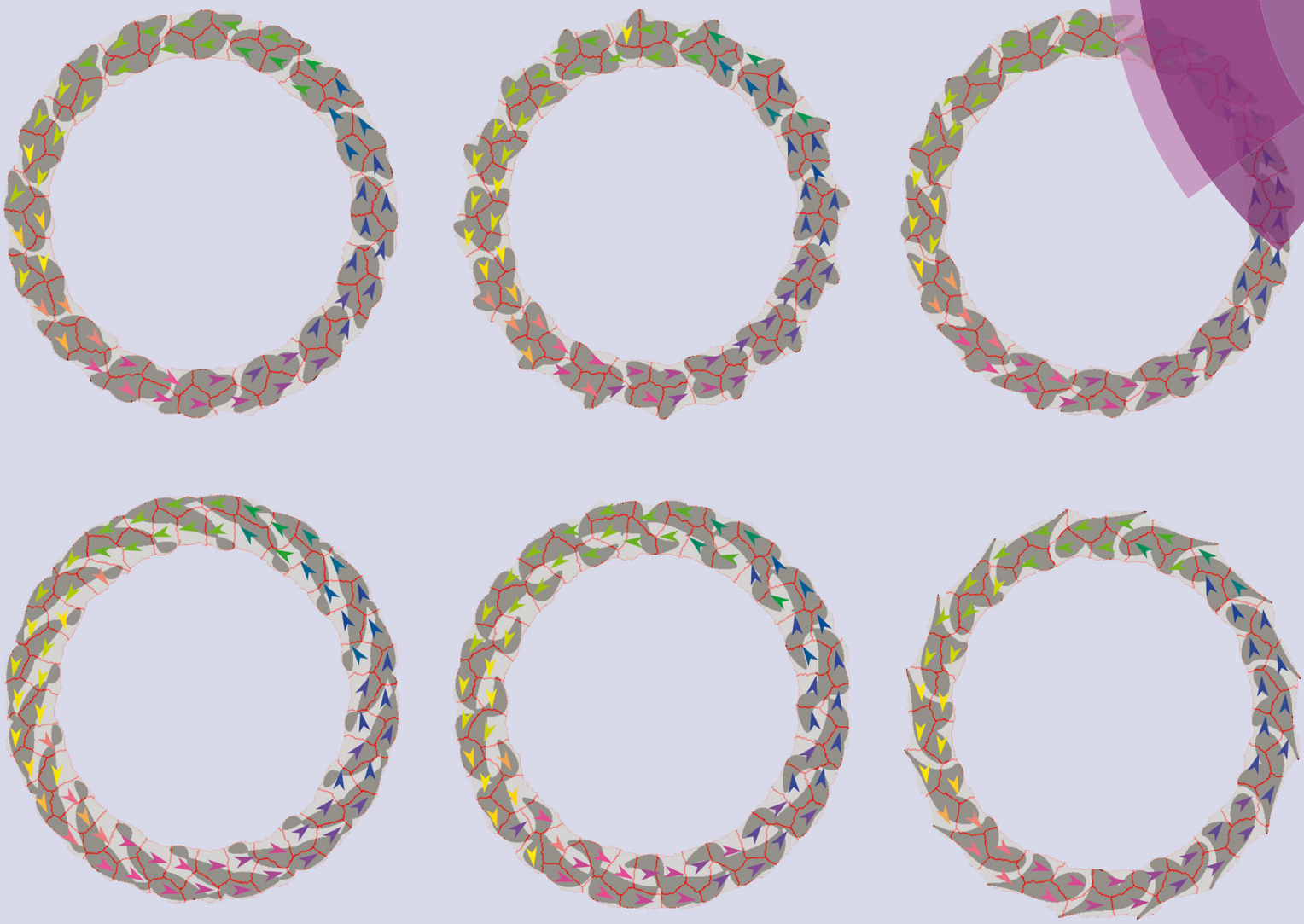
Volume 8 | Number 7 | July 2016 | Pages 731–806

Integrative Biology

Interdisciplinary approaches for molecular and cellular life sciences

www.rsc.org/ibiology



ISSN 1757-9694

ROYAL SOCIETY OF CHEMISTRY

PAPER

Philipp J. Albert and Ulrich S. Schwarz

Optimizing micropattern geometries for cell shape and migration with genetic algorithms

Integrative Biology

PAPER

View Article Online
View Journal | View Issue

Cite this: *Integr. Biol.*, 2016, 8, 741

Optimizing micropattern geometries for cell shape and migration with genetic algorithms†

Philipp J. Albert and Ulrich S. Schwarz*

Received 11th April 2016,
Accepted 13th June 2016

DOI: 10.1039/c6ib00061d

www.rsc.org/ibiology

Adhesive micropatterns have become a standard tool to control cell shape and function in cell culture. However, the variety of possible patterns is infinitely large and experiments often restrict themselves to established designs. Here we suggest a systematic method to establish novel micropatterns for desired functions using genetic algorithms. The evolutionary fitness of a certain pattern is computed using a cellular Potts model that describes cell behavior on micropattern. We first predict optimal patterns for a desired cell shape. We then optimize ratchet geometries to bias cell migration in a certain direction and find that asymmetric triangles are superior over the symmetric ones often used in experiments. Finally we design geometries which reverse the migration direction of cells when cell density increases due to cell division.

Insight, innovation, integration

Adhesive micropatterns are increasingly used to control cellular functions, but little attention has been devoted before to their rational design. Here we introduce a computational procedure based on genetic algorithms to predict adhesive geometries that promote a desired cell behaviour, for example migration on a ratchet pattern into a given direction. Our approach is computationally cheap because it builds on a very efficient cellular Potts model. As long as a good model description exists for cell behaviour, our approach can be used to predict adhesive micropatterns for virtually any task of interest.

1 Introduction

Micropatterns (MP) have become a standard tool to study cells in culture under controlled conditions[1,2] and are increasingly combined with quantitative analysis and modeling.[3–6] The application of MP range from single cell experiments up to systems consisting of several hundreds of cells. For single cells, it was first found that the geometry of extracellular matrix ligand distribution determines cell fate.[7] Later studies focused strongly on the effect of MP-geometry on the internal organisation of the cells.[8] It was further demonstrated that cell shape directly relates to cell mechanics.[9,10] To address cell mechanics, MP are increasingly combined with soft elastic substrates to measure traction forces of single cells[11–15] and multicellular systems.[16,17] The extension to multicellular systems occurs naturally when cells divide during the time course of the experiment. Here, MP have revealed an influence of the extracellular matrix (ECM) geometry on the cell division axis.[18,19] Larger multicellular systems have been investigated mainly in regard to the influence of geometry on collective cell migration.[20,21] Collective migration experiments are often combined with a removable barriers approach,[22] *e.g.* to study migration into nonadhesive regions[16] or the formation of leader cells.[23]

The influence of geometry on cell migration has been addressed in much detail for single cells. MP polarize the internal organization of cells[24] and determine where lamellipodia form.[25,26] On square and triangular patterns lamellipodia form most likely at one of the corners.[25] When the same cell line is grown on teardrop shaped MP, lamellipodia form most likely at the blunt end and cells move in this direction for a certain time once the pattern has been removed.[26]

