THE EPITHELIOME: AGENT-BASED MODELLING OF THE SOCIAL BEHAVIOUR OF CELLS

D C Walker[1], J Southgate[2], G Hill[2], M Holcombe[3], D R Hose[1], S M Wood[1], S Mac Neil[4], R H Smallwood[1*]

[1] Medical Physics and Engineering, University of Sheffield, Floor I, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF, United Kingdom.
[2] Department of Biology, University of York
[3] Department of Computer Science, University of Sheffield
[4] Department of Engineering Materials, University of Sheffield
* corresponding author: r.smallwood@shef.ac.uk

KEYWORDS

Computational modelling, software agent, epithelium, urothelium, cell culture, calcium

ABSTRACT

We have developed a new computational modelling paradigm for predicting the emergent behaviour resulting from the interaction of cells in epithelial tissue. As proof-of-concept, an agent-based model, in which there is a one-to-one correspondence between biological cells and software agents, has been coupled to a simple physical model. Behaviour of the computational model is compared with the growth characteristics of epithelial cells in monolayer culture, using growth media with low and physiological calcium concentrations. Results show a qualitative fit between the growth characteristics produced by the simulation and the *in vitro* cell models.

1. INTRODUCTION

1.1 Modelling Epithelial Systems

The components of biological tissue span orders of magnitude in size from sub-micron genes and molecules, to entire organs, tens of centimetres in diameter. Similarly, the time scale of processes within the tissue can vary from chemical reactions that occur in milliseconds, to the slow processes of tissue morphogenesis and pathogenesis which take place over many months. Such complexity, coupled with the enormous volume of data provided by the sequencing of the human genome, renders the task of understanding biological processes almost insurmountable. Computational models provide us with an invaluable tool with which to further our understanding of biological systems, and handle the immense amounts of data available.

In comparison with organs such as the heart (Noble 2002), modelling of epithelial tissue is in its infancy. Examples of work relating to epithelia include the electrical models developed by Walker *et*

al. for cervical squamous epithelium (Walker D C 2002) and Jones *et al.* for gastric and oesophageal columnar epithelia (Jones *et al*., 2003). In common with the vast majority of computational biological models, these are based on an assumed, fixed tissue structure, and are used to predict physical characteristics, in this case, electrical impedance across a frequency range. In reality, however, there is no fixed 'scaffold' that determines the structure achieved during tissue morphogenesis. The organisation into complex tissues and organs is an emergent property of the constituent cells of an organism, and the genes that determine the behaviour of those cells. It is this concept of structure as an emergent property of the interaction or 'social behaviour' of a large number (10^6-10^7) of individual cells that we wish our model to address explicitly.

Although vastly outnumbered by predetermined structural models, the concept of growth, division and differentiation of individual cells has been addressed by a number of workers (see Smallwood *et al*. , 2003, for a more detailed review). In terms of epithelial tissue, the models developed by Stekel and Rashbass (Stekel D 1995) and Morel *et al*. (Morel D 2001) are particularly interesting. However, although the first and possibly most comprehensive attempt to date to simulate epithelial development, the Stekel model has the drawback of being essentially empirical in nature. Rules are selected to produce the desired outcome, and are not linked to the underlying biological mechanisms.

Morel *et al (Morel D 2001)* took their epithelial model a step further by incorporating two distinct hierarchies: a kinematic model of cell cycle regulation, incorporating both intracellular components (e.g. cyclins) and response to extracellular stimuli (e.g. growth factors); and a 2-D Voronoi graph-based tissue architecture model, with individual cells represented by polygons. Although potentially a powerful model, to the authors' knowledge, only simulations of basal cell dynamic behaviour have been carried out to date.

Whereas the models described above have either been purely illustrative in nature, or have focussed on a limited range of cell behaviours, our ultimate aim is to produce a full 3-D model of normal functional epithelium that identifies the critical processes (proliferation, differentiation, cell-cell and cell-substrate interactions) and the underlying inter- and intra-cellular regulatory signalling mechanisms, relating ultimately to changes in gene expression (i.e. the *transcriptome*). This 'normal' model can then be used as a basis to examine the anomalies in these processes that eventually lead to pathological conditions, including proliferative disorders, metaplasia and neoplasia. Thus, we will consider the epithelium as a complete biological system, linking '*omics'* from genotype through to the normal and pathological phenotype.

