Journal of Theoretical Biology 443 (2018) 66–81

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

Approximate Bayesian computation reveals the importance of repeated measurements for parameterising cell-based models of growing tissues

Jochen Kursawe[a,1,*], Ruth E. Baker[a], Alexander G. Fletcher[b,c]

[a] Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK
[b] School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, UK
[c] Bateson Centre, University of Sheffield, Sheffield, S10 2TN, UK

ARTICLE INFO

Article history:
Received 28 September 2017
Revised 15 January 2018
Accepted 17 January 2018

Keywords:
Cell-based models
Approximate Bayesian computation
Parameter inference
Vertex models
Drosophila wing imaginal disc

ABSTRACT

The growth and dynamics of epithelial tissues govern many morphogenetic processes in embryonic development. A recent quantitative transition in data acquisition, facilitated by advances in genetic and live-imaging techniques, is paving the way for new insights to these processes. Computational models can help us understand and interpret observations, and then make predictions for future experiments that can distinguish between hypothesised mechanisms. Increasingly, cell-based modelling approaches such as vertex models are being used to help understand the mechanics underlying epithelial morphogenesis. These models typically seek to reproduce qualitative phenomena, such as cell sorting or tissue buckling. However, it remains unclear to what extent quantitative data can be used to constrain these models so that they can then be used to make quantitative, experimentally testable predictions. To address this issue, we perform an *in silico* study to investigate whether vertex model parameters can be inferred from imaging data, and explore methods to quantify the uncertainty of such estimates. Our approach requires the use of summary statistics to estimate parameters. Here, we focus on summary statistics of cellular packing and of laser ablation experiments, as are commonly reported from imaging studies. We find that including data from repeated experiments is necessary to generate reliable parameter estimates that can facilitate quantitative model predictions.

1. Introduction

The growth and dynamics of epithelial tissues are central to numerous processes in embryonic development. Experimental studies are generating an increasing volume of quantitative and semi-quantitative data on these processes Pargett and Umulis (2013), providing the potential for significant increases in our understanding of epithelial morphogenesis. Computational models can be used to interpret such data, develop theories for the underlying biophysical mechanisms, and make predictions for future experiments to distinguish between competing theories. Increasingly, 'cell-based' modelling approaches are used that exploit the availability of data across subcellular, cellular and tissue scales Fletcher et al. (2017). Usually, these models are calibrated and tested according to their ability to reproduce qualitative phenomena and, as such, their predictive capability is limited. It remains largely unclear to what extent cell-based models can be used in a quantitative manner.

One important challenge that arises when applying cell-based models quantitatively is parameter inference from experimental data. Since cell-based models predict high-dimensional data, for example the shape or position of each modelled cell, it is necessary to select low-dimensional summary statistics to infer model parameters. However, the optimal choice of summary statistic or experimental design to infer parameters is non-intuitive. A common approach to parameter inference for computational models is to identify parameters of best fit. However, this approach is problematic since it does not quantify the uncertainty associated with such estimates. Here, we conduct parameter inference, compare summary statistics, and estimate uncertainty for a vertex model of the *Drosophila* wing imaginal disc, a classical model system for the study of tissue growth (Farhadifar et al., 2007; Gibson and Perrimon, 2005; Hufnagel et al., 2007; Mao et al., 2011). Vertex models are a widely used class of cell-based model in which epithelial cell sheets are approximated by tessellations of

* Corresponding author.
E-mail address: jochen.kursawe@manchester.ac.uk (J. Kursawe).
[1] Present address: Faculty of Biology, Medicine and Health, University of Manchester, Michael Smith Building, Oxford Road, Manchester, M13 9PL, UK.

https://doi.org/10.1016/j.jtbi.2018.01.020
0022-5193/

polygons representing the apical surfaces of cells, and vertices (where three or more cells meet) move in response to forces due to growth, interfacial tension and the hydrostatic pressure within each cell (Farhadifar et al., 2007; Fletcher et al., 2013; Kursawe et al., 2015).

Many cell-based models include mechanical parameters that influence dynamic and steady-state behaviour. The optimal choice for such parameters for a given biological system is often not intuitive, since these parameters reflect mechanical properties of single cells or their pairwise interactions. To measure the mechanical properties of a single cell it may be necessary to remove it from the surrounding tissue, however this generally has a significant impact on its mechanical properties. Hence, parameterising cell-based models using tissue-scale data is an area of growing interest Jagiella et al. (2017). In the case of vertex models, efforts have been made to estimate mechanical parameters from tissue-scale measurements. Farhadifar et al. (2007) achieve this by manual fitting of their model using a combination of summary statistics: the relative occurrence of cells with different numbers of sides; the average area of cells with different numbers of sides; the maximum displacement of vertices following laser ablation of individual cell edges; and the area and perimeter changes for cells whose edges were ablated. Other authors have used similar statistics to parameterise vertex models of the *Drosophila* wing imaginal disc (Canela-Xandri et al., 2011; Mao et al., 2011; Schilling et al., 2011) and the *Xenopus laevis* animal cap Nestor-Bergmann et al. (2017). These authors arrive at different parameter estimates, which is not surprising given that the analysed tissues, summary statistics, and model implementations differ.