However, for a combination of several such patterns, single cell polarization is only one out of several important factors, as has been demonstrated with ratchet patterns.[27] A ratchet pattern is composed of a asymmetric building block repeated several times, *e.g.* in a linear fashion, on a circle of variable radius or in the shape of a square. For a square arrangement of teardrop patterns, it was found that the direction of migration depends not only on single cell polarization, but also on the availability of adhesive ligands on the neighboring pattern.[28] If the teardrop pattern is arranged in a square in such a way that a new pattern is not easily accessible from the blunt side, then

Institute for Theoretical Physics and BioQuant, Heidelberg University, Philosophenweg 19, 69120 Heidelberg, Germany.
E-mail: Ulrich.Schwarz@bioquant.uni-heidelberg.de; Fax: +49-6221-54-9331; Tel: +49-6221-54-9399

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6ib00061d

cells can also migrate in the other direction. If the availability of ligands is ensured by appropriate pattern positioning, then the cells can maintain their single cell polarization on teardrop patterns also over the gap.[29]

Importantly, these effects strongly depend on cell type. For example, it has been shown that epithelial cells migrate on very similar square arrangements of teardrop patterns in the other direction[30] than mesenchymal cells such as 3T3-fibroblasts,[28] most likely because they protrude and stabilize lamellipodia in different ways. The behaviour of these epithelial cells can be further modulated using teardrop patterns with elongation (spear-shaped).[31] For a linear arrangement of triangular shapes, it was found that cell migration of mesenchymal cells such as 3T3-fibroblasts proceeds mainly in the direction of the tip,[32,33] most likely because lamellipodia are suppressed at the blunt end if the triangles are somehow larger than the typical spread area. Here again it was observed that these effects are cell-type specific, because different cell types can traverse the same ratchet pattern in different directions.[32]

Surprisingly, the range of MP usually chosen for cell studies is rather limited and often shaped by past successes. Little attention has been devoted to a systematic approach to design MP for specific functions. Fig. 1A shows that experiments usually address the so-called direct problem, that is they show how cell shape and mechanics adapt to a given MP. For example, the crossbow pattern seen here is often used to study cell processes related to cell polarization, because it forces the cell to organize its cytoskeleton in an asymmetric manner.

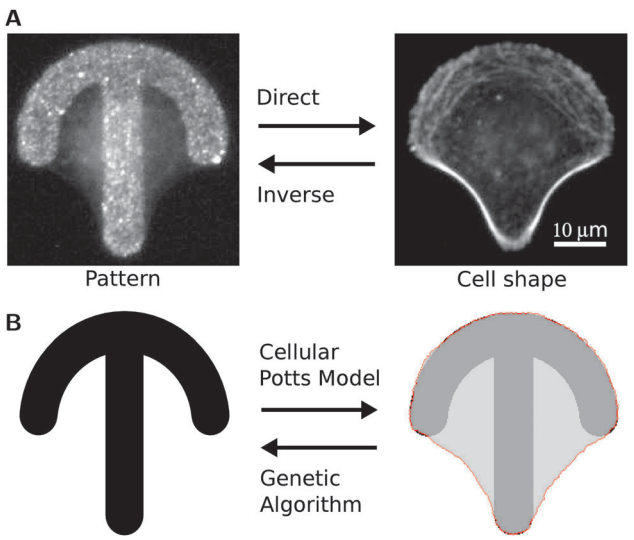


Fig. 1 Optimizing micropatterns. (A) Given a micropattern, the corresponding cell shape can be obtained in experiments by letting a cell spread on the pattern. The inverse problem of obtaining the pattern leading to a given cell shape is an open issue. Image courtesy of Dr Vytaute Starkuviene-Erfle. (B) Direct and inverse problem solved with a model and optimization approach. For the direct problem of predicting cell behavior on micropatterns (MP) a cellular Potts model (CPM) is used. The inverse problem of optimizing a pattern layout for a specific task is addressed with genetic algorithms (GA). This algorithm suggests layouts systematically and tests with the CPM for how well they are suited for a desired task.

Although prominent patterns like the crossbow are very instructive, in principle the number of possible geometries is infinite and one might wonder which other patterns might be interest for experiments. The challenge of predicting the best MP for a desired biological function (*e.g.* a desired shape or migration direction) can be called the inverse problem, compare Fig. 1A.

In principle not only the direct problem, but also the inverse problem can be addressed experimentally, for example by high-throughput screening. However, such an approach is unrealistically cost- and time-consuming and would still require a rational approach to generate relevant adhesive patterns. We therefore set out to address this challenge using a model and optimization approach. As indicated in Fig. 1B, we approach this problem using the cellular Potts model (CPM) for the direct problem and genetic algorithms (GA) for the inverse problem, respectively.

CPMs where originally developed to describe cell sorting by the differential adhesion hypothesis.[34,35] They have found a wide range of applications,[36,37] including tumor invasion,[38] cell arrangement in the Drosophila retina,[39] keratocyte migration[40] and single cells on dot MP.[41] The versatile nature of the CPM used to solve the direct problem allows to optimize MP for virtually any task. Here we apply a version of the CPM that we have developed before to describe spreading and shape of single cells[42] and migrating multicellular systems of variable sizes.[43]

Once the direct problem is solved with a suitable model like the CPM, one can address the inverse problem with an optimization approach. For this purpose, we use GA, which are a well established tool in the class of natural problem solvers and which in contrast to other optimization methods (such as simulated annealing or descent methods) do not try to improve one solution, but rather to develop many solutions in parallel.[44–47] The main motivation to use GA to optimize the layout of MP is the lack of knowledge about the fitness landscape. It may be very rough with high barriers and its effects might depend on how the cells dynamically probe it. GA are also robust to noise, which is inevitable in biological experiments. For example, the exact seeding location of cells cannot be controlled and their movement has a strong stochastic component. With GA, it is not rigorously guaranteed to find the best solution in finite time, but a good working solution is sufficient in our context. In fact it has been used before to optimize patterns of motors in motility assays to transport microtubules efficiently[48] or guiding tracks for molecular shuttles.[49]

In the following we describe how we combine CPM and GA to identify optimal adhesive patterns for certain cell functions of interest. We pick three illustrative examples to demonstrate our approach. In the first (trivial) example we give a cell shape to the GA and let them predict the optimal pattern shape resulting in this cell shape. In the second example we ask for the optimal geometry which biases cell migration into one direction. For the third example we let the GA design a geometry that biases cell migration in one direction for low cell densities but reverses this direction for high densities. We envision that in the future, our approach can be used to strongly improve the power of MP for cell culture.

2 Methods

2.1 Genetic algorithms for optimizing adhesive geometries

Genetic Algorithm are inspired by natural evolution. A set of individuals, in our case the pattern layouts, compete in an selective environment. Only the fittest individuals are allowed to reproduce and to live on into the next generation. In the implementation of a GA the selection pressure is generated by a fitness measure, *e.g.* how well a given pattern layout leads to a desired shape. Those patterns which perform best are more likely to be selected for reproduction which can occur *e.g.* by averaging the shapes of the patterns. Reproduction results in a new generation which is again evaluated for its fitness and so forth until a desired fitness value is reached.

The principle of a GA is shown in Fig. 2A. The optimization by a GA starts out with a population consisting of several randomized individuals. In our case each individual is a pattern consisting of elementary building blocks called genes. *E.g.* the [Y] shaped patterns is made by three genes each encoding a straight line but with different orientations. How patterns are encoded is discussed in detail below. The algorithm usually starts out with a random configuration, but for better illustration the pattern in Fig. 2A are very regular.

