

Research

Open Access

A statistical approach to estimating the strength of cell-cell interactions under the differential adhesion hypothesis

Mathieu Emily*^{1,2} and Olivier François¹

Address: ¹TIMC-TIMB, Université Joseph Fourier, INP Grenoble, Faculty of Medicine, 38706 La Tronche cedex, France and ²Bioinformatics Research Center (BiRC), University of Aarhus, Hoegh-Guldbergs Gade 10, 8000 Aarhus C, Denmark

Email: Mathieu Emily* - memily@daimi.au.dk; Olivier François - olivier.francois@imag.fr

* Corresponding author

Published: 18 September 2007

Received: 23 April 2007

Theoretical Biology and Medical Modelling 2007, **4**:37 doi:10.1186/1742-4682-4-37

Accepted: 18 September 2007

This article is available from: <http://www.tbiomed.com/content/4/1/37>

© 2007 Emily and François; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The Differential Adhesion Hypothesis (DAH) is a theory of the organization of cells within a tissue which has been validated by several biological experiments and tested against several alternative computational models.

Results: In this study, a statistical approach was developed for the estimation of the strength of adhesion, incorporating earlier discrete lattice models into a continuous marked point process framework. This framework allows to describe an ergodic Markov Chain Monte Carlo algorithm that can simulate the model and reproduce empirical biological patterns. The estimation procedure, based on a pseudo-likelihood approximation, is validated with simulations, and a brief application to medulloblastoma stained by beta-catenin markers is given.

Conclusion: Our model includes the strength of cell-cell adhesion as a statistical parameter. The estimation procedure for this parameter is consistent with experimental data and would be useful for high-throughput cancer studies.

Background

The development and the maintenance of multi-cellular organisms are driven by permanent rearrangements of cell shapes and positions. Such rearrangements are a key step for the reconstruction of functional organs [1]. *In vitro* experiments such as Holtfreter's experiments on the pronephros [2] and the famous example of an adult living organism Hydra [3] are illustrations of spectacular spontaneous cell sorting. Steinberg [4-7] used the ability of cells to self-organize in coherent structures to conduct a series of pioneering experimental studies that characterized cell adhesion as a major actor of cell sorting. Following his experiments, Steinberg suggested that the interaction between two cells involves an adhesion surface energy which varies according to the cell type. To

interpret cell sorting, Steinberg formulated the Differential Adhesion Hypothesis (DAH), which states that cells can explore various configurations and finally reach the lowest-energy configuration. During the past decades, the DAH has been experimentally tested in various situations such as gastrulation [8], cell shaping [9], control of pattern formation [10] and the engulfment of a tissue by another one. Some of these experiments have been recently improved to support the DAH with more evidence [11].

In the 80's and the 90's, the DAH inspired the development of many mathematical models. These models, recently reviewed in [12], rely on computer simulations of physical processes. In summary, these models act by min-

imizing an energy functional called the Hamiltonian, and they can be classified into four main groups according to the geometry chosen to describe the tissues.

First, *cell-lattice models* consider that each cell is geometrically described by a common shape, generally a regular polygon (square, hexagon, etc ...) (see [13] for example). Although these models may not be realistic due to the simple representation of each cell, their computation is straightforward and fast. The second class of models has been called *centric models*. In comparison with the cell-lattice models, centric models are based on more realistic cell geometries by using tessellations to define cell boundaries from a point pattern where points characterize cell centers [14]. While the main benefit of this class of models is the use of a continuous geometry, tessellation algorithms are known to be computationally slow [12]. The third class of models are the *vertex models*. These models are dual to the centric models [15,16], and they have the same characteristics in terms of realism and computational behavior. The fourth class of models, called *sub-cellular lattice models*, has been developed as a trade-off between the simulation speed of cell-lattice models and the geometrical flexibility of the centric models. The first sub-cellular lattice model was introduced by Graner and Glazier (*GG model*) [17].

Tuning the internal parameters of centric or lattice models is usually achieved by direct comparison of the model output and the real data that they are supposed to mimic. An important challenge is to provide automatic estimation procedures for these parameters based on statistically consistent models and algorithms. For example, it is now acknowledged that cell-cell interactions play a major role in tumorigenesis [18]. Better understanding and estimating the nature of these interactions may play a key role for an early detection of cancer. In addition, the invasive nature of some tumors is directly linked to the modification of the strength of cell-cell interactions [19]. Estimating this parameter could therefore be a step toward more accurate prognosis.

In this study, we present a statistical approach to the estimation of the strength of adhesion between cells under the DAH, based on a continuous stochastic model for cell sorting rather than a discrete one. Our model is inspired by the pioneering works of Sulsky *et al.* [20], Graner and Sawada (*GS model*) [21] and from the *GG model* [17]. In the new model, the geometry of cells is actually similar to the centric models: assuming that cell centers are known, the cells are approximated by Dirichlet cells. Using the theory of Gibbsian marked point processes [22], the continuous model can still be described through a Hamiltonian function (Section "A continuous model for DAH"). The Gibbsian marked point processes theory contains

standard procedures to estimate interaction parameters. In addition, it allows us to provide more rigorous simulation algorithms including better control of mixing properties, and it also provides a tool for establish consistency of estimators (Section "Inference procedure and model simulation"). In Section "Results and Discussion", results concerning the simulation of classical cell sorting patterns using this new model are reported, and the performances of the cell-cell adhesion strength estimator derived from this model are evaluated.

A continuous model for DAH

In this section, a new continuous model for differential adhesion is introduced. Like previous approaches, the model is based on a Hamiltonian function that describes cell-cell interactions. The Hamiltonian function incorporates two terms: an interaction term and a shape constraint term. The interaction term refers to the DAH through a differential expression of Cellular Adhesion Molecules (CAMs) weighted by the length of the membrane separating cells. This model is inspired by cell-cell interactions driven by cadherin-catenin complexes [23] which are known to be implicated in cancerous processes [24]. The main characteristic of interactions driven by cadherin-catenin complexes is that the strength of adhesion is proportional to the length of the membrane shared by two contiguous cells. This particularity is due to a zipper-like crystalline structure of cadherin interactions [25]. The constraint term relates to the shape of biological cells and prevent non-realistic cell shapes.

The proposed model uses a Dirichlet tessellation as a representation of cell geometry. The Dirichlet tessellation is entirely specified from the locations of the cell centers. Formally, we denote by x_i ($i = 1, \dots, n$) the n cell centers, where x_i is assumed to belong to a non-empty compact subset \mathcal{X} of \mathbb{R}^2 . The Dirichlet cell of x_i is denoted by $\text{Dir}(x_i)$, and is defined as the set of points (within \mathcal{X}) which are closer to x_i than to any other cell centers. Let us denote a (marked) cell configuration as

$$\varphi = \{(x_1, \tau_1), \dots, (x_n, \tau_n)\}, \quad (1)$$

where the (x_i) are the cell centers, and the (τ_i) are the corresponding cell types (or marks). The marks belong to a finite discrete space M . In the section "Results and Discussion", we consider the case where cells may be of one of the three types: $M = \{\tau_1, \tau_2, \tau_E\}$, in analogy with cell types used in [26].

The interaction term corresponds to pair potentials and it controls the adhesion forces between contiguous cells. This term is defined as follows

$$H_{\text{inter}}(\varphi) = \sum_{i \sim_{\varphi} j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j)$$

where $|\text{Dir}(x_i \cap x_j)|$ denotes the length of the contact zone between cell x_i and cell x_j . Function J is assumed to be symmetric and nonnegative

$$J : M \times M \rightarrow [0, \infty)$$

The symbol $i \sim_{\varphi} j$ means that the cells x_i and x_j share a common edge in the Dirichlet tiling built from the configuration of points in φ .

The shape constraint term corresponds to singleton potentials. It controls the form of each cells and puts a penalty on abnormally large cells. It is defined as follows

$$H_{\text{shape}}(\varphi) = \sum_i h(\text{Dir}(x_i), \tau_i)$$

where the function h is assumed to be nonnegative

$$h : \mathcal{D} \times M \rightarrow [0, \infty)$$

One specific form of the term $h(\text{Dir}(x_i), \tau_i)$, used as an example in this paper, will be described in the section "Results and Discussion". The energy functional of our model is defined as a combination of the interaction term and the shape constraint as follows

$$H(\varphi) = \theta H_{\text{inter}}(\varphi) + H_{\text{shape}}(\varphi) \tag{2}$$

where θ is a positive parameter. This parameter can be interpreted as an adhesion strength intensity, as it determines the relative contribution of cell-cell interactions in the energy. It may reflect the general state of a tissue, and its inference is relevant to applications of the model to experimental data.