We plan to make this model compatible with other packages being developed under the umbrella of the Human Physiome project (Bassingthwaighte 1995, (Bassingthwaighte 2000)) by creating an interface using an application-specific mark-up language similar to the existing Cell ML (Hedley *et al.* 2001).

1.2 Rule-based modelling

In undertaking such an extensive project, the selection of the most appropriate modelling paradigm is critical. Likewise, choosing a suitable starting point to model of a system that spans several orders of magnitude is also crucial. Naively, it may seem that the genome is the most obvious starting point, as it is the information encoded in the genes that ultimately controls the cellular behaviour, and hence the final tissue architecture and function. However, although the sequencing of the human genome may be essentially complete (Venter *et al.* 2001) the task of unravelling the often context-dependent coding of genes for particular proteins (proteomics), remains an immense undertaking that may take many years to complete. For this reason, although the roles of particular genes in epithelial tissues are already known (Eckhert *et al*. 1997), we are unlikely to be able to elucidate the complete set of genes essential for epithelial development and function, or the subset of genes linked to specific pathological changes, for some time to come. Nevertheless, the variety of behaviours of individual epithelial cells critical to tissue formation, e.g. growth, division, migration, bonding, signalling and differentiation are fairly well categorised, although in many cases the molecular or genetic mechanisms that underlie them remain elusive. For this reason, the individual epithelial cell, and a rule-based modelling paradigm, are the ideal starting points for the development of a model. As a first approximation, simple rules sets can be assigned to cells depending on their differentiation status (e.g. stem cell, transit amplifying, committed), to be executed depending on the internal state, and external environment of that cell. This rule-based model structure is similar to that adopted by Stekel *et al.* for their epithelial model (Stekel D 1995), but it is our intention to extend and elaborate upon this basic rule-based framework to include increasingly detailed hierarchies of control. For example, intracellular signalling pathways involved in cell-cycle control will be included, as will the expression of genes responsible for regulating these pathways. Thus, a rule-based model provides us with a convenient framework into which we can slot modules representing the actual controlling mechanisms behind these rules.

When models are implemented in such a way that rules are applied to a population of individual interacting entities, or agents, this concept has often been termed ‘agent based modelling’. This paradigm has been successfully employed to model colonies of ants, bees (Gheorghe M 2001) and even the human immune system (Holcombe M 1998), demonstrating that complex social systems have evolved where there is no concept of a higher control, but each individual component of the system, or agent, executes a set of rules in response to various environmental cues. It is the method of representing biological cells as autonomous software agents, in combination with a simple physics-based model, that we have exploited to form the basis of the initial stages of our model development.

At this stage, all software development has been carried out using object-oriented code in Mathworks Matlab versions 5.3 and 6 and all models are run on a standard Pentium 4 PC.

2. METHODS

2.1 Initial conditions and basic model structure

Model set-up is carried out using a menu-based graphical user interface. The model consists of a number of cells interacting with their environment, which comprises a two-dimensional, square substrate with user-defined dimensions and modifiable exogenous calcium ion concentration. The user can select the number of cells seeded and whether to place them randomly on the substrate, or in specified locations. The seed cell radii and epithelial type (either keratinocyte or urothelial) can also be selected by the user. All seeded cells are designated to be members of the stem or transit amplifying (TA) cell classes.

Model execution is based on an iterative process, with each tick or agent model iteration representing approximately 30 minutes in real time. This was achieved by selecting mean cell cycle duration and migration rates, so that the cell cycle step, or distance migrated by cells at each time step is, on average, similar to that undergone by a real cell during this time period. These data were based on a mean cell cycle duration of 60 hours for keratinocytes (Dover R 1983) and 15 hours for urothelial cells (Southgate J 1994) and a migration speed of approximately 1μm/ minute, as measured from time lapse sequences of urothelial cell culture. Obviously, these parameters would be expected to vary as a result of culture conditions as well as cell type, and it is an important future aim of the *in vitro* work to be carried out in conjunction with the computational modelling, to elucidate the difference between genetically pre-programmed as opposed to environmentally dependent parameters. However, at this early stage in model development, relevant values were selected to lie within the expected range.