To mimic evolution the fitness of each individual (=pattern) is evaluated. In the shown example the fitness is determined by how well a cell shape (red outline) matches the shape of a target cell (not shown). In the recombination step individuals are selected to cross their genomes and form offspring. The likelihood of being selected depends on the fitness of each individual. This mimics the selective pressure of the environment. For illustration the fitness based selection is not shown in Fig. 2A. Instead, two patterns generate two offspring without selection.

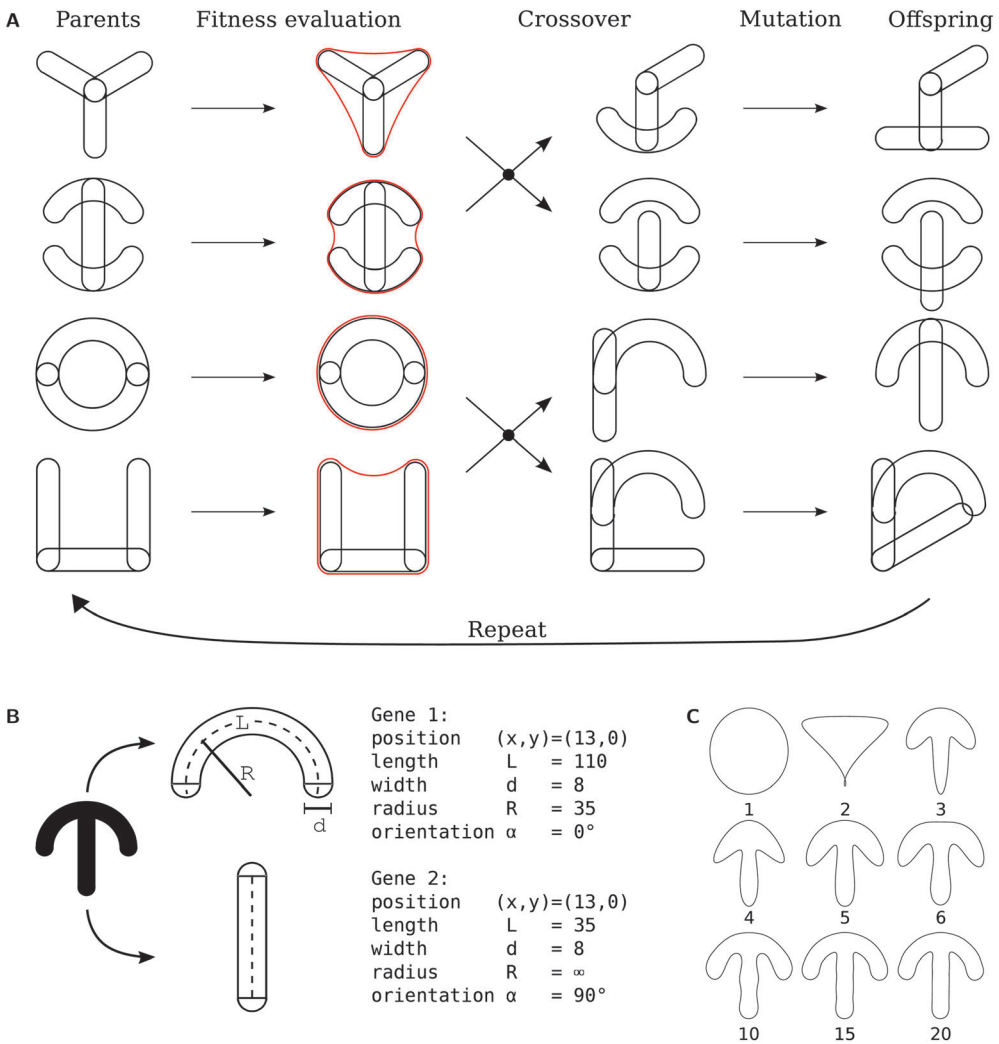


Fig. 2 Principles of genetic algorithms (GAs). (A) Pictorial representation of an iteration step in a GA. Members of a parent population are evaluated for their fitness, *e.g.* how well a cell on a member matches the shape of a target cell (cell shapes on the pattern are indicated by a red outline). According to their fitness they are selected for recombination in a crossover process. This generates a new generation which is subjected to mutation resulting in the offspring generation. The offspring become the new parents and the iteration starts again. (B) Representation of a [crossbow] pattern by arcs. The [crossbow] is made from two arcs, each arc is encoded in a gene by six numbers: the position (*x*,*y*), length *L*, width *d*, radius *R* and orientation α. Straight arcs have an infinite radius. (C) Representation of a [crossbow] pattern by Fourier descriptors. The order N_{max} of the Fourier expansion is indicated below the shapes.

The crossover of the genomes in the recombination step generates new pattern shapes which have attributes of both parents. In the last step mutation randomly changes the properties of some genes and one arrives at the final offspring generation. The offspring become the new parent and the iteration starts again. The algorithm is explained in detail in the ESI,† Section A.

The crossover and mutation operations are the main search operators of the algorithm which explore the fitness landscape while the fitness based parent selection ensures that good changes persist. The mutation step is important to maintain diversity in the population. We employ another mechanism called elitism to ensure that good individuals survive and their genome is not changed by mutation or recombination. During each iteration step 10% of the best individuals are carried over to the next generation without any changes to their genome replacing 10% of the worst individuals.

For a GA the pattern shapes need to be encoded in a genome. Many experiments have been performed with patterns made from arc like building blocks. The [crossbow] shown in Fig. 2B is made from two arcs, one curved part at the top and a straight at the bottom. Both arcs can be described by their position (x,y), length L, width d, radius R and orientation α as shown in Fig. 2B. Straight arcs have an radius of infinity. Thus, six numbers are sufficient to represent one arc. GA often use a binary representation for a genome. With each number represent by 16 bits the crossbow is described by a binary genome of 192 genes. In contrast, representing the pattern on a lattice similar to an image would result in a much larger genome (*e.g.* a 50 × 50 lattice representation amounts to a genome with 2500 binary genes).

GA are not limited to binary genomes and we take the six numbers describing one arc as a gene. As crossover operation we let parents exchange arcs and a mutation varies one of the six numbers describing an arc. A detailed description of the pattern representation as a genome, crossover and mutation operations can be found in the ESI,† Sections B and C.

An alternative approach to describe two-dimensional shapes is by Fourier descriptors (FD). They are a standard tool in image processing[50] and are often used to describe cell shapes obtained by optical microscopy.[51,52] For Cartesian FD a shape is represented in the complex plane by

$$z(s) = \sum_{\nu=-N_{\max}}^{N_{\max}} z_\nu \exp(2\pi i \nu s). \qquad (1)$$

Here, the closed curve $z(s)$ with parameter $s \in [0,1]$ is an expansion of the actual shape in a Fourier series truncated at $N_{\max}$. The complex valued coefficients $z_\nu = x_\nu + iy_\nu$ hold the information about the shape. To describe a pattern one can take the coefficients z_ν as genome. Each gene consists then of two real valued numbers representing the real and imaginary part of a coefficient. Crossover operations of the GA can exchange coefficient of two parents to form an offspring and mutation operations change the value of a coefficient.

As shown in Fig. 2C complex shapes can be represented by a small number of coefficients. FD superimpose ellipses to generate shapes and are therefore not suited to represent sharp kinks. However, this limitation is not relevant to describe MP or cells.