Since one considers finite configurations φ on the compact set $\mathcal{X} \times \mathcal{M}$, the energy functional $H(\varphi)$ is finite ($|H(\varphi)| < \infty$). Indeed, one can notice that the area of the cell $|\text{Dir}(x_i)|$ is bounded by the area of the compact set \mathcal{X} . Coupling with the fact that h is a real-valued function, it comes that H_{shape} is bounded. Similarly, the length of a common edge $|\text{Dir}(x_i \cap x_j)|$ is bounded by the diameter of the compact set \mathcal{X} , and providing that J is a real-valued function, H_{inter} is also bounded. Moreover, since J and h are positive functions and $\theta > 0$, $H(\varphi)$ is even positive.

Before giving an inference procedure for θ , we describe the connections of our continuous model to earlier models, for which no such procedure exists. The new continuous model improves on three previous approaches by Sulsky *et al.* [20], Graner and Sawada [21] and Graner and Glazier [17]. Sulsky *et al.* proposed a model of cell sorting [20] according to a parallel between cell movements and fluid dynamics. A Dirichlet tessellation was used for modeling cells, the following Hamiltonian was introduced

$$H_S = \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| e_{i,j} \tag{3}$$

where $e_{i,j}$ is the interaction energy between cells x_i and x_j . As in our new continuous model, the length of the membrane also contributes to the energy. Graner and Sawada described another geometrical model for cell rearrangement [21]. Graner and Sawada introduced "free Dirichlet domains", which are an extension of Dirichlet domains, to overcome the excess of regular shapes in classical Dirichlet tilings. In addition to this geometrical representation, Graner and Sawada proposed an extension to Sulsky's Hamiltonian accounting for the interaction between cells and the external medium

$$H_{GS} = \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| e_{i,j} + \sum_i |\text{Dir}(x_i \cap M)| e_{i,M}$$

where $|\text{Dir}(x_i \cap M)|$ is the length of the membrane between cell x_i and the extracellular medium. This term is equal to 0 if the extracellular medium is not in the neighbourhood of x_i . While the length of the membrane is explicitly included in the models, no statistical estimate for the interaction strength was proposed in these two approaches.

In the GG model [17], a cell is not defined as a simple unit, but instead consists of several pixels. The cells can belong to three types: high surface energy cells, low surface energy cells or medium cells, which were coded as 1, 2 and -1 in the original approach. According to the DAH, Hamiltonian H_{GC} was defined as follows

$$H_{GC} = \sum_{(i,j) \sim (i',j')} J(\tau(\sigma_{ij}), \tau(\sigma_{i'j'})) (1 - \delta_{\sigma_{ij}, \sigma_{i'j'}}) + \lambda \sum_{\sigma} (a(\sigma) - A_{\tau(\sigma)})^2 \Gamma(A_{\tau(\sigma)}), \tag{4}$$

where (i, j) are the pixel spatial coordinates, σ_{ij} represents the cell to which the pixel (i, j) belongs, $\tau(\sigma_{ij})$ denotes the type of the cell σ_{ij} , and the function J characterizes the interaction intensity between two cell types (δ denoted the Kronecker symbol). The neighbourhood relationship used by Graner and Glazier is of second order which means that diagonal pixels interact. The term

$(1 - \delta_{\sigma_{ij}, \sigma'_{i'j'}})$ indicates that the interaction between two pixels within the same cell is zero. Shape constraints are modeled by the second term where λ corresponds to an elasticity coefficient, $a(\sigma)$ is the cell area and $A_{\tau(\sigma)}$ is a prior area of a cell of type $\tau > 0$. The function Γ denotes the Heaviside function and is included in the formula so that medium cells (coding -1) are not subject to the shape constraint. This model is simulated using the Boltzmann dynamics with various parameter settings and is able to reproduce many biologically relevant patterns [26]. The model introduced in this paper is a formal extension of the continuous version of the GG model [17] and also of the models introduced by Sulsky *et al.* [20] and Graner and Sawada [21]. Let us now explain in which sense this extension works. In the GG model, a cell σ is in the neighbourhood of a cell σ' as soon as a single pixel of σ is adjacent to a pixel from σ' . With this in mind, the GG model's Hamiltonian can be rewritten as

$$H_{GG} = \sum_{\sigma \sim \sigma'} |\sigma \cap \sigma'| J(\tau(\sigma), \tau(\sigma')) + \lambda \sum_{\sigma} (a(\sigma) - A_{\tau(\sigma)})^2 \Gamma(A_{\tau(\sigma)})$$

where $|\sigma \cap \sigma'|$ is the number of connected pixels between σ and σ' . The quantity $|\sigma \cap \sigma'|$ can be identified as the Euclidean length of the interaction surface between the two cells σ and σ' . Identifying cells to their centers, $|\sigma \cap \sigma'|$ can be approximated as $|\text{Dir}(x_i \cap x_j)|$. In addition, a cell area in our model matches with the area of a Dirichlet cell, which means that $a(\sigma)$ corresponds to $|\text{Dir}(x_i)|$. Using these notations, the GG energy function can be rewritten in a form similar to our Hamiltonian

$$H(\varphi) = \sum_{i \sim_{\varphi} j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}). \tag{5}$$

The second term in Equation 5 is a particular case of the shape constraint term (see Equation 2) taking

$$h(\text{Dir}(x_i), \tau_i) = \lambda (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \quad i = 1 \dots n. \tag{6}$$

To conclude this section, the new continuous model, introduced in this paper, unifies main features inspired from the three previous approaches. First, it borrows from Sulsky *et al.* the Dirichlet geometry for cells. Next it considers interactions between cells and surrounding medium as Graner and Sawada did. And finally it borrows from Graner and Glazier an additional constraint on the shape of cells. In addition, one strength of the new model

is the introduction of a new parameter which quantifies adhesion within a tissue.

Inference procedure and model simulation

An important benefit of the continuous approach is that it allows to develop consistent statistical estimation procedures for the adhesion strength parameter θ . To achieve this, we use the theory of Gibbsian marked point processes which provides a natural framework for parameter estimation (see [22,27]). Gibbsian models, according to the statistical physics terminology, have been introduced and largely studied in [28] or [29]. The idea of modeling cell configurations with point processes has been introduced in the literature by [30] and [22].

Given the energy functional defined in equation 2, we introduce a new marked point processes that have a density f , with respect to the homogeneous Poisson process of intensity 1 (as in [31], p360, l.12), of the following form

$$f(\varphi, \theta) = \frac{\exp(-H(\varphi))}{Z(\theta)} \tag{7}$$

where $Z(\theta)$ is the partition function, and θ is the parameter of interest. The probability measure for the marks is assumed to be uniform on the space of marks M . As noted in the previous section, our energy functional $H(\varphi)$ is positive and bounded. Then $H(\varphi)$ is stable in the sense of [28] (definition 3.2.1, p33). It follows that the proposed point process is well-defined as $Z(\theta)$ is bounded. A realization of such a process is called a configuration and is denoted as φ . When φ has exactly n points, we can write

$$\varphi = \{(x_1, \varphi_1), \dots, (x_n, \varphi_n)\},$$

as in Equation 1. A cell-mark couple (x_i, τ_i) is then called a *point*. We can notice that the model proposed in this study belongs to the class of the *nearest-neighbour markov point processes* introduced by [32] (see Appendix 1).