At each tick of the model clock, each cell is interrogated in turn, and depending on the state of its internal parameters (e.g. current position in cell cycle, flag indicating bond to substrate) and its environment (number and proximity of neighbouring cells, concentration of calcium ions) a number of rules are executed that may change the state of the cell's internal parameters. During this process, cells can receive messages from other cells in their immediate vicinity, or send messages to other cells by reading from and writing to a structure designated as a communication matrix. An additional global data structure, which can be accessed by all cells in the model, contains information relating to the exact position and dimensions of every cell. Thus, each cell knows the location (distance and azimuth) of its close neighbours within approximately 10 cell radii, and can moderate its behaviour accordingly. This information is updated every time a cell moves or changes shape. If a cell divides during a model iteration, one daughter cell overwrites the data structure of the parent, whilst the second daughter is added to end of the string of existing cell structures. Further details of rule sets pertaining to particular cell behaviours are given below.

After the rule sets have been applied to every cell in the model in turn, a numerical technique is then used to correct for any overlap between neighbouring cells that may have resulted from cell growth, spreading or division during the previous agent based iteration. This involves identifying any overlap

between every pair of neighbouring cells in the model, and applying a force proportional to this overlap to separate these cells. Further details of the formulation of this physical model are given in section 2.5.

2.2 Modelled cell types

Each cell in our model is represented as a Matlab data structure assigned to a particular object class according to its designated cell type. The current permitted types are:

1. **Stem cells** (specifically epithelial stem cells). These are predicted to exist in many types of epithelial tissue, although a definitive biochemical/antigenic marker has yet to be found. The fundamental property of such cells is that they have an unlimited capacity for self-renewal, and thus are critical in maintaining a constant cell number in normal homeostatic tissue. There is some evidence that in certain types of epithelium, stem cells may have a particular spatial distribution, may cycle more slowly or may express different integrins and thus exhibit distinct contact properties to transit amplifying cells (Jensen U B 1999), (Potten C S 2002). We have currently ignored such characteristics, some of which remain speculative, and restrict our stem cells to differ from TA cells in terms of their self-renewal properties only. Ten percent of the initial seed cells are designated to be stem cells (Potten C S 1988). The permitted behaviour set for this cell type are:
    - Bonding (to substrate and other cells)
    - Spreading (once bonded to substrate)
    - Lateral migration (once bonded to substrate)
    - Apoptosis (if failure to bond to substrate)
    - Cell growth and division
2. **Transit Amplifying, or TA cells**. These form approximately 90% of the basal compartment of epithelial tissue. Unlike stem cells, they are capable of only a limited number of division cycles before becoming committed to terminal differentiation. The number of rounds of division allowed appears to vary between different epithelial tissue types. For instance, it is not unusual for keratinocytes in monolayer culture to express cytokeratin isotypes associated with differentiation, although this is not observed in the culture of urothelial cells (Southgate J 1994). It is, of course, quite possible that this disparity can be accounted for by the different conditions used to culture these cell types, in contrast to being hard-coded for each cell type at a genetic level. However, whilst this point remains in doubt, the number of allowed TA cell divisions allowed are limited to 3 for keratinocyte models and 30 for urothelial cell models. The permitted behaviour set for TA cells is identical to stem cells in all other respects.
3. **Mitotic cells** These are a subset of both stem and TA cells that are progressing through the M phase of the cell cycle. In cell culture, mitotic cells are observed to round-up thus losing adhesion with their neighbours, and to some extent with the underlying substrate. For this reason, model cells assume a rounded morphology, breaking all intercellular bonds on entering the M phase

(defined as the final 20% of the cell cycle) and are not permitted to form new bonds or actively migrate until mitosis is complete.