The initial population of a GA consists of individuals with random genomes. Examples are shown in Fig. S2A for an arc representation and in Fig. S2B for FD in the ESI.† With FD it is possible to enforce the symmetry of the patterns as described in the ESI.† Patterns with a reflection symmetry are shown in Fig. S2C and patterns with a three-fold rotational symmetry in Fig. S2D (ESI†).

2.2 Cellular Potts model for cell dynamics

The central part of a GA is the fitness functions which evaluates individual patterns and selects them for reproduction. The choice of the fitness function depends on the optimization target, *e.g.* the optimal pattern for a predefined cell shape or optimal pattern to bias cell migration. Thus, to calculate the fitness function a versatile model is needed that predicts dynamic cell behavior on MP correctly. We choose a two-dimensional cellular Potts model (CPM) to describe cells because of its ability to describe single and multicellular systems dynamically and because of its computational simplicity. The central part of a CPM model is an energy functional that we choose to be[42]

$$H = \sigma A + \lambda_s l + \sum_{\text{arc}i} \frac{k}{2L_{0,i}}\left(L_i - L_{0,i}\right)^2 - \frac{E_0}{A_{\text{ref}} + A_{\text{ad}}} A_{\text{ad}}. \qquad (2)$$

The first term is the contractile energy originating from the myosin activity in the actin cortex. A denotes the cell area projected onto the substrate. The surface tension σ controls the strength of the myosin contraction. The second term is a line energy. It acts in the contour of the cell where the plasma membrane folds back on itself. It scales with the contour length l of the cell and its strength is controlled by the simple line tension λ_s. The third term accounts for the reinforced contour observed at edges of cells extending above non-adhesive regions on concave MP.[9] Each edge has a length L_i and rest length $L_{0,i}$. The strength of the elastic contributions of these edge bundles is controlled by the elastic rigidity k. The last term drives spreading of a cell on MP. It decreases the energy when the adhesive area A_{ad} is increased. The ratio $W = E_0/A_{\text{ref}}$ sets the strength of the decrease and is equal to the adhesive area energy density. The constant A_{ref} is set through the typical cell size A_0. The gain from adhesive energy saturates with increasing adhesive area A_{ad}. This mirrors the finite number of adhesion receptors available to the cell. We propagate the cell shape by minimizing the energy functional eqn (2) with Metropolis dynamics. This model accurately predicts cell spreading dynamics, shape and traction forces on MP for standard cell types.[42]

In order to address also cell migration, cell division and cell–cell adhesion, the CPM for spreading of adherent cells has to be extended by a set of additional rules.[43] Cell migration is introduced through an effective migration machinery acting inside the cell. It decreases the surface tension at the cell front describing actin polymerization and increase it at the back. The movement direction is set through a velocity-alignment model.

Cells divide with a size dependent rate which decreases for denser tissue accounting for contact inhibition of proliferation.[53] Cell–cell interaction at adherence junction is described by a reduction of energy when new contacts between cells are formed. Likewise, the breakage of adherence junction requires energy. With these extensions, division and migration of single and multiple cells can be described in good quantitative agreement with experiments for standard cell types.[43] The parameters used in this work are for MCF10A cells as identified previously.[43]

3 Results

3.1 Optimizing single cell shape

As a first and simple illustrative example we use GAs to find the optimal pattern for a given cell shape. In Fig. 3A the averaged cell shape of a cell on a [crossbow] is shown which we use as target shape. Patterns are generated with the GA and the shape of a cell on a pattern is predicted by the CPM and compared to the target shape. The agreement with the target shape sets the fitness of a given pattern. Hence, a fitness function describing the agreement between the target shape and a simulated shape is needed. To measure the difference between shapes we use FD

$$D^2 = \sum_{\substack{\nu=-N_{\max} \\ \nu\neq 0}}^{N_{\max}} \left(|z_\nu| - \left|z_\nu^{(\text{target})}\right| \right)^2. \qquad (3)$$

The z_ν and $z_\nu^{(\text{target})}$ are the coefficients of the FD describing the simulated and target cell shape, respectively. The calculation of these coefficients from a given shape is defined in eqn (S8) in the ESI.† The zeroth order term is excluded to make the measure translation-invariant. Taking the absolute value of the coefficients renders the measure rotation-invariant since all orientation information is contained in the phase of the coefficients.[54] The series extends until $N_{\max} = 10$. Suppressing higher order coefficients neglects finer structures of the shapes such as fluctuations of the contour. For the fitness we choose $F = \exp(-D)$. A typical fitness trajectory is shown in Fig. 3B. After 30 generations the fitness becomes stationary. Throughout the whole population the fitness of the individuals is quite diverse. This diversity is ensured by a adaptive mutation strength defined in eqn (S14) in the ESI.† If the diversity of a population measured by the variance of the fitness decreases, the mutation rate is increased.

Fig. 3C shows optimal patterns encoded by arcs predicted by the GA. The GA optimization is performed three times with different initial conditions resulting in the three different shapes. All predicted patterns are very similar to the original one shown in Fig. 1B used to generate the target shape. However, since we use a contour model the pattern geometry inside of the cell has little influence on the shape of the cell. Therefore, the central bar is not always vertical or extended to the top. Pattern encoded by FD work equally well as shown in Fig. 3D. Again, only the contour of the patterns matches the pattern defining the cell shape and patterns can be asymmetric.

Our approach allows us to find the optimal pattern to a given cell shape in a way that is computationally very efficient.

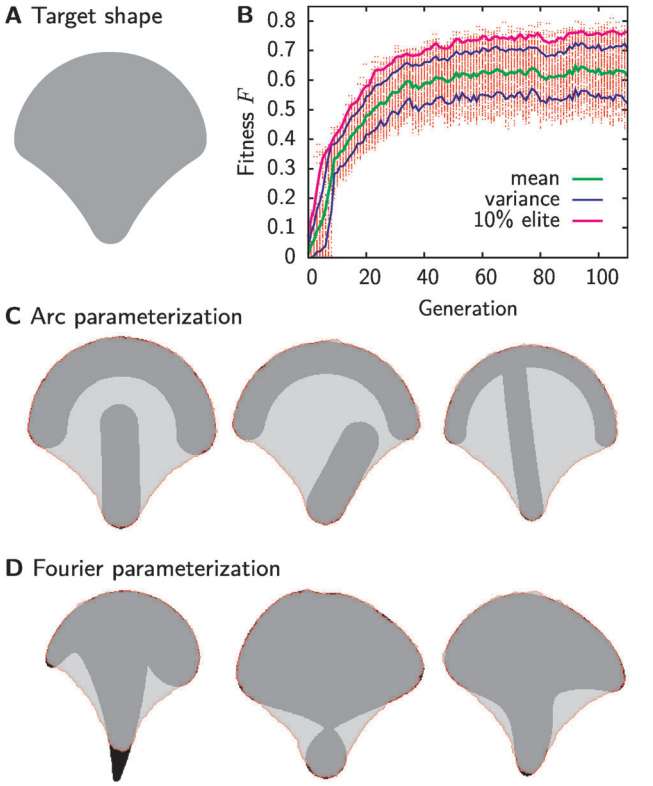


Fig. 3 Shape predictions. (A) Target shape of a cell on a [crossbow] pattern. (B) Fitness trajectory of the genetic algorithm. Red dots indicate the fitness of individual members (128 in total), the green line the mean fitness, the blue line the variance of the population and the magenta line the mean fitness of the fittest 10%. (C) Patterns encoded by two arcs predicted by three separate runs of the genetic algorithm. The cells on these patterns match the target shape of a cell on a [crossbow] pattern. (D) Same as previous but with cells encoded by FD with $N_{\max} = 5$ (compare eqn (1)).