In statistics, estimating θ is usually based on a maximum-likelihood approach. However, this approach cannot be used because the computation of the partition function is in general a very hard problem apart for very small n . Hence, as in [22], we resort to a classical approximation: the pseudo-likelihood method, first introduced by Besag in the context of the analysis of dirty pictures [33] (see also [34]). For any configuration φ , the pseudo-likelihood is defined as the product over all elements of φ of the following conditional probabilities

$$PL(\varphi, \theta) = \prod_{\{x_i, \tau_i\} \in \varphi} \text{Prob}(\{x_i, \tau_i\} | \varphi \setminus \{x_i, \tau_i\}, \theta)$$

In this formula, the conditional probability of observing $\{x_i, \tau_i\}$ at x_i , given the configuration outside x_i , can be described as

$$\text{Prob}(\{x_i, \tau_i\} | \varphi \setminus \{x_i, \tau_i\}, \theta) = \frac{\exp(-H_\varphi(\{x_i, \tau_i\}, \theta))}{\int_{\mathcal{X}} \sum_{m \in M} \exp(-H_{\varphi \setminus \{x_i, \tau_i\} \cup \{y, m\}}(\{y, m\}, \theta)) dy}$$

where M corresponds to the set of the possible cell types (or marks), and where $H_\varphi(\{x_i, \tau_i\})$ represents the contribution of the marked cell $\{x_i, \tau_i\}$ in the expression of the Hamiltonian $H(\varphi)$, i.e.

$$H_\varphi(\{x_i, \tau_i\}, \theta) = \theta \sum_{j \sim_\varphi i} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) + h(\text{Dir}(x_i), \tau_i).$$

Taking the logarithm of the pseudo-likelihood leads to

$$\text{LPL}(\varphi, \theta) = \sum_{\{x_i, \tau_i\} \in \varphi} \left(-H_\varphi(\{x_i, \tau_i\}, \theta) + \log \left(\int_{\mathcal{X}} \sum_{m \in M} \exp(-H_{\varphi \setminus \{x_i, \tau_i\} \cup \{y, m\}}(\{y, m\}, \theta)) dy \right) \right) \quad (8)$$

and maximizing $\text{LPL}(\theta)$ provides an estimate of θ , namely

$$\hat{\theta}(\varphi) = \text{argmax}_\theta \text{LPL}(\varphi, \theta)$$

which can be computed using standard numerical techniques.

In order to evaluate both the statistical cell configurations according to the distribution of the Gibbsian marked point process and evaluate the statistical performances of the estimator $\hat{\theta}$, an MCMC algorithm have been implemented. The algorithm differs from the GS and GG algorithms notably since these methods were time-dependent and account for the path from the initial to final state. We apply a Metropolis-Hastings algorithm for point processes as described in [31].

At each iteration, the algorithm randomly chooses between three operations: inserting a cell within the region \mathcal{X} , deleting a cell or displacing a cell within \mathcal{X} . One iteration is detailed in the appendix (Appendix 2). From Equation 7, one can remark that only the variation in the energy is needed to compute the acceptance probability. Insertion, deletion and displacement of a cell in the configuration has been implemented using local changes as described in [35] and [36].

A second kind of benefit carried out by the use of marked point processes is to provide theoretical conditions that warrant the convergence of the simulation algorithm.

Proposition 1 Let \mathcal{X} be a compact subset of \mathbb{R}^2 and M be a finite discrete space. Let φ be a point configuration

$$\varphi = \{(x_1, \tau_1), \dots, (x_n, \tau_n)\}$$

Let us consider a Gibbsian marked point process as defined in Equation 2, and

$$H(\varphi) = \theta \sum_{i \sim_\varphi j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) + \sum_i h(\text{Dir}(x_i), \tau_i),$$

where J characterizes the interaction intensity and h the constraint on the shape of cells.

Assuming that J and h are nonnegative real-valued functions, the Markov chain generated by the simulation algorithm of the continuous model (see Appendix 2) is ergodic.

The proof of proposition 1 can be derived along the same lines as [31] (Section 4, p. 364). It can be sketched as follows. First, it is clear that the transition probabilities of the proposed algorithm satisfy Equations 3.5–3.9 in [31] (p. 361–362). Next, in order to ensure the irreducibility of the Markov chain, the density of the process has to be hereditary (Definition 3.1 in [31], p. 360). The nearest-neighbour markov property of our model ensures its hereditary. Then by adapting the proof of Corollary 2 in Tierney ([37], Section 3.1, p. 1713), it follows that the chain is ergodic.

Results and Discussion

Simulation of biological patterns

In this section, we report simulation results obtained with three marks $M = \{\tau_1, \tau_2, \tau_E\}$. We provide evidence that our model has the ability to reproduce at least three kinds of biologically observed patterns: checkerboard, cell sorting and engulfment. The constraint shape function h is borrowed from the GG model, and is defined as in Equation 6. The parameter λ controls the intensity of the shape constraint. It also acts on the density of points within the studied region \mathcal{X} . In the following of this paper we consider \mathcal{X} to be the unit disc and λ has been fixed to 10,000.

Biological tissue configurations are often interpreted in terms of surface tension parameters. For instance, checkerboard patterns are usually associated with negative surface tensions, whereas cell sorting patterns are associated with positive surface tensions [17]. When two distinct cell types are considered, the surface tension between cells with the distinct types can be defined as

$$\gamma_{12} = J(\tau_1, \tau_2) - \frac{J(\tau_1, \tau_1) + J(\tau_2, \tau_2)}{2}$$

The two marks τ_1 and τ_2 characterize "active cell types", as defined in [17], with distinct phenotypes responsible for the adhesion process. For example, phenotypes may represent different levels of expression of cadherins. In addition, active cells are surrounded by an extracellular medium modeled by cells of type τ_E . One hundred cells of type τ_E were uniformly placed on the frontier of the unit disc \mathcal{X} .

These three types are similar to the ℓ , d and M types of Glaizer and Graner [26]. Simulations were generated from the Metropolis algorithm presented in the previous section. A unique configuration was used to initialize all the simulations. This configuration is displayed in Figure 1. It consisted of about 1,000 uniformly located active cells, and the marks were also uniformly distributed in the mark space M . The target areas for active cells were equal to $A_{\tau_1} = A_{\tau_2} = 5 \times 10^{-3}$. At equilibrium, configurations were expected to consist of about $\pi/5 \cdot 10^{-3} \approx 628$ cells in the unit disc. No area constraint affected the τ_E cells and we set $A_{\tau_E} = -1$. The interaction term affecting two contiguous extra-

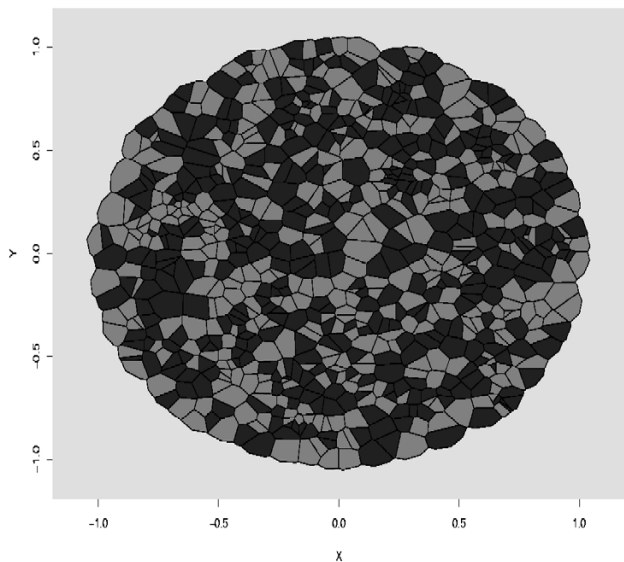


Figure 1
The initial configuration for simulating Checkerboard, Cell Sorting and Engulfment patterns. It consists of about 1,000 active cells surrounded by an extracellular medium. The active cells are randomly located in the unit sphere, and their types are randomly sampled from M . Cells of type τ_1 are colored in black while cells of type τ_2 are colored in grey. One hundred cells of type τ_E were uniformly placed on the frontier of the unit disc.

cellular cells was set to the value $J(\tau_E, \tau_E) = 0$. The adhesion strength parameter θ was fixed to $\theta = 10$.

Checkerboard patterns can be interpreted as arising from negative surface tensions. In the GG model, checkerboard patterns were generated using parameter settings that corresponded to a surface tension equal to $\gamma_{12} = -3$. Figure 2 displays the configuration obtained after 100,000 cycles of the Metropolis-Hastings algorithm, where the interaction intensities were fixed at $J(\tau_1, \tau_2) = 0$, $J(\tau_1, \tau_1) = J(\tau_2, \tau_2) = 1$ and $J(\tau_E, \tau_1) = J(\tau_E, \tau_2) = 0$. These interaction intensities correspond to a surface tension equal to $\gamma_{12} = -1$ which was of the same order as the one used in the GG model. Moreover we have $\gamma_{1E} = -1/2$ and $\gamma_{2E} = -1/2$.