4. **Post-mitotic cells** These are produced by the division of TA cells that have undergone the maximum permitted number of mitotic cycles. As explained above, they are rarely observed in monolayer culture of urothelial cells, but are relatively common in keratinocyte culture. The extension of our model to be truly three dimensional will allow these cells to assume the characteristics of cells committed to terminal differentiation, and migrate upwards to reside in the supra-basal layer. Currently, post-mitotic cells are permitted access to the same rule sets as stem and TA cells, with the obvious exception of cell growth and division.
5. **Dead cells** Failure of normal cells in culture to adhere to the substrate results in the programmed cell death (apoptosis) of that cell. The dead cells are then fragmented and phagocytosed by their neighbours. Model cells that have failed to bond to the substrate before the end of their G1 phase become dead cells, shrink and are removed from the model. In reality, most cells adhere very quickly after seeding, and hence dead cells are a rare occurrence.

In addition to the internal parameters unique to the designated class of a cell, every cell structure in the model contains a set of parameters that defines the exact location of its centre in x, y and z Cartesian co-ordinates, and the length of its axes in each direction. Hence, although the model remains essentially two dimensional at this stage, each cell is defined as a three dimensional ellipsoid. This allows three dimensional changes in cell morphology, such as cell spreading and rounding, to take place.

2.3 Rules dictating cell behaviour

2.3.1 Cell cycle

As discussed above, at this relatively early stage of development, our model is restricted to modelling cellular-based behaviours only – the sub-cellular mechanisms underlying these processes are essentially ignored. The central hinge of our model is the cell cycle, a schematic of which is shown in Figure 1. All cells capable of cycling (i.e. stem and TA cells) have an internal clock, and provided that they are not inhibited, progress one tick through the cycle at each model iteration. At model initiation, or when mitosis occurs to create new daughter cells, the new TA cells are designated a fixed S-G2-M phase length comprising approximately 50% of the total cell cycle duration according to the model cell type (keratinocytes = 60 hours or 120 model iterations, urothelial cells = 15 hours or 30 iterations) and a G1 phase length that is selected randomly from a normal distribution with mean equal to half the total cycle length and standard deviation equal to a tenth of the mean. This reflects the situation that is observed in real cells where the variation in the cell cycle duration between cells of the same lineage is derived almost entirely from variation in G1 duration.

The G1 phase of the cell cycle is the growth phase during which cells effectively double their volume by synthesis of proteins within the cytoplasm. At the first tick of the G1 phase of a model cell, the cell's volume is calculated along with the volume increase required at each subsequent tick in order

that the volume is exactly doubled at the end of G1. This volume increment is added to the cell at each G1 tick.

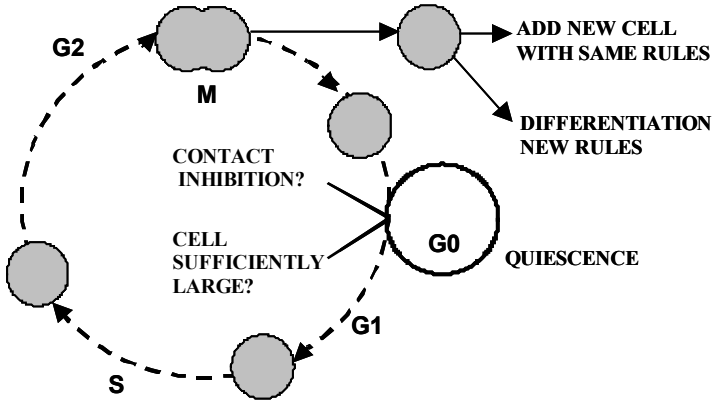

Figure 1 Model of simplified cell cycle. Segments of dotted line indicate incremental 'ticks' in cycle

Approximately half way through G1, a checkpoint has been introduced that, at present, governs whether a cell proceeds through the remainder of the cell cycle based on the number of contacts the cell has formed with neighbouring cells, and its morphology. It has been observed that in the culture of several epithelial cell types, cells situated at the centre of cell colonies grown in physiological or high calcium media withdraw from the cell cycle and become quiescent. It is believed that this phenomenon, termed contact inhibition is mediated via intercellular signalling – cells are aware of the existence of their nearest neighbours via the cadherin-mediated, and hence calcium-dependent physical bonds (St Croix B 1998), (Orford K 1999). The requirement for a model cell to become contact-inhibited is that it has four or more bonds with neighbouring cells (or, alternatively three bonds and be adjacent to the 'edge' of the substrate). It has also been observed that cells require a spread morphology in order to progress further in the cell cycle. For this reason, model cells also check that they are bonded to the substrate and sufficiently spread at this checkpoint. Cells which have not attached at this stage undergo apoptosis.