This in turn allows us to easily assess the importance of different possible recombination and mutation operations. We found that one point crossover is the best recombination method. For mutation we found that adding random variables drawn from a Gaussian distribution is beneficial over complete randomization of a gene value. Adapting the width of the Gaussian distribution to the diversity in the population also yields better results. Crossover and mutation are described in detail in the ESI† by eqn (S12) and (S14).

3.2 Optimizing single cell migration

As a more demanding application of our approach, we now turn to optimal patterns for single cell migration and ask which ratchet patterns can bias migration into one desired direction. Ratchet-shaped structures as shown in Fig. 4A have been demonstrated experimentally to bias cell migration into one direction on triangular ratchets[32,33] and teardrop shapes.[28–31] However, it is not known if those shapes actually yield the optimal directional bias. We use the GA-approach to predict the optimal shape to bias cell migration towards one direction. To generate ratchet geometries, we arrange patterns into a square lattice as shown in Fig. S2E in the ESI† for patterns represented

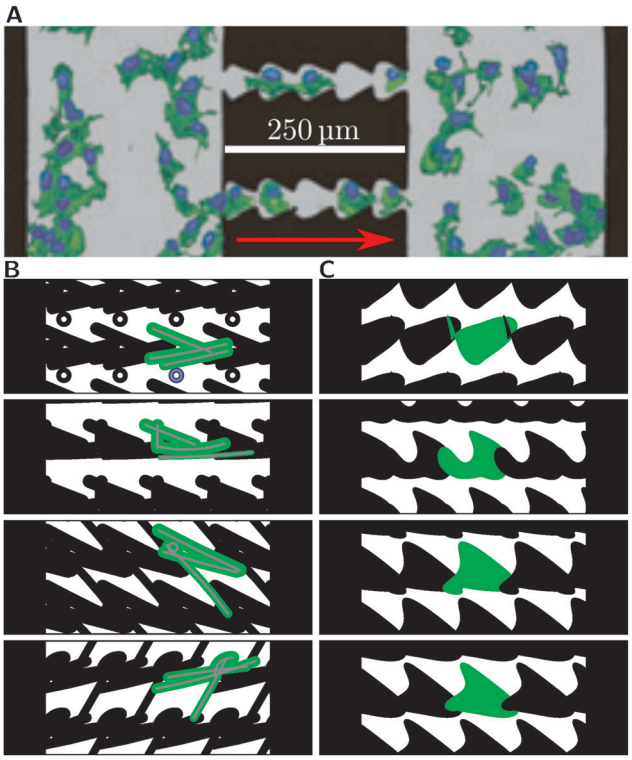


Fig. 4 Optimizing single cell migration on ratchet patterns with reservoirs. (A) Experimental observation of cell migration on a symmetric ratchet pattern with triangular shapes and cell reservoirs on both sides. Cells are stained for actin (green) and the nucleus (blue) and migrate preferentially to the right. Reprinted by permission from Macmillan Publishers Ltd: *Nat. Phys.*,[32] copyright (2009). (B) Patterns optimized by the GA to bias cell migration to the right. The four patterns originate from four independent runs of the GA. The patterns are generated from four arcs per elementary unit. The elementary unit is shown in color, has a width of 50 and is arranged into a cubic 4×2 lattice with periodic boundaries. Green and blue mark essential and non-essential parts for the observed cell behaviour, respectively. The patterns are connected to reservoirs and cells are only seeded in one of the reservoirs. (C) Same as previous but with patterns generated from FD.

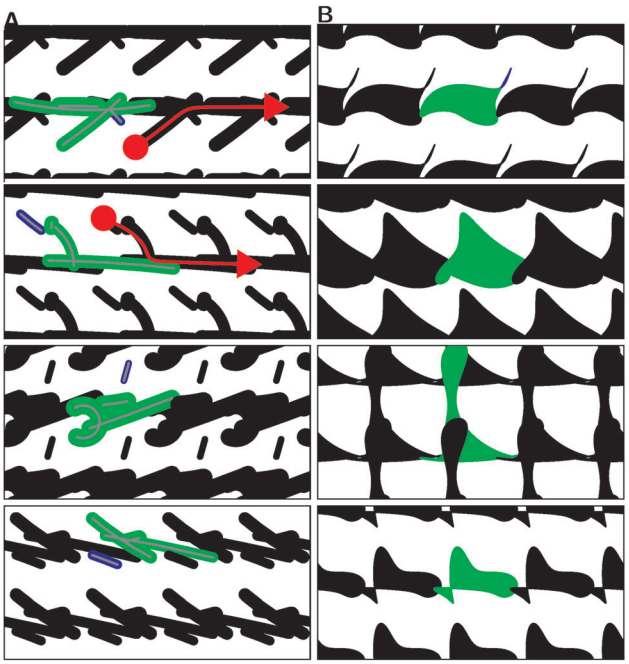


Fig. 5 Optimizing cell migration on ratchet patterns with random seeding. (A) Patterns optimized by the GA to bias cell migration to the right. The cells are seeded at random positions. The first two patterns show possible cell paths. (B) Same as previous but with patterns generated from FD.

by arcs and in Fig. S2F for pattern represented by FD. The first setup we investigate are ratchets connected to a reservoir on both sides with periodic boundaries. Cells are only seeded on the reservoir. As a second system we use no reservoirs and seed cells at random positions. As fitness measure we take

$$F = \frac{d_r}{d_0}, \tag{4}$$

where d_r is the distance migrated by a cell to the right during a fixed time (sufficient to cross the ratchet two times). It is normalized by the distance d_0 migrated on a homogeneously adhesive substrate when the migration direction is kept fixed. At any time there is only one cell on the ratchet but the seeding and migration measurement is repeated 32 times with random starting positions for each pattern and the results for d_r are averaged. If cells migrate to the left and the averaged fitness of a pattern becomes negative, we set it to zero.

Fig. 4 shows optimized geometries for patterns represented by arcs (Fig. 4B) and FD (Fig. 4C). The patterns are connected to reservoirs at the sides and cells are initially seeded only on the reservoirs. In Fig. 5A and B shows the same for patterns without reservoirs and cells seeded at random initial positions. In both cases, we mark one elementary unit in color; for the arc-case, in addition the arc-backbones are shown in grey. The green and blue colors differ between parts of the patterns which are essential for the observed cell dynamics (green) and which are not (blue). The identification of the non-essential (blue) parts results from a more detailed analysis (compare below the discussion of Fig. 6). These feature can in principle be removed in an experimental realization, but they are not suppressed by the GA. For both Fig. 4 and 5, one clearly sees how the whole structure is generated by placing the elementary units next to each other on a cubic lattice.