In contrast, cell sorting patterns arise from positive surface tensions between active cells. In the GG model, cell sorting patterns were generated using parameter settings that corresponded to surface tensions around $\gamma_{12} = +3$. In our model, simulations were conducted using the following interaction intensities:

$J(\tau_1, \tau_2) = 1$, $J(\tau_1, \tau_1) = J(\tau_2, \tau_2) = 0$ and $J(\tau_E, \tau_1) = J(\tau_E, \tau_2) = 0$. These values correspond to $\gamma_{12} = +1$. Surface tension with extracellular medium is equal to $\gamma_{1E} = 0$ and $\gamma_{2E} = 0$. The configuration obtained after 100,000 steps cycles of Metropolis-Hastings is displayed in Figure 3.

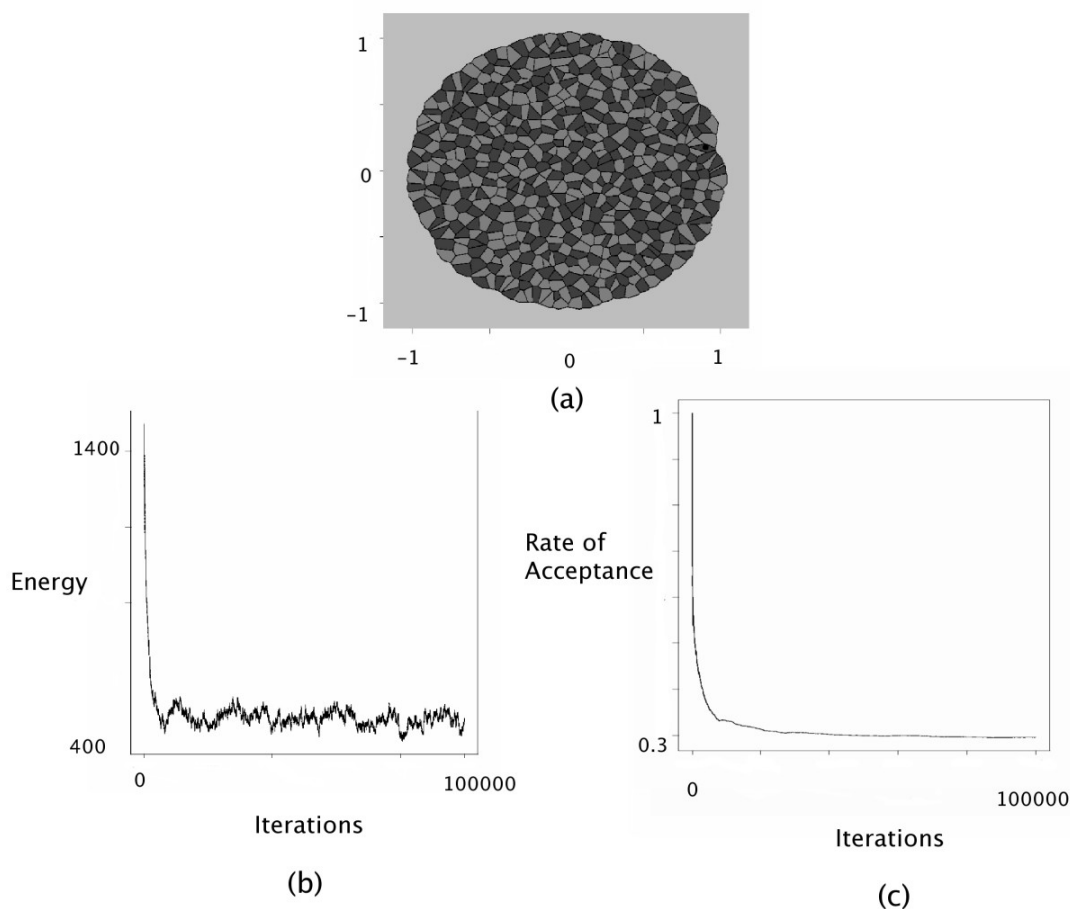
Simulations of engulfment were conducted using the following parameters: $J(\tau_1, \tau_2) = 1$, $J(\tau_1, \tau_1) = J(\tau_2, \tau_2) = 0$, $J(\tau_E, \tau_1) = 0$, $J(\tau_E, \tau_2) = 1$. These interaction intensities provide positive surface tensions between active cells, which contribute to the formation of clusters. The fact that $J(\tau_E, \tau_2)$ is greater than $J(\tau_E, \tau_1)$ ensure that τ_1 cells are more likely to be close to the extracellular medium and to surround the τ_2 cells. It is reflected by the extracellular surface tensions: $\gamma_{1E} = 0$ and $\gamma_{2E} = 1$. The results are displayed in Figure 4.

At the bottom of Figures 2, 3, 4, the evolution of the energy as well as the rate of acceptance is plotted as a function of the number cycles of Metropolis-Hastings algorithm. These curves exhibit a flat profile, which suggests that stationarity was indeed reached.

Statistical estimation of the adhesion strength parameter

In this section, we study the sensitivity of simulation results to the adhesion strength parameter θ , and we report the performances of the maximum pseudo-likelihood estimator $\hat{\theta}$.

To assess the influence of θ on simulations, three values were tested: $\theta = 1$, $\theta = 5$ and $\theta = 10$. The results are presented for simulations of checkerboard, cell sorting and

**Figure 2**

Checkerboard simulation. The interaction intensities were chosen as follows: $J(\tau_1, \tau_1) = 1, J(\tau_2, \tau_2) = 1, J(\tau_1, \tau_2) = 0, J(\tau_1, \tau_E) = 0, J(\tau_2, \tau_E) = 0$ and $J(\tau_E, \tau_E) = 0$. (a) The configuration obtained after 100,000 iterations with $\theta = 10$. (b) The decrease of the energy as a function of the iteration steps. (c) The evolution of the acceptance rate as a function of the iteration steps.

engulfment patterns. In each case, the interaction intensities were set as in the previous paragraph.

We ran the Metropolis algorithm for 100,000 cycles. This number is sufficient to provide a flat profile of energy and rate of acceptance. The final configurations, in checkerboard, cell sorting and engulfment, are displayed in Figure 5. Either for checkerboard or for cell sorting simulations, we observe a gradual evolution when θ increases. For $\theta = 1$, the marks seem to be randomly distributed, for $\theta = 5$ a small inhibition is visible in the checkerboard simulation, small clusters appear in the cell sorting pattern and black cells start to surround white cells in the engulfment simulation. Finally, for $\theta = 10$ the stronger inhibition between cells with the same types provides a more pronounced checkerboard pattern, larger clusters are obtained in cell sorting and black cells completely engulf white cells.

For each value of θ , 100 replicates of cell sorting, checkerboard and engulfment were generated from which the mean and the variance of $\hat{\theta}$ were estimated. Each replicate consisted in 100,000 cycles started from independent initial configurations and sampled from uniform distributions. The number of active cells was sampled from the interval [500,1500]. Cells were uniformly located within the unit disk and types were uniformly assigned to each cell. Table 1 summarizes the results obtained for θ in the range [1, 20]. For cell sorting, the bias is weak for all values of θ , while for checkerboard the bias seems to be slightly higher. The results are similar regarding the variance. It is higher for checkerboard than for cell sorting. Under the engulfment model, the estimator $\hat{\theta}$ seemed to systematically slightly overestimate θ . Variance under the engulfment model is of the same order as the variance in