Cells which do not fulfil the criteria to pass through the G1 checkpoint enter a quiescent, or G0 phase of the cell cycle. The cell will remain in this state until enough of its intercellular bonds are broken to allow it to re-enter the cycle (this may occur if a neighbouring bonded cell reaches the mitotic phase of it cycle, rounds up, and breaks its intercellular bonds), and the cell has spread sufficiently (Section 2.3.3). A corollary of these rules is that they allow us to introduce the concept of wound repair in a confluent monolayer model.

If a cell passes the checkpoint, or fulfils the requirements to leave the G0 state, it will then progress to the S and G2 phases. Biological cells synthesise new DNA during the S phase, and prepare for mitosis during the second lag phase, or G2. However, as our model does not currently consider sub-cellular phenomena explicitly, during these phases model cells simply increment their cycle clock by a single tick at each model iteration. Further checks relating to phenomena such as occupation of growth factor receptors (G1 phase) and successful DNA replication (G2 phase) will be incorporated at a later stage of model development.

On approaching the final 20% of the cycle, defined to constitute the mitotic or M phase, the cell alters its morphology to become rounded, breaks any bonds it may have with neighbouring cells and is redefined as a member of the mitotic cell class. Cells in this class are defined as two distinct but overlapping spheres, which move apart over the duration of the M phase to produce two adjacent volumes that are eventually designated as two separate daughter cells. The first daughter cell structure overwrites the position of the parent in the model, and the second is added to the end of the existing list of cells. Stem cells always undergo symmetric division, i.e. produce one daughter that is also a stem cell, and a second that is a transit amplifying cell, whereas transit amplifying cells always produce transit amplifying daughters.

2.3.2 Bonding

2.3.2.1 Bonding to substrate

New cells, whether seeded onto the substrate, or the product of mitosis, have a rounded morphology and are not bonded to the substrate – i.e. they are not polarised. At each subsequent model iteration, there is a probability-based calculation of the cell binding to the substrate. This calculation is biased to reflect the fact that it is extremely unusual for cells in culture not to adhere within the first hour or so after seeding. Once adherent, a cell is then permitted to engage in a number of cell behaviours, including spreading and migration.

2.3.2.2 Calcium-dependent intercellular bonding

As stated above, the location of all cells in the model is continually updated and stored in a global data structure, which allows every cell to know the relative positions of all its neighbours. For each cell, at every model iteration, a binding probability is calculated for each neighbour within a set distance (defined to be half the specified radius of the original seeded cells, which in the case of the models discussed in this paper is 10μm). The probability of bonding for a pair of cells is inversely proportional to the separation between the cell edges and related to the environmental calcium concentration via a sigmoid function with the inflection point at 1.0mM. This relationship is based on published data (Baumgarter W 2000 on the relationship between strength of E-cadherin-mediated bonds and calcium ion concentration. We have assumed that the bond strength can be correlated with bond number, hence bonds are common at calcium concentrations greater than 1.0mM and rare for lower concentrations. A random number between 0 and 1 is then automatically generated and bonds will be formed with cells whose probability constant is greater than this number. All cells that form intercellular bonds automatically become bound to the substrate, if they are not already attached.

The first cell records the identity flags of the successfully bonded neighbouring cells within its data structure and sends a signal to these neighbours via the global communication matrix. On receiving the signal, a cell adds the flag of the sender to its own data structure. Hence every cell in the model that is able to form bonds will contain a list of flags indicating the cells to which it is linked. Bonds between cells will be recognised when the data is assembled for the next physical correction and non-adjacent

linked cells will feel an attractive force proportional to their separation, resulting in them moving together until their edges are adjacent.

