cristina, Combes, Alexander N, Ju, Adler, Ferrer-Vaquer, Anna, Hadjantonakis, Anna-Katerina, Zong, Hui, Little, Melissa H, & Costantini, Frank D. 2013. Luminal mitosis drives epithelial Cell dispersal within the branching ureteric bud. *Developmental Cell*, **27**(3), 319–330.

Pathmanathan, P, Cooper, Jonathan, Fletcher, Alexander, Mirams, Gary, Murray, P, Osborne, J, Pitt-Francis, Joe, Walter, A, & Chapman, SJ. 2009. A computational study of discrete mechanical tissue models. *Physical biology*, **6**(3), 036001.

Perfahl, Holger, Hughes, Barry D, Alarcón, Tomás, Maini, Philip K, Lloyd, Mark C, Reuss, Matthias, & Byrne, Helen M. 2016. 3D Hybrid Modelling of Vascular Network Formation. *Journal of Theoretical Biology*.

Qiao, Jizeng, Nigam, Sanjay K, & Sakurai, H. 1999. Branching morphogenesis independent of mesenchymal–epithelial contact in the developing kidney. *Proceedings of the National Academy of Sciences of the United States of America*.

Qiao, Jizeng, Nigam, Sanjay K, Bush, K T, Steer, D L, Stuart, R O, Sakurai, H, & Wachsman, W. 2001. Multiple fibroblast growth factors support growth of the ureteric bud but have different effects on branching morphogenesis. *Mechanisms of Development*, **109**(2), 123–135.

Rejniak, Katarzyna A, & Anderson, Alexander RA. 2011. Hybrid models of tumor growth. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, **3**(1), 115–125.

Riccio, Paul, Cebrian, Cristina, Zong, Hui, Hippenmeyer, Simon, & Costantini, Frank D. 2016. Ret and Etv4 Promote Directed Movements of Progenitor Cells during Renal Branching Morphogenesis. *PLoS Biology*, **14**(2), e1002382.

Roeder, Ingo, Horn, Matthias, Glauche, Ingmar, Hochhaus, Andreas, Mueller, Martin C, & Loeffler, Markus. 2006. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nature Medicine*, **12**(10), 1181–1184.

Scialdone, Antonio, Mugford, Sam T, Feike, Doreen, Skeffington, Alastair, Borrill, Philippa, Graf, Alexander, Smith, Alison M, & Howard, Martin. 2013. Arabidopsis plants perform arithmetic division to prevent starvation at night. *Elife*, **2**.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Scott, Jacob G, Hjelmeland, Anita B, Chinnaiyan, Prakash, Anderson, Alexander R A, & Basanta, David. 2014. Microenvironmental variables must influence intrinsic phenotypic parameters of cancer stem cells to affect tumourigenicity. *PLoS Computational Biology*, **10**(1), e1003433.

Serls, Amanda E, Doherty, Shawna, Parvatiyar, Pankhuri, Wells, James M, & Deutsch, Gail H. 2005. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. *Development*, **132**(1), 35–47.

Shakya, Reena, Watanabe, Tomoko, & Costantini, Frank D. 2005. The role of GDNF/Ret signaling in ureteric bud cell fate and branching morphogenesis. *Developmental Cell*, **8**(1), 65–74.

Short, Kieran M, Combes, Alexander N, Lefevre, James, Ju, Adler L, Georgas, Kylie M, Lamberton, Timothy, Cairncross, Oliver, Rumballe, Bree A, McMahon, Andrew P, Hamilton, Nicholas A, Smyth, Ian M, & Little, Melissa H. 2014. Global quantification of tissue dynamics in the developing mouse kidney. *Developmental Cell*, **29**(2), 188–202.

Sottoriva, Andrea, & Tavare, Simon. 2010. Integrating Approximate Bayesian Computation with complex agent-Based models for cancer research. *Pages 57–66 of: Proceedings of COMPSTAT'2010*. Heidelberg: Physica-Verlag HD.

Srinivas, S, Goldberg, M R, Watanabe, Tomoko, D'Agati, V, Al Awqati, Q, & Costantini, Frank D. 1999. Expression of green fluorescent protein in the ureteric bud of transgenic mice: A new tool for the analysis of ureteric bud morphogenesis. *Developmental Genetics*, **24**(3-4), 241–251.

Sunnåker, Mikael, Busetto, Alberto Giovanni, Numminen, Elina, Corander, Jukka, Foll, Matthieu, & Dessimoz, Christophe. 2013. Approximate bayesian computation. *PLoS Computational Biology*, **9**(1).

Tee, James B, Choi, Yohan, Shah, Mita M, Dnyanmote, Ankur, Sweeney, Derina E, Gallegos, Tom F, Johkura, Kohei, Ito, Chiharu, Bush, Kevin T, & Nigam, Sanjay K. 2010. Protein kinase A regulates GDNF/RET-dependent but not GDNF/Ret-independent ureteric bud outgrowth from the Wolffian duct. *Developmental Biology*, **347**(2), 337–347.

Tee, James B, Choi, Yohan, Dnyanmote, Ankur, Decambre, Marvalyn, Ito, Chiharu, Bush, Kevin T, & Nigam, Sanjay K. 2013. GDNF-independent ureteric budding: role of PI3K-independent activation of AKT and FOSB/JUN/AP-1 signaling. *Biology Open*, **2**(9), 952–959.

Toni, Tina, Welch, David, Strelkowa, Natalja, Ipsen, Andreas, & Stumpf, Michael P H. 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, **6**(31), 187–202.

Turing, Alan M. 1952. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **237**(Aug.), 37–72.

bioRxiv preprint doi: https://doi.org/10.1101/096032; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Varner, Victor D, & Nelson, Celeste M. 2014. Cellular and physical mechanisms of branching morphogenesis. *Development*, **141**(14), 2750–2759.

Vo, Brenda N, Drovandi, Christopher C, Pettitt, Anthony N, & Simpson, Matthew J. 2015. Quantifying uncertainty in parameter estimates for stochastic models of collective cell spreading using approximate Bayesian computation. *Mathematical biosciences*, **263**, 133–142.

von Neumann, John. 1966. *Theory of Self-Reproducing Automata*. University of Illinois Press.

Watanabe, Tomoko, & Costantini, Frank D. 2004. Real-time analysis of ureteric bud branching morphogenesis in vitro. *Developmental Biology*, **271**(1), 98–108.

Zubkov, Vladimir, Combes, Alexander N, Short, K M, Lefevre, J, Hamilton, Nicholas A, Smyth, I M, Little, Melissa H, & Byrne, Helen M. 2015. A spatially-averaged mathematical model of kidney branching morphogenesis. *Journal of Theoretical Biology*, **379**(Aug.), 24–37.