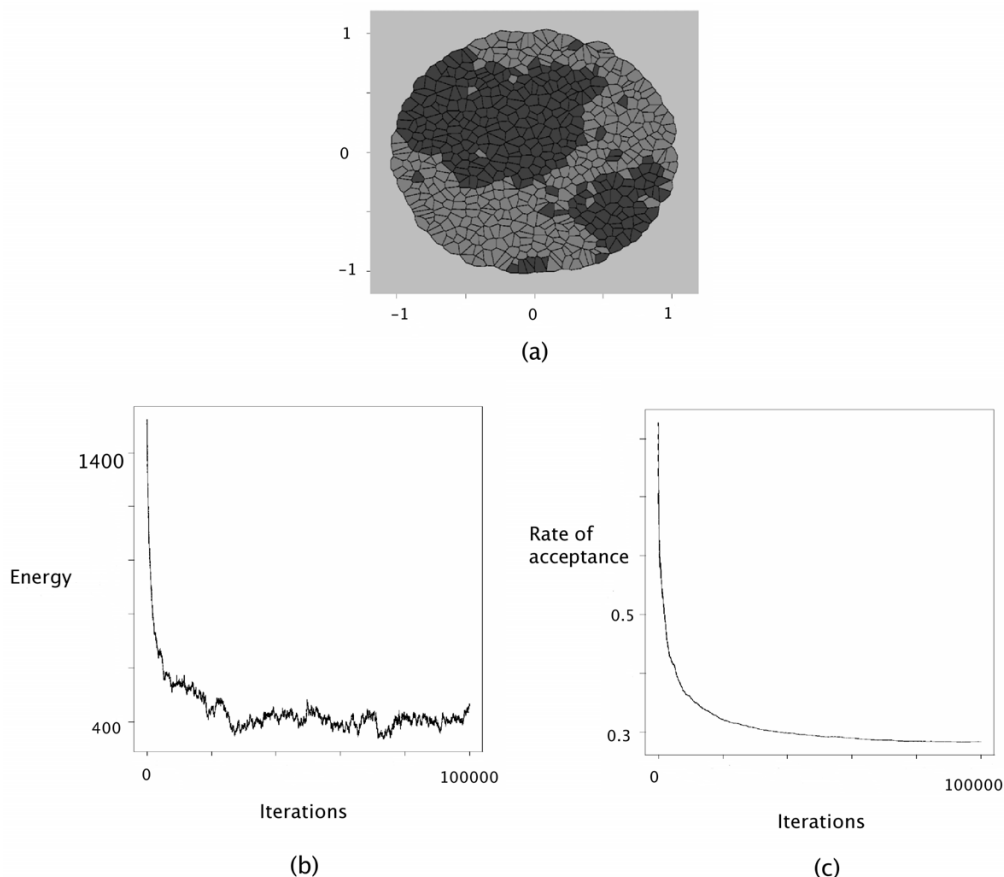


Figure 3
 Cell Sorting simulation. The interaction intensities were chosen as follows: $J(\tau_1, \tau_1) = 0, J(\tau_2, \tau_2) = 0, J(\tau_1, \tau_2) = 1, J(\tau_1, \tau_E) = 0, J(\tau_2, \tau_E) = 0$ and $J(\tau_E, \tau_E) = 0$. (a) The configuration obtained after 100,000 iterations with $\theta = 10$. (b) The decrease of the energy as a function of the iteration steps. (c) The evolution of the acceptance rate as a function of the iteration steps.

cell sorting. Finally, in the three model, the variance increased as θ increased. The estimates can be considered as accurate for moderate values of $\theta (\approx 10)$, as the pseudo-likelihood may provide significant bias in cases of strong interaction [38].

Experimental data

Estimation of the adhesion strength was also performed on a real data example. We used data from Pizem *et al.* ([39]), who measured survivin and beta-catenin markers in Human medulloblastoma. These markers are known to be involved in complexes that regulate adhesion between contiguous cells. An image analysis, analogous to the analysis performed in [40], was achieved to extract the locations of cell nuclei and the levels of expression of markers in cells. The expression levels were used to define cell types as displayed in Figure 6. The resulting image is relevant to a cell sorting pattern, and we used the set of J parameters that corresponded to this pattern.

The estimate of θ was computed as $\hat{\theta} \approx 5.27$. This value provides evidence that the model is able to detect large clusters (black cell clusters here) and that white cells may be surrounded by black cells. The estimated value $\hat{\theta}$ was then tested as input to the simulation algorithm, and the resulting spatial pattern is displayed in Figure 7. Comparing the real tissue and the cell sorting pattern simulated with the estimated interaction strength makes clear that the model provides a good fit to the data and that θ estimation is consistent.

Conclusion

In this study, we presented an approach to cell sorting based on marked point processes theory. It proposes a continuous geometry for tissues using a Dirichlet tessellation and an energy functional expressed as the sum of two terms: an interaction term between two contiguous cells weighted by the length of the membrane and a cell shape

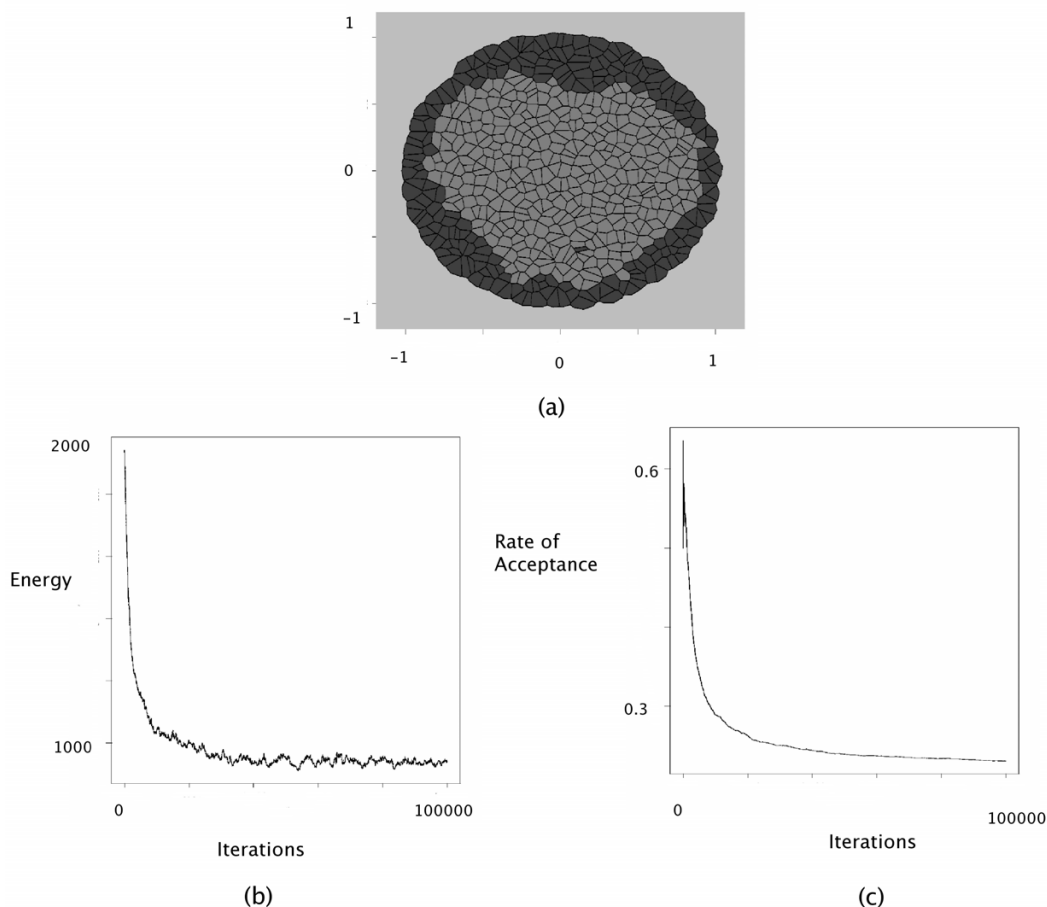


Figure 4

Engulfment simulation. The interaction intensities were chosen as follows: $J(\tau_1, \tau_1) = 0$, $J(\tau_2, \tau_2) = 0$, $J(\tau_1, \tau_2) = 1$, $J(\tau_1, \tau_E) = 0$, $J(\tau_2, \tau_E) = 1$ and $J(\tau_E, \tau_E) = 1$. (a) The configuration obtained after 100,000 iterations with $\theta = 10$. (b) The decrease of the energy as a function of the iteration steps. (c) The evolution of the acceptance rate as a function of the iteration steps.

constraint term. Such models, where interactions are weighted by the length of the membrane, have already been considered in the literature, first by Sulsky *et al.* [20] and next by Graner and Sawada [21]. Based on Honda's studies that showed that the geometry of Dirichlet cells was in agreement with biological tissues [41,42], these earlier models also used a continuous geometry of cells. These authors were interested in formulating a dynamical model which determines not only the equilibrium state but the path from the initial state to final state. These two approaches introduced systems of differential equations to simulate cell patterns.