2.3.3 Spreading

Cells seeded onto the substrate and new daughter cells produced by mitosis are not attached to the substrate and are spherical in shape. In order to be able to progress through the cell cycle, real biological cells spread by forming numerous attachments to the substrate, thus assuming a flattened morphology (Huang S 1999). This process is emulated in the model by increasing the surface area of attached cells by a small amount each model iteration, whilst maintaining a constant volume. Cells continue to spread until their radius has increased by a factor of approximately 1.5.

2.3.4 Migration

Cells that are attached to the substrate but have no intercellular bonds are free to migrate laterally on the substrate surface. The allowed migration distance per model iteration is fixed at approximately 2.5 times its radius. This figure was obtained from measurement of urothelial cell movement from successive frames of a time-lapse recording of a wound closure assay. Migration directions are randomly assigned to new cells. At each subsequent iteration, a cell will tend to carry on migrating in the same direction, but may randomly choose another direction within 60° of its original trajectory. If the migration path is blocked by a neighbouring cell or the edge of the substrate, the cell will move by a smaller distance so that its edge is adjacent to that of its neighbour. If no movement at all is possible in the chosen direction, the cell will alter its trajectory by 30 degrees and try again, continuing to do so up to 10 times, after which it will remain stationary. Such checks, and similar ones that are executed when selecting the direction of the separation in mitotic cells, for instance, are essential to ensure that cells do not fall off the edge of the model, or overlap one another by more than a single cell radius, as this could lead to errors during the physical correction.

2.3.5 Apoptosis

Stem and TA cells that have failed to attach to the substrate by the end of their G1 phase are redefined as dead cells. These cells shrink, and are removed from the model.

2.4 Communication between cells

Individual cells within the model can communicate via a global Communication Matrix, which is a data structure N cells long by N cells wide. In practice, cells only communicate with their nearest neighbours, so this matrix is sparse. Currently, the communication matrix is used to transmit information about intercellular bonding. If cell m forms a bond with cell n, a *'stick'* message is posted in row m, column n of the communication matrix, and cell n will receive the message and update its internal parameters accordingly. If, however, later in the model cell m reaches it M phase and wishes to break all intercellular bonds, a *'split'* message is posted in the *mth* row, and columns corresponding to all bound neighbours. Once a message is received it is removed from the matrix.

2.5 Physical Correction

Cell growth, movement and mitosis events that occur during each agent-based iteration will invariably result in some degree of overlap between neighbouring cells. In addition, there will often be small separations between the edges of neighbouring cells that are bonded to one another. In order to correct for these anomalies, each agent-based model iteration is followed by a numerically-based physical correction. For the purposes of brevity, only an overview of the method is given in this paper.

For each cell in turn, the location and size are extracted from its internal parameters, and along with the global position information, used to identify any overlap that may exist between it and any neighbours. Any gaps between the cell and bonded neighbours are also identified. Only cells that are overlapped or bonded are designated to be linked, hence the computation required for the correction is kept to a minimum. The forces existing between each pair of linked cells is set to be proportional to their overlap (or separation), i.e:

$$F_{ij} \propto sep_{ij} - \sum r_{ij} \qquad (2.1)$$

where F_{ij} is the force between cells i and j, sep_{ij} is the separation between centres and r_i and r_j are the radii. Thus, each pair of overlapping cells will experience a negative, or repulsive force, which tends to move them apart, whilst cells that are bonded, but not touching, will feel a positive attractive force.

The relevant position and separation information is obtained for each linked pair of cells and used to populate a number of relevant matrices and vectors with dimensions dictated by the number of linked cells in this model. Matrices and vectors relating to the x and y directions are constructed separately. An iterative, numerical method is then used to obtain the equilibrium positions of all the cells in x and y independently. Once a solution has been found the new position information is inserted into the individual cell structures, and the next iteration of the agent model commences. Hence this physical correction is an iterative process contained within the main iterative structure of the primary agent model.