Typical fitness values (eqn (4)) for patterns connected to reservoirs are around 0.45. Cells spend time migrating vertically on the reservoirs which reduces the fitness. Without reservoirs and cells seeded at random positions the fitness is close or above one. A value above one means that cells migrate faster on the patterns than on a continuously adhesive substrate. This is an effect of the almost stripe like geometry of some patterns. An increased persistence for cells on stripes has been observed before[31] and was investigated with a velocity alignment model.[55]

Almost all geometries predicted by us have a triangular shape with the tip directed to the right. Such shapes have been experimentally demonstrated to bias cell migration in the direction of the tip, compare Fig. 4A.[32,33] However, in contrast to these earlier approaches, the triangles predicted here and shown in Fig. 4B and C are not arranged in a symmetric fashion. They are rotated and two corners lie on a horizontal

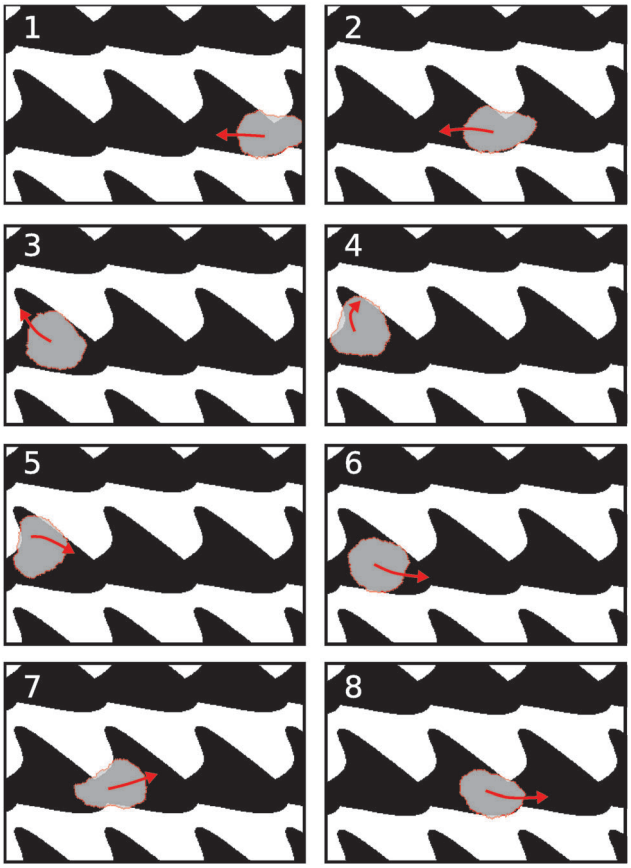


Fig. 6 Example of a cell migrating on ratchet MP predicted with a CPM. Initially the cell moves to the left. When it deviates from the horizontal track, it turns at one of the corners of a triangle, which thus acts like a dead end. Compare Movie S1 in the ESI.†

line forming an almost straight horizontal edge. In general, patterns generated by FD perform better to bias cell migration, most likely because they are better suited to generate triangular shapes than arc patterns.

When cells are seeded at random positions, asymmetric triangles are still the most prominent shape predicted. However, there are also triangles facing the other way, *e.g.* the first two patterns of Fig. 5A (when the appendices sticking out of the continuous horizontal track are ignored). Their angle is very acute which has been reported to bias cell migration away from the acute end.[29,56] When seeded randomly, cells move into the direction of more available adhesive area, which is for most random initial positions away from the acute tip. In addition, the narrow track prevents vertical migration. The appendices sticking out of the pattern catch randomly seeded cells and since they are dead ends, cells are biased to migrate to the right as indicated by the red arrows in Fig. 5A. For patterns with cells seeded in reservoirs no appendices are predicted.

The most significant difference of the patterns predicted here to previously used patterns in experiments is the asymmetric shape. The mechanism how asymmetric triangles bias migration is explained in Fig. 6, which shows snapshots of a cell migrating on an asymmetric triangular MP. The cell initially migrates to the left against the bias direction. When it deviates from the horizontal path it either encounters a vertical barrier or it moves into a dead end depending on the exact shape of the MP. In both cases the migration direction is usually reversed and the cell moves either along the long edge of the triangle or downwards along the vertical barrier. In both cases the downward movement is stopped at the horizontal continuous track. If the cell is not already biased to the right, *e.g.* when it moves vertically downwards, moving to the right provides more adhesive area and it is biased towards right. The cell migrating in Fig. 6 is also shown in Movie S1 in the ESI.†

Equilateral triangles are not predicted by the GA, most likely because they offer less of a dead end structure to turn cells. In addition, symmetric triangles have a continuous track in the center. When cells migrate in vertical direction they move past the continuous horizontal track. In fact the benefit of asymmetric triangles can also be quantified with the CPM. Arranging the triangles of Fig. 6 in a symmetric way reduces the fitness from 0.45 to 0.2, indicating that the asymmetry is indeed beneficial to bias cell migration (for a movie comparing the two types of triangles see Movies S1 and S2, ESI†).

There exists another mechanism to bias cell migration into one direction which uses the patterns shown in the top panel of Fig. 4C. Here, triangles are arranged in a way that cells have to migrate sidewards when moving to the next triangle. When cells encounter a pointed end of a triangle, in our CPM it is more likely that they form a sideway oriented protrusions (assuming that the leading edge points into the same direction as the triangle). Within our model forces generated by the migratory machinery decay towards the sides of the leading edge. The angle between center of mass of a cell and point on the membrane determines the strength of protrusions.[43] If the leading edge is forced to be narrow by the pattern the forces at the side are still large and the cell can easily form a protrusion sideways towards the next pattern. If the cell encounters the blunt end of a triangle, its leading front is already extended and protruding forces towards the side are low, making a transition less likely. This mechanism is also visualized in Movie S3 (ESI†) which shows a cell that initially migrates to the left and gets turned by the pattern. This mechanism of generating sideway protrusions at pointed ends of patterns has been observed experimentally for epithelial cells.[30]

3.3 Reversal of migration direction at high cell densities

We next ask if the GA can predict a geometry that biases migration of single cells in one direction but reverses this direction when the cell density is increased. To avoid correlations between cells generated by the periodic boundary conditions, we arrange the patterns on a circle. The fitness is constructed with the angular velocities by

$$F = -\frac{\omega_{\mathrm{low}}}{\omega_{0,\mathrm{low}}}\frac{\omega_{\mathrm{high}}}{\omega_{0,\mathrm{high}}}. \quad (5)$$

ω_{low} and ω_{high} denote the angular velocities for a low density of cells (up to four cells) and a high density (when the cell area is below $0.6A_0$ of the typical cell size A_0 of single cells).

Both velocities are normalized by the values of the angular velocities ($\omega_{0,\mathrm{low}}$ and $\omega_{0,\mathrm{high}}$) of cells migrating on a ring pattern with the cell migration direction prescribed externally. The minus ensures a positive fitness when the angular velocities have opposite signs. If the fitness is negative it is set to zero. Cells divide with a size dependent rate as described previously.[43] The measurement of the angular velocities is repeated 32 times with random initial positions for each pattern and the results are averaged to obtain the final fitness of a pattern.

Fig. 7 shows the results of the optimization. For low densities cells move clockwise and switch to a counterclockwise motion for high densities. The key element for the direction reversal are the appendices sticking out of the patterns. For low densities they are not occupied. The cells move in-line with one cell after the other occupying the bulk adhesive areas of the patterns. With increasing cell density cells sit next to each other. They occupy the appendices and form bridges above the nonadhesive areas. At that point the movement direction reverses (compare Movies S4 and S5, ESI†). The typical fitness of the pattern is between 0.5 and 0.7 indicating a very efficient bias mechanism. The interaction between cells is important for the reversal. Without any cell–cell adhesion the optimal fitness obtained by the GA drops below 0.3. Cells cannot bridge nonadhesive areas without cell–cell contacts which seems to be essential for reversed migration. Without contacts, cell migration stops for higher densities (compare Movie S6 in the ESI†).