Although the previous approaches contained the main ingredients to model simulation, they were not well-adapted to perform statistical estimation of interaction parameters. Furthermore, Graner and Sawada reported two limitations of their approach. First, because the GS

model is not stochastic, it does not explore the set of possible configurations ([21], p.497, l.10). Next Graner and Sawada stressed that their simulation algorithm suffers from instability because of its lack of theoretical control ([21], p.497, l.15). Graner and Glazier proposed Boltzmann dynamics and were interested in the time needed to achieve desired configurations. However, there is no warranty that their Markov chain has correct mixing properties, and the sensitivity of their method to the discretization scale remains to be studied. Because of discretization, detailed balance condition and cell connexity did not seem to hold in the GG model. GG's approach cannot be easily adapted to define inference procedures.

Our study is not the first attempt to propose statistical procedures for estimating interaction strength parameters in tissues. In [13], two statistics have been introduced to measure the degree of spatial cell sorting in a tissue where

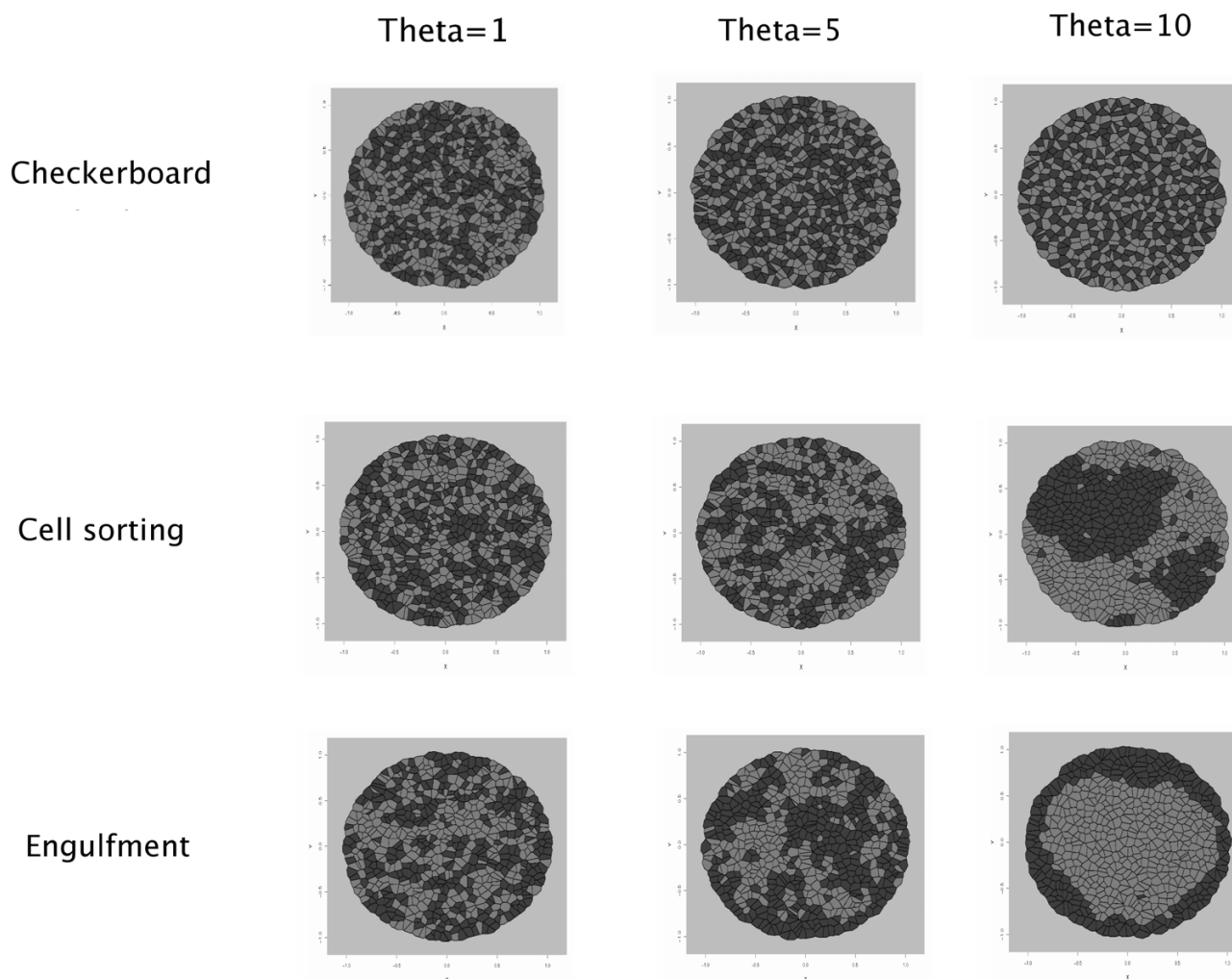


Figure 5
Influence of θ in simulations. Final configurations using three different values for θ . Simulations gradually corresponds to either a checkerboard, large clusters or engulfment.

cells are of types black and white. Cell sorting can be quantified by the fraction of black cells in the nearest neighborhood of single black cell and the number of isolated black cells. Although these two statistics have been recently used to study the role of cadherins in tissue segregation [43], their practical application requires cells to be pixels within a lattice ([13] and [43]). Their capacity to quantify cell sorting has been studied using a cell-lattice model where all cells have the same geometry, hypothesis which does not fit with the zipper-like structure of cadherins [25].

In contrast to these approaches, the mathematical background of marked point processes allows the establishment of a statistical framework. In this study, we have shown that our model was able to reproduce biologically relevant cell patterns such as checkerboard, cell sorting

and engulfment. Checkerboard pattern formation was investigated in a simulation study of the sexual maturation of the avian oviduct epithelium [44]. Cell sorting is a standard pattern of mixed heterotypic aggregates. Experimental observations of this phenomena were reported by Takeuchi *et al.* [45] and Armstrong [1]. Engulfment of a tissue by another one was studied by Armstrong [1] and Foty *et al.* [46]. This phenomenon is a direct consequence of adhesion processes between the two cell types and the extracellular medium. These cell patterns were also simulated by pioneering studies ([17,20,21]).

Furthermore, the present model has been built so that it includes the strength of cell-cell adhesion as a statistical parameter. We proposed and validated an inference procedure based on the pseudo-likelihood. The statistical errors remain small in cell sorting simulations. In check-

Table 1: Mean and Variance of $\hat{\theta}$, maximum of Pseudo-likelihood, for checkerboard, cell sorting and engulfment simulations. The evaluation was achieved using 100 simulations of each case.

| | Checkerboard | | Cell Sorting | | Engulfment | |
|---------------|--------------|----------|--------------|----------|------------|----------|
| | Mean | Variance | Mean | Variance | Mean | Variance |
| $\theta = 1$ | 0.99 | 0.39 | 1.12 | 0.34 | 1.27 | 0.32 |
| $\theta = 3$ | 3.34 | 0.48 | 3.03 | 0.43 | 3.08 | 0.41 |
| $\theta = 5$ | 5.15 | 0.92 | 5.37 | 0.45 | 5.23 | 0.48 |
| $\theta = 8$ | 8.42 | 1.03 | 8.11 | 0.92 | 8.21 | 0.67 |
| $\theta = 10$ | 10.58 | 1.36 | 10.26 | 0.67 | 10.22 | 1.02 |
| $\theta = 12$ | 12.12 | 1.55 | 12.14 | 0.92 | 12.18 | 0.94 |
| $\theta = 15$ | 15.51 | 1.72 | 14.94 | 1.64 | 15.48 | 1.12 |
| $\theta = 20$ | 19.52 | 2.15 | 20.08 | 1.51 | 20.28 | 1.82 |

erboard simulations, bias and variance are slightly higher than for cell sorting but still reasonable. The bias is also weak in engulfment simulations. Further improvements of this approach would require a longer study of the properties of the point process model. In particular, the other interaction parameters can also be estimated in the same way that θ is. Although we did not assess the performances of these estimators, we believe that they would be