In high density culture conditions, it may be impossible to find a solution where there is no overlap between adjacent cells. In order to circumvent this problem, an error value, corresponding to the sum of any overlaps with adjacent cells is recorded for every cell at every physical iteration. This list of values is termed the *error vector*. When the difference in the norm of this error vector between two consecutive iterations falls below five percent, the physical correction is terminated and the current solution is returned. Any remaining overlap is interpreted as squashing of cells, and during the next agent based iteration, cells with less than 3 bonds will round up, reducing their radius to accommodate any overlap. Thus at high cell density, in low calcium cultures in particular, cells will adopt an increasingly round morphology and will eventually be prevented from passing the mid G1 checkpoint.

3 COMPARISON WITH *IN VITRO* MODELS

We are in the unusual situation where we have access to *in vitro* cell culture models that are direct parallels to our computer-based simulations. This provides us with the valuable opportunity to test the ability of our model to accurately simulate the growth characteristics of real cell cultures. For the purpose of this report, we have carried out a descriptive comparison with a well-characterised normal human epithelial cell culture system that has been developed for bladder epithelial (urothelial) cells (Southgate, Hutton et al. 1994; Southgate, Masters et al. 2002).

It is generally accepted that the growth rate of epithelial cells in culture may be modulated by the concentration of calcium ions in the growth medium. In our in vitro system, normal human urothelial (NHU) cells are grown as monolayer cultures in a low calcium medium (0.09mM), whereas physiological calcium is used in order to induce stratification (Southgate et al., 1994).The effect of a physiological calcium environment on the proliferation of urothelial cells is not known. This provides us with an interesting scenario with which to test our computational model against an *in vitro* experiment.

3.1 Computational Model Simulations

Computer simulations were carried out using extracellular calcium environments of 0.09mM and 2.0mM. The model substrate size was set at 1.2mm by 1.2mm and the number of cells seeded at the start of the simulation was 100, giving an initial seeding density of approximately $7x10^3$ cells per cm^2. The cell cycle times in the model were designated to represent urothelial cells (i.e. mean cycle duration of 15 hours, or 30 model iterations). The simulations were run for a total of 160 iterations, equivalent to 80 hours in real time. The total cell number was recorded at every model iteration for both the low and high calcium simulations for later comparison with the *in vitro* assays. In addition, an image file of the substrate and cells was also saved in a *jpeg* format to allow visualisation of the simulated monolayer growth.

3.2 Culture of normal human urothelium

NHU cell lines were established from surgical specimens, as detailed previously (Southgate, Hutton et al. 1994; Southgate, Masters et al. 2002). Briefly, freshly-isolated normal human urothelial cells were suspended in low calcium growth medium (Keratinocyte Serum-Free Medium containing 50µg/ml bovine pituitary extract, 5ng/ml epidermal growth factor and 30ng/ml cholera toxin), seeded into tissue culture flasks at an initial minimum seeding density of 2.6×10^3 cells per cm^2 and maintained at 37°C in a humidified atmosphere of 5% CO_2 in air. As cultures reached confluence, the NHU cells were harvested from the substrate using trypsin and EDTA, and seeded into six fresh flasks at a split ratio of between 1/3 to 1/6. NHU cell lines were used for experiments between passages 1 and 4. For some experiments, the calcium concentration of the medium was increased from 0.09mM to 2mM by the addition of $CaCl_2$. Changes in cell population number were estimated by MTT assay, which is

based on the ability of metabolically-active cells to convert a yellow formazan salt (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, MTT) to insoluble blue crystals. Cells in 6 replicate cultures were incubated with 0.5mg/ml MTT at 37°C for 4 hours, the formazan crystals were collected, dissolved in DMSO and the absorbance measured at 570 nm.

4 RESULTS

4.1 General growth characteristics

Following each physical adjustment to the position of the cells, a plot of the substrate and cells was produced, with each cell represented as a three dimensional spheroid with a radius/height relationship reflecting the dimensions of the model cell. This provided a visual representation of the model culture growth, similar to that which might be seen by a biologist looking down a microscope. Three snapshots of the simulated pattern of growth in low and physiological calcium concentrations after 50 iterations (25 hours), 75 iterations (37.5 hours) and 100 iterations (50hours) are shown in Figure 2.