A variety of approaches have been taken to fit summary statistics obtained from vertex models to experimental data. These include ad hoc (Canela-Xandri et al., 2011; Schilling et al., 2011) and least-squares (Mao et al., 2011; Nestor-Bergmann et al., 2017) approaches. Merzouki et al. (2016) infer parameters by inducing tissue deformation; they compare stress-strain curves obtained from simulations with those obtained from experiments on monolayers of Madine–Darby Canine Kidney cells Harris et al. (2012). In these experiments, a free monolayer is suspended between rods and the stress curve is recorded as the distance between the rods was increased. A similar approach is taken by Xu et al. (2015a, 2015b), who highlight that the stress-strain curves obtained in this way are affected by the amount and orientation of cell divisions. Similarly, Wyatt et al. (2015) show that cell divisions in this tissue are oriented so as to dissipate stress. Studies of cell mechanical properties in epithelia also include force-inference methods (Chiou et al., 2012; Ishihara and Sugimura, 2012; Ishihara et al., 2013), where a heterogeneous generalisation of the vertex model is fitted to microscopy data to estimate the apical tension and pressure forces on each cell in the tissue. Such approaches are particularly suitable to estimate mechanical heterogeneity in a tissue without the need to physically manipulate the sample, though they typically assume the tissue to be in mechanical equilibrium, which is generally not the case during embryogenesis.

In summary, although some progress has been made in estimating the parameters of vertex models, none of the studies described above quantify the uncertainty of *in vivo* parameter estimates, nor do they address parameter identifiability. Hence, it remains unclear to what extent previously used summary statistics are sensible choices for parameter inference. Tissue manipulation approaches are not applicable *in vivo* since they require the removal of a tissue from its substrate (Harris et al., 2012; Merzouki et al., 2016; Xu et al., 2015a), while force-inference methods do not directly estimate model parameters. There is thus a need to establish to what extent each parameter can be estimated, and the associated uncertainty quantified, in such models. For the first time we apply Bayesian inference to estimate vertex model parameters in a simulation of tissue growth. Bayesian inference enables uncertainty quantification by calculating the probability of the model parameters given observed data. Specifically, we apply Approximate Bayesian computation (ABC) (Beaumont, 2010; Beaumont et al., 2002), a method that is commonly used when the likelihood of the model is not analytically tractable. To evaluate the inference method and compare summary statistics we apply the inference method to virtual data generated from the model, which allows us to compare inferred parameter values to the ground truth.

The remainder of this paper is structured as follows. In Section 2 we describe our *in silico* study and the inference method in detail. In Section 3 we infer model parameters using a range of summary statistics. We find that the uncertainty in parameter estimates generated using summary statistics of cell packing, or from tissue responses to laser ablations, depends heavily on the amount of data used to calculate summary statistics, and we investigate how much data is required for reliable parameter estimation. We identify the mean area of cells of each polygon class, in combination with the mean cell elongation, as a suitable summary statistic for vertex model parameter inference, and analyse how the quality of parameter estimates varies over parameter space. In Section 4 we discuss the implications of our findings.

2. Methods

2.1. Vertex model

For our *in silico* study, we consider a simplified vertex model of wing imaginal disc growth in the dipteran fly *Drosophila melanogaster*. The wing imaginal disc has been intensely studied as an experimental system for research on tissue growth and proliferation. During larval stages, this epithelial tissue undergoes a period of intense proliferation, increasing from around 40 to over 50,000 cells, before later giving rise to the adult wing (Aegerter-Wilmsen et al., 2010; Farhadifar et al., 2007). We now outline the technical details of our model; note that, motivated by the need to run a large number of simulations for parameter inference, we include several simplifying assumptions.

We represent the wing disc epithelium as a dynamic tessellation of polygons that approximate cell apical surfaces, with a vertex at each point where three cells meet. The position of each vertex, i, evolves in time according to the overdamped force equation

$$\tilde{\mu}\frac{\mathrm{d}\tilde{\mathbf{x}}_i}{\mathrm{d}\tilde{t}} = -\tilde{\nabla}_i\tilde{E}. \tag{1}$$

Here and throughout, we use tildes to denote dimensional quantities. In Eq. (1), $\tilde{\mu}$ denotes the friction strength, $\tilde{\mathbf{x}}_i(\tilde{t})$ is the position of vertex i at time $\tilde{t}$, $\tilde{E}$ denotes the total energy associated with the tissue, and $\tilde{\nabla}_i$ denotes the gradient operator with respect to $\tilde{\mathbf{x}}_i$. The number of vertices in the system may change over time due to cell division and removal (see below). For a homogeneous (unpatterned) epithelial tissue, the total energy $\tilde{E}$ is given by Farhadifar et al. (2007)

$$\tilde{E} = \sum_\alpha \frac{\tilde{K}}{2}(\tilde{A}_\alpha - \tilde{A}_{0,\alpha})^2 + \sum_\alpha \frac{\tilde{\Gamma}}{2}\tilde{P}_\alpha^2 + \sum_{\langle i,j\rangle,\mathrm{int}} \tilde{\Lambda}\tilde{l}_{i,j} + \sum_{\langle i,j\rangle,\mathrm{ext}} \tilde{\Lambda}_B\tilde{l}_{i,j}. \tag{2}$$

where the first two sums run over every cell α in the tissue, the third sum runs over every cell edge (pair of neighbouring vertices) $\langle i, j\rangle$ internal to the tissue and the third term runs over all cell edges at the boundary of the tissue, at time $\tilde{t}$. In Eq. (2), the variables $\tilde{A}_\alpha$ and $\tilde{P}_\alpha$ denote the area and perimeter of cell α, respectively, and the parameter $\tilde{A}_{0,\alpha}$ denotes a 'target' or preferred area for that cell. The four sums respectively represent the bulk elasticity of each cell, the presence of a contractile acto-myosin cable