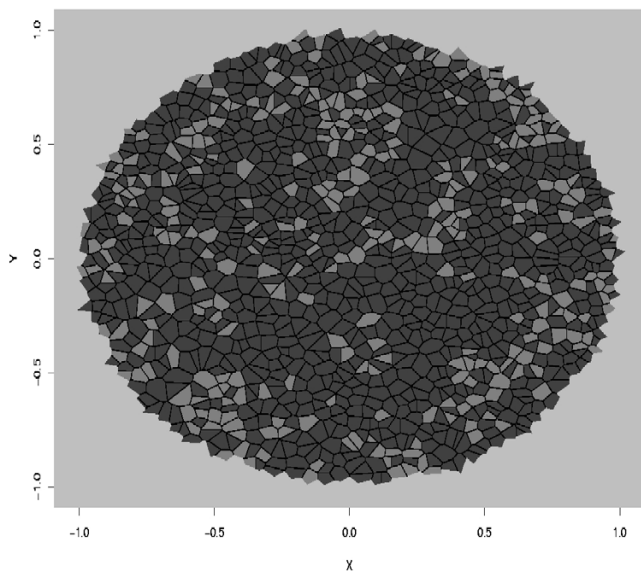


Figure 6
Experimental data. The statistical procedure was conducted using the interactions parameters for cell sorting patterns: $J(\tau_1, \tau_1) = 1, J(\tau_2, \tau_2) = 1, J(\tau_1, \tau_2) = 0, J(\tau_1, \tau_E) = 0, J(\tau_2, \tau_E) = 0$ and $J(\tau_E, \tau_E) = 0$. λ was set to 10,000. We obtained $\hat{\theta} \approx 5,27$.

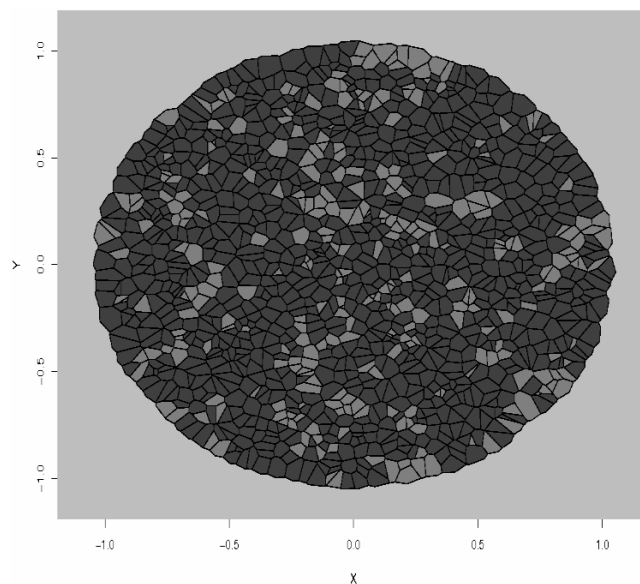


Figure 7
Simulated data. Using the interactions parameters for cell sorting patterns: $J(\tau_1, \tau_1) = 1, J(\tau_2, \tau_2) = 1, J(\tau_1, \tau_2) = 0, J(\tau_1, \tau_E) = 0, J(\tau_2, \tau_E) = 0$ and $J(\tau_E, \tau_E) = 0$. The parameter θ used for the simulation is the estimated value from experimental data $\theta = 5.27$ (see Figure 6). The estimated interaction parameter $\hat{\theta} \approx 5.51$.

useful for analyzing tissues arrays, as generated by high-throughput cancer studies [47].

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

OF and ME both provided the basic ideas of the project. ME was responsible for the development of the proposed method and carried out the simulation analysis. ME and OF equally contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Appendix

Appendix I

In this section we prove that the point process introduced in this paper is a *nearest-neighbour markov point process* which has been defined in [32] (p.107 Definition 4.9).

We use Theorem 4.13 (p.108) in [32] (analogue of the Hammersley-Clifford-Ripley-Kelly theorem [48] for nearest-neighbour markov point processes) which is as follows

Theorem 1 Let H be an hereditary subset of the set of finite configurations in $\mathcal{X} \times \mathcal{M}$. Let \sim_φ be a neighbour relation with consistency conditions (C1)–(C2) (4.7 p.106 in [32]) hold, $\varphi \in H$. Then a function $g : H \rightarrow [0, \infty)$ is a Markov function if and only if

$$g(x) = \prod_{\phi \subset \varphi} \Psi(\phi | \varphi)$$

for all $\varphi \in H$, where Ψ is an interaction function.

As proved in [32] (Appendix A1, p116), the set of Dirichlet configurations is hereditary and satisfies properties (C1) and (C2) (see paragraphs 2.5 p94 and 4.7 p106 in [32]).

Let φ be a finite configuration in \mathcal{X} and \sim_φ its Dirichlet neighbourhood. Let Ψ be a function defined over all cliques $\phi \in \varphi$ as follows

$$\Psi(\{x_i, \tau_i\}) = \frac{\exp(-\lambda h(\text{Dir}(x_i, \tau_i)))}{Z(\theta)} > 0 \text{ for each point } (x_i, \tau_i) \in \varphi$$

$$\Psi(i, j) = \frac{\exp(-\theta | \text{Dir}(x_i \cap x_j) | I(\tau_i, \tau_j))}{Z(\theta)} > 0 \text{ for each couple } \{(x_i, \tau_i), (x_j, \tau_j)\} \text{ such as } x_i \sim_\varphi x_j$$

$\Psi(\text{clique}) = 1$ for all cliques with three or more points

As mentioned in the section "A new model for DAH", one can remark that for each marked cell (x_i, τ_i) in a configuration $\varphi \in \mathcal{X} \times \mathcal{M}$

$$\Psi(\{x_i, \tau_i\}) = \frac{\exp(-\lambda h(\text{Dir}(x_i, \tau_i)))}{Z(\theta)} > 0$$

, and for each couple $\{(x_i, \tau_i), (x_j, \tau_j)\}$, such as $x_i \sim_\varphi x_j$

$$\Psi(i, j) = \frac{\exp(-\theta | \text{Dir}(x_i \cap x_j) | I(\tau_i, \tau_j))}{Z(\theta)} > 0$$

This shows that Ψ is an *interaction function* in the sense of [32] (definition 4.11 p108). One can easily note that $\prod_{c \in \text{cliques}(\varphi)} \Psi(c) = f(\varphi, \theta)$. Then Theorem 1 holds for the interaction function proposed in this paper which leads to f is a Markov function.

Since the neighbourhood depends on the configuration, our point process is a *nearest-neighbour markov point process* as defined in [32] (Definition 4.9 p107).

Appendix 2

Here is the description of one iteration in the MCMC algorithm proposed in this paper. Let us denote φ the config-

uration just before the current iteration and ψ the proposed new configuration.

- With probability $p(\varphi)$ – Displacement
 - One point $\{x_i, \tau_i\}$ is chosen with probability $d(\varphi, \{x_i, \tau_i\})$
 - The proposal distribution for the moving point is $b(\varphi, \{x, \tau\})$
 - $\psi = \varphi \subset \{x_i, \tau_i\} \cup \{x, \tau\}$
- Else with probability $(1 - p(\varphi))q(\varphi)$ – Insertion
 - The proposal distribution for the new point is $b(\varphi, \{x, \tau\})$
 - $\psi = \varphi \cup \{x, \tau\}$
- Else with probability $(1 - p(\varphi))(1 - q(\varphi))$ – Deletion
 - The deleted point $\{x_i, \tau_i\}$ is chosen with probability $d(\varphi, \{x_i, \tau_i\})$
 - $\psi = \varphi \subset \{x_i, \tau_i\}$
- The proposal configuration ψ is accepted with the acceptance probability A [31], as follows

$$A(\psi | \varphi) = \min(1, f(\psi)/f(\varphi))$$

In the algorithm proposed in this paper, we used

$$p(\varphi) = 1/2 \text{ and } q(\varphi) = 1/2 \text{ for all } \varphi \in \mathcal{X} \times \mathcal{M}$$

$d(\varphi, \{x_i, \tau_i\}) = 1/n$ for all $\{x_i, \tau_i\} \in \varphi$ and where n is the number of points in φ $b(\varphi, \{x, \tau\}) = \frac{1}{\rho(\mathcal{X})\%(\mathcal{M})}$ where ρ

ϱ is the intensity of the underlying marked Poisson process

Acknowledgements

The authors are grateful to anonymous referees for very useful comments. The authors would like to thank Jean-Michel Billiot and Remy Drouilhet for insightful discussions and to acknowledge Jose Pizem for providing experimental data. This work has been partly granted by the ALPB project supported by the Institut d'Informatique et de Mathématiques Appliquées de Grenoble.