4 Conclusions

We have used GA to rationally design adhesive MP for specific tasks. This can be considered to be an inverse problem for which one first has to solve the forward direction, that is one has to predict the behavior of cells on MP. Here we use a CPM which has been demonstrated before to predict cellular behavior on MP in good agreement with experiments.[42,43] To illustrate the algorithm we first used it to predict the optimal pattern for a given cell shape. For this task the solution is known and our predictions are very similar to the initial pattern used to generate the target shape (compare Fig. 3). Since we use a CPM which is a contour model, only the contour of the patterns is predicted correctly. Nevertheless, this task provides an effective benchmark for the GA.

As a second system we investigated the bias of single cell migration on ratchet MP in a linear arrangement. Earlier experiments focused on symmetric triangular[32,33] and teardrop patterns.[28–31] In contrast to these symmetric shapes our algorithm predicts asymmetric triangles as optimal shape to bias

Fig. 7 Optimal patterns predicted by the GA biasing cell migration into a clockwise motion for low cell densities and into an counterclockwise for high densities. (A) Pattern generated by a combination of two shapes both parametrized by FD. (B) Pattern generated from a single FD. The cell movement direction is indicated by arrows. The arrows are color-coded by the direction. The circular patterns are made from 16 unit cells, each unit cell has a dimension of 50. The outer diameter of the circle is 330. Both patterns are generated with FD with N_{max} = 5. For a visualization of the complete movement see Movies S4 and S5 in the ESI.†

single cell migration. A direct comparison of the symmetric and asymmetric patterns confirmed that the asymmetric ones perform better. Here, the full potential of our approach is revealed since it is not obvious that asymmetric shapes perform better. We use two different methods (arc and FD representation) to encode the pattern geometries into a genome. Both yield similar triangular shapes, confirming that the GA approach gives results independent of the details of the parametrization.

In the future, these predictions should be tested in experiments. When presenting our results, we chose not to remove small features of the obtained patterns that do not influence the cell behavior (marked in blue in Fig. 4 and 5). It is a general property of GA that features that do not interfere with the task at hand are not removed by the mutation and selection procedure. If these features are considered to be a nuisance for the experimental implementation, for example because feature size is too small for micropatterning, they could be removed. In general, their identification requires an analysis of cell behaviour with the direct model, as shown in 6. In the FD-method, they tend to be suppressed by using high order FD coefficients. Nevertheless here we also presented our results with the arc method, because this is the traditional approach in designing MP for cell adhesion.

We also note that one has to expect different experimental outcomes for different cell types. Here we have used a parametrization that has worked well before for the human mammary epithelial cell line MCF10A cells[42,43] and indeed we obtain results similar to the ones reported before for this cell line.[30,31] However, experimental results might be different for mesenchymal cell types such as 3T3-fibroblasts, for which our CPM might has to be adapted. Here we have focused on general aspects of linear ratchet patterns in order to demonstrate the potential of our method. Future applications, for example for single cell migration on square ratchets, for which the details of ligand accessibility become more important, might require more detailed models for the direct problem, for example regarding the growth and regulation of sites of adhesion.[57]

We note that our work does not only predict specific patterns to be tested in experiments, it also provides deeper insight into cell dynamics. In particular, our identification of non-symmetrical ratchet patterns as optimal solution for unidirectional cell migration demonstrates the importance of continuous tracks and the dead end structures leading to cell turning. In general, our analysis of cell dynamics on MP shows that the spatial combination of convex and concave regions is the most important element to determine the outcome, rather than small features that cell tends to move over. However, this conclusion might depend on our implementation of the CPM, which models high persistency of cell migration.

As a third task we let the algorithm design patterns that bias cell migration in one direction for low cell densities, but reverse this direction for high densities which has not been addressed experimentally before. The results are patterns which are only partly occupied by single cells. When the cell density increases, cells sit next to each other on the narrow tracks and completely occupy the patterns with bridges between nonadhesive regions (compare Fig. 7). The bridge formation seems to be very important for these patterns to work. Removing the adhesion between cells does not result in bridge formation and efficient reversal of the transport for high cell densities. The exact mechanism of direction reversal is not clear, but could be investigated in detail in experiments with the patterns suggested here.

For the future MP for virtually any task can be designed with our approach in a computationally efficient manner. Possible targets are mainly limited by the level of detail of the forward description. Here, the CPM proves to be very versatile, although other computational approaches like phase field models could also be employed in an optimization scheme for cell dynamics on MP.[6] The CPM has been employed ranging from the single cell scale with very detailed description of the migration machinery[40] to large multicellular systems. Further extension could include the formation of filopodia, which seem to be important in bridging the gaps between MP.[33] Cell division[43] and traction forces[42] are also accurately described by a CPM. The interplay between them has been investigated on continuously adhesive substrates experimentally.[58] It is well known that division axis, cell shape, stress and strain in a tissue are related.[59] MP could be designed to change these relationships. For example, it would be interesting to design MP geometries with GA which have the most probable division axis and prominent stress direction rotated by 90° to investigate their interaction further. Another very interesting extension for future work would be three-dimensional models, for example to investigate the role of the nucleus for cell migration on topographically structured ratchet patterns.[27]

References

1 G. M. Whitesides, E. Ostuni, S. Takayama, X. Jiang and D. E. Ingber, *Annu. Rev. Biomed. Eng.*, 2001, **3**, 335–373.
2 M. Théry, *J. Cell Sci.*, 2010, **123**, 4201–4213.
3 P. Kollmannsberger, C. M. Bidan, J. W. C. Dunlop and P. Fratzl, *Soft Matter*, 2011, **7**, 9549–9560.
4 U. S. Schwarz and S. A. Safran, *Rev. Mod. Phys.*, 2013, **85**, 1327–1381.
5 M. L. Zorn, A.-K. Marel, F. J. Segerer and J. O. Rädler, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2015, **1853**, 3143–3152.
6 P. J. Albert and U. S. Schwarz, *Cell Adhes. Migr.*, 2016, 1–13.
7 C. S. Chen, M. Mrksich, S. Huang, G. M. Whitesides and D. E. Ingber, *Science*, 1997, **276**, 1425–1428.
8 M. Théry, A. Pépin, E. Dressaire, Y. Chen and M. Bornens, *Cell Motil. Cytoskeleton*, 2006, **63**, 341–355.
9 I. B. Bischofs, F. Klein, D. Lehnert, M. Bastmeyer and U. S. Schwarz, *Biophys. J.*, 2008, **95**, 3488–3496.
10 C. Labouesse, A. B. Verkhovsky, J.-J. Meister, C. Gabella and B. Vianay, *Biophys. J.*, 2015, **108**, 2437–2447.
11 A. D. Rape, W.-H. Guo and Y.-L. Wang, *Biomaterials*, 2011, **32**, 2043–2051.
12 Q. Tseng, I. Wang, E. Duchemin-Pelletier, A. Azioune, N. Carpi, J. Gao, O. Filhol, M. Piel, M. Théry and M. Balland, *Lab Chip*, 2011, **11**, 2231–2240.
13 P. W. Oakes, S. Banerjee, M. C. Marchetti and M. L. Gardel, *Biophys. J.*, 2014, **107**, 825–833.

14 N. Hampe, T. Jonas, B. Wolters, N. Hersch, B. Hoffmann and R. Merkel, *Soft Matter*, 2014, **10**, 2431–2443.
15 K. Mandal, I. Wang, E. Vitiello, L. A. C. Orellana and M. Balland, *Nat. Commun.*, 2014, **5**, 5749.
16 S. R. K. Vedula, H. Hirata, M. H. Nai, A. Brugués, Y. Toyama, X. Trepat, C. T. Lim and B. Ladoux, *Nat. Mater.*, 2014, **13**, 87–96.
17 S. He, C. Liu, X. Li, S. Ma, B. Huo and B. Ji, *Biophys. J.*, 2015, **109**, 489–500.
18 M. Thery, V. Racine, A. Pepin, M. Piel, Y. Chen, J.-B. Sibarita and M. Bornens, *Nat. Cell Biol.*, 2005, **7**, 947–953.
19 J. Fink, N. Carpi, T. Betz, A. Betard, M. Chebah, A. Azioune, M. Bornens, C. Sykes, L. Fetler, D. Cuvelier and M. Piel, *Nat. Cell Biol.*, 2011, **13**, 771–778.
20 K. Doxzen, S. R. K. Vedula, M. C. Leong, H. Hirata, N. S. Gov, A. J. Kabla, B. Ladoux and C. T. Lim, *Integr. Biol.*, 2013, **5**, 1026–1035.
21 F. J. Segerer, F. Thüroff, A. Piera Alberola, E. Frey and J. O. Rädler, *Phys. Rev. Lett.*, 2015, **114**, 228102.
22 M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin and P. Silberzan, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 15988–15993.
23 S. Rausch, T. Das, J. R. Soiné, T. W. Hofmann, C. H. Boehm, U. S. Schwarz, H. Boehm and J. P. Spatz, *Biointerphases*, 2013, **8**, 1–11.
24 M. Théry, V. Racine, M. Piel, A. Pépin, A. Dimitrov, Y. Chen, J.-B. Sibarita and M. Bornens, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 19771–19776.
25 A. Brock, E. Chang, C.-C. Ho, P. LeDuc, X. Jiang, G. M. Whitesides and D. E. Ingber, *Langmuir*, 2003, **19**, 1611–1617.
26 X. Jiang, D. D. A. Bruzewicz, A. P. Wong, M. Piel and G. M. Whitesides, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 975–978.
27 D. Caballero, J. Comelles, M. Piel, R. Voituriez and D. Riveline, *Trends Cell Biol.*, 2016, **25**, 815–827.
28 G. Kumar, C.-C. Ho and C. C. Co, *Adv. Mater.*, 2007, **19**, 1084–1090.
29 G. Kumar, C. Co and C. Ho, *Langmuir*, 2011, 3803–3807.
30 K. Kushiro, S. Chang and A. R. Asthagiri, *Adv. Mater.*, 2010, **22**, 4516–4519.
31 K. Kushiro and A. R. A. Asthagiri, *Langmuir*, 2012, **28**, 4357–4362.
32 G. Mahmud, C. J. Campbell, K. J. M. Bishop, Y. A. Komarova, O. Chaga, S. Soh, S. Huda, K. Kandere-Grzybowska and B. A. Grzybowski, *Nat. Phys.*, 2009, **5**, 606–612.
33 D. Caballero, R. Voituriez and D. Riveline, *Biophys. J.*, 2015, **107**, 34–42.
34 F. Graner and J. A. Glazier, *Phys. Rev. Lett.*, 1992, **69**, 2013–2016.
35 J. A. Glazier and F. Graner, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1993, **47**, 2128.
36 *Single-Cell-Based Models in Biology and Medicine (Mathematics and Biosciences in Interaction)*, ed. A. Anderson and K. Rejniak, Birkhäuser, Basel, 1st edn, 2007, ch. II, pp. 77–168.
37 M. Scianna and L. Preiosi, *Multiscale Model. Simul.*, 2012, **10**, 342–382.
38 A. Szabó and R. M. H. Merks, *Front. Oncol.*, 2013, **3**, 87.
39 J. Käfer, T. Hayashi, A. F. M. Marée, R. W. Carthew and F. Graner, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 18549–18554.
40 A. F. M. Marée, V. a. Grieneisen and L. Edelstein-Keshet, *PLoS Comput. Biol.*, 2012, **8**, e1002402.
41 B. Vianay, J. Käfer, E. Planus, M. Block, F. Graner and H. Guillou, *Phys. Rev. Lett.*, 2010, **105**, 3–6.
42 P. J. Albert and U. S. Schwarz, *Biophys. J.*, 2014, **106**, 2340–2352.
43 P. J. Albert and U. S. Schwarz, *PLoS Comput. Biol.*, 2016, **12**, e1004863.
44 D. B. Fogel, *IEEE Trans. Neural Networks*, 1994, **5**, 3–14.
45 W. B. Langdon and R. Poli, *Foundations of genetic programming*, Springer, Berlin, Heidelberg [u.a.], 2002, pp. XV, 260 S.
46 A. E. Eiben and J. E. Smith, *Introduction to evolutionary computing*, Springer, Berlin, Heidelberg [u.a.], Corr. 2. p ed., 2007, pp. XV, 299 S.
47 M. Affenzeller, S. Winkler, S. Wagner and A. Beham, *Genetic algorithms and genetic programming*, CRC Press, Boca Raton [u.a.], 2009, pp. XXVII, 365 S.
48 B. Rupp and F. Nédélec, *Lab Chip*, 2012, **12**, 4903–4910.
49 T. Sunagawa, A. Tanahashi, M. E. Downs, H. Hess and T. Nitta, *Lab Chip*, 2013, **13**, 2827–2833.
50 B. Jähne, *Digital image processing*, Springer, Berlin, Heidelberg [u.a.], 5th edn, 2002, pp. XII, 585 S.
51 *Fourier descriptors and their applications in biology*, ed. P. E. H. Lestrel, Cambridge University Press, Cambridge [u.a.], 1997, pp. XI, 466 S.
52 Z. Pincus and J. A. Theriot, *J. Microsc.*, 2007, **227**, 140–156.
53 A. Puliafito, L. Hufnagel, P. Neveu, S. Streichan, A. Sigal, D. K. Fygenson and B. I. Shraiman, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 739–744.
54 D. Zhang and G. Lu, Proc. of 5th Asian Conference on Computer Vision (ACCV), 2002, pp. 646–651.
55 B. A. Camley and W.-J. Rappel, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2014, **89**, 62705.
56 T. Vignaud, L. Blanchoin and M. Théry, *Trends Cell Biol.*, 2012, **22**, 671–682.
57 P. P. Provenzano and P. J. Keely, *J. Cell Sci.*, 2011, **124**, 1195–1205.
58 H. Tanimoto and M. Sano, *Phys. Rev. Lett.*, 2012, **109**, 248110.
59 T. P. J. Wyatt, A. R. Harris, M. Lam, Q. Cheng, J. Bellis, A. Dimitracopoulos, A. J. Kabla, G. T. Charras and B. Baum, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 5726–5731.