References

1. Armstrong PB: **Cell sorting out: the self assembly of tissues in vitro.** *Critical Reviews in Biochemistry and Molecular Biology* 1989, **24**:119-149.
2. Holfreter J: **Experimental studies on the development of the pronephros.** *Rev Can Biol* 1944, **3**:220-250.

3. Gierer A, Berking S, Bode H, David CN, Flick K, Hansmann G, Schaller H, Trenkner E: **Regeneration of hydra from reaggregated cells.** *Nat New Biol* 1972, **239**:98-101.
4. Steinberg MS: **On the mechanism of tissue reconstruction by dissociated cells, I. Population kinetics, differential adhesiveness, and the absence of directed migration.** *Proceedings of the National Academy of Science* 1962, **48**:1577-1582.
5. Steinberg MS: **Mechanism of tissue reconstruction by dissociated cells, II. time-course of events.** *Science* 1962, **137**:762-763.
6. Steinberg MS: **On the mechanism of tissue reconstruction by dissociated cells, III. Free energy relations and the reorganization of fused, heteronomic tissue fragments.** *Proceedings of the National Academy of Science* 1962, **48**:1769-1776.
7. Steinberg MS: **Reconstruction of tissues by dissociated cells.** *Science* 1963, **141**:401-408.
8. McClay DR, Ettensohn CA: **Cell adhesion in morphogenesis.** *Annu Rev Cell Biol* 1987, **3**:319-345.
9. Nubler-Jung K, Mardini B: **Insect epidermis: polarity patterns after grafting result from divergent cell adhesions between host and graft tissue.** *Development* 1990, **110**:1071-1079.
10. Newman SA, Comper WD: **'Generic' physical mechanisms of morphogenesis and pattern formation.** *Development* 1990, **110**:1-18.
11. Foty RA, Steinberg MA: **The differential adhesion hypothesis: a direct evaluation.** *Developmental Biology* 2005, **278**:255-263.
12. Brodland GV: **Computational modeling of cell sorting, tissue engulfment, and related phenomena: a review.** *Applied Mechanics Reviews* 2004, **57**:47-76.
13. Mochizuki A, Iwasa Y, Takeda Y: **A stochastic model for cell sorting and measuring cell-cell-adhesion.** *Journal of Theoretical Biology* 1996, **179**:129-146.
14. Honda H, Tanemura M, Imayama S: **Spontaneous architectural organization of mammalian epiderms from random cell packing.** *The Journal of Investigative Dermatology* 1996, **106**(2):312-315.
15. Nagai T, Honda H: **A dynamic cell model for the formation of epithelial tissue.** *Philosophical Magazine B* 2001, **81**(7):699-719.
16. Honda H, Tanemura M, Nagai T: **A three-dimensional vertex dynamics cell model of space-filling polyhedra simulating cell behavior in a cell aggregate.** *Journal of Theoretical Biology* 2004, **226**:439-453.
17. Graner F, Glazier JA: **Simulation of biological cell sorting using a two-dimensional extended Potts model.** *Physical Review Letters* 1992, **69**(13):2013-2016.
18. Barker N, Clevers H: **Catenins, Wnt signaling and cancer.** *BioEssays* 2000, **22**:961-965.
19. Lozano E, Betson M, Brage VMM: **Tumor progression: Small GTPases and loss of cell-cell adhesion.** *BioEssays* 2003, **25**:452-463.
20. Sulsky D, Childress S, Percus JK: **A model of cell sorting.** *Journal of Theoretical Biology* 1984, **106**(3):275-301.
21. Graner F, Sawada Y: **Can surface adhesion drive cell-rearrangement? Part II: a geometrical model.** *Journal of Theoretical Biology* 1993, **164**(4):477-506.
22. Van-Lieshout MNM: *Markov Point Processes and their Applications* Imperial College Press; 2000.
23. Geiger B, Ayalon O: **Cadherins.** *Annual Review of Cell Biology* 1992, **8**:307-332.
24. Foty RA, Steinberg MA: **Cadherin-mediated cell-cell adhesion and tissue segregation in relation to malignancy.** *International Journal of Developmental Biology* 2004, **48**:397-409.
25. Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, Grubel G, Legrand JF, Als-Nielsen J, Colman DR, Hendrickson WA: **Structural basis of cell-cell adhesion by cadherins.** *Nature* 1995, **374**:327-337.
26. Glazier JA, Graner F: **Simulation of differential adhesion driven rearrangement of biological cells.** *Physical Review E* 1993, **47**(3):2128-2154.
27. Møller J, Waagepetersen RP: *Statistical Inference and Simulation for Spatial Point Processes* Boca Raton: Chapman and Hall/CRC; 2003.
28. Ruelle D: *Statistical Mechanics: rigorous results* New-York: Benjamin; 1969.
29. Preston CJ: **Random Fields.** *Lecture Notes in Mathematics* 1976, **534**:1-200.
30. Bertin E, Billiot JM, Drouilhet R: **Spatial Delaunay Point Processes.** *Stochastic Models* 1999, **15**:181-199.
31. Geyer CJ, Møller J: **Simulation procedures and likelihood inference for spatial point processes.** *Scandinavian Journal of Statistics* 1994, **21**:359-373.
32. Baddeley A, Møller J: **Nearest-Neighbour Markov point processes and random sets.** *International Statistical Review* 1989, **2**:89-121.
33. Besag J: **Statistical analysis of non-lattice data.** *The Statistician* 1975, **24**:192-236.
34. Comets F: **On consistency of a class of estimation for exponential families of Markov random fields on the lattice.** *The Annals of Statistics* 1992, **20**:455-468.
35. Watson DF: **Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes.** *The Computer J* 1981, **24**:167-172.
36. Bertin E: **Diagrammes de Voronoi 2D et 3D, applications en analyse d'images.** In *PhD thesis* Université Joseph Fourier – Grenoble I; 1994.
37. Tierney L: **Markov chains for exploring posterior distributions.** *The Annals of Statistics* 1994, **22**:1701-1762.
38. Diggle PJ, Fiksel T, Grabarnik P, Ogata Y, Stoyan D, Tanemura M: **On parameter estimation for pairwise interaction point processes.** *International Statistical Review* 1994, **62**:99-117.
39. Pizem J, Cör A, Zdravec-Zaletel L, Popovic M: **Survivin is negative prognostic marker in medulloblastoma.** *Neuropathol Appl Neurobiol* 2005, **31**:422-428.
40. Emily M, Morel D, Marcelpoil R, Francois O: **Spatial correlation of gene expression measures in tissue microarray core analysis.** *Computational and Mathematical Methods in Medicine* 2005, **6**:33-39.
41. Honda H: **Description of cellular patterns by Dirichlet domains: The two-dimensional case.** *Journal of Theoretical Biology* 1978, **72**:523-543.
42. Honda H: **Geometrical models for cells in tissues.** *International Review of Cytology* 1983, **81**:191-248.
43. Takano R, Mochizuki A, Iwasa Y: **Possibility of tissue segregation caused by cell adhesion.** *Journal of Theoretical Biology* 2003, **221**:459-474.
44. Honda H, Yamanaka H, Eguchi G: **Transformation of a polygonal cellular pattern during sexual maturation of the avian oviduct epithelium: Computer simulation.** *Journal of Embryology and Experimental Morphology* 1986, **98**:1-19.
45. Takeuchi I, Kakutani T, Tasaka M: **Cell behavior during formation of prestalk/prespore pattern in submerged agglomerates of Dictyostelium.** *Developmental Genetics* 1988, **9**:607-614.
46. Foty RA, Pflieger CM, Forgacs G, Steinberg MS: **Surface tensions of embryonic tissues predict their mutual envelopment behavior.** *Development* 1996, **122**:1611-1620.
47. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP: **Tissue microarray for high-throughput molecular profiling of tumor specimens.** *Nature Medicine* 1998, **4**:844-847.
48. Ripley BD, Kelly FP: **Markov point processes.** *Journal of the London Mathematical Society* 1977, **15**:188-192.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