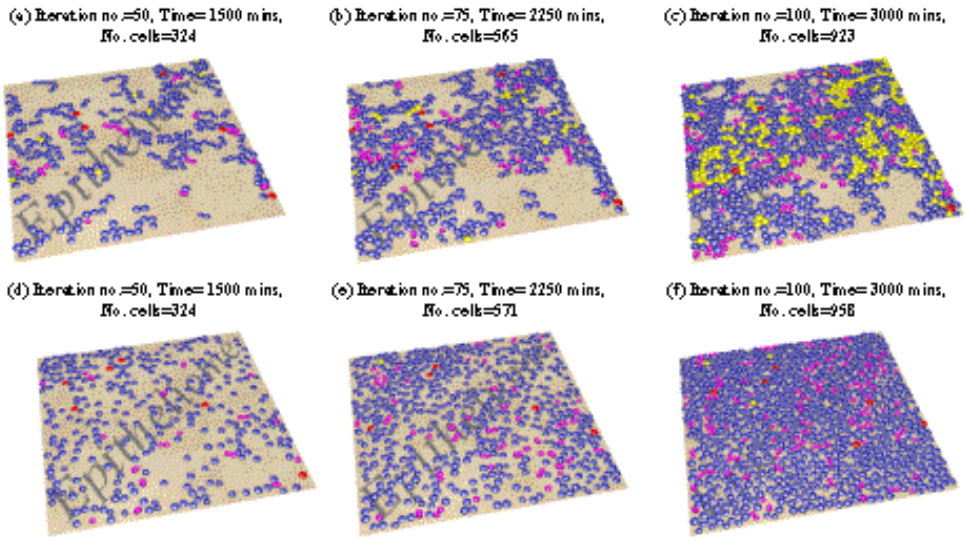

Figure 2 Pattern of simulated culture growth in physiological calcium (a-c) and low calcium (d-f). Snapshots correspond to iteration 50 (a,d), iteration 75 (b,e) and iteration 100 (c, f). Red cells represent stem cells, blue cells transit amplifying (TA), pink cells are mitotic and yellow cells are quiescent.

It is evident that there is a marked difference in growth pattern between the two models. At the start of both models, cells migrate randomly around the substrate, according to the rules outlined above. In the case of physiological calcium concentration, intercellular bonds are formed quickly when cells come within close proximity of one another, migration is suppressed and colonies appear early on (figure 2a). As the simulation progresses, small colonies begin to coalesce, forming progressively larger groups (figure 2b and c). Conversely, in the low calcium model, cells come into close proximity without the formation of intercellular bonds. Migration continues unabated and at any moment in time cells are fairly evenly distributed over the substrate, even at relatively high cell densities (figure 2d-f). Despite both models having the same number of seed cells at the start of the simulation, comparing

figure 2c and 2f, there is a larger number of cells in existence in the low calcium model after 100 iterations.

Figure 3 shows micographs of the in vitro cell distribution approximately 3 days after plating. In physiological calcium, the formation of colonies is apparent (figure 3a), whilst cells are more evenly distributed in the low calcium medium. Mitotic cells, which are highly rounded in comparison to the spread morphology of non-dividing cells, are indicated by arrowheads.

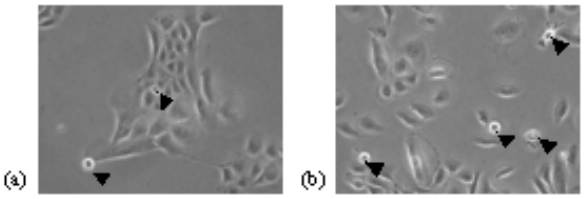

Figure 3 Micrographs of cell distribution in (a) physiological calcium and (b) low calcium medium, approximately 3 days after plating. Arrows indicate mitotic cells. Magnification is approximately x200.

4.2 Contact Inhibition

Quiescent cells in the G0 phase of the cell cycle are shown in yellow on the surface plots (Figure 2). In the case of the physiological calcium simulation, the first quiescent cell was visible at the centre of a small colony after only 14 model iterations. The number of quiescent cells increased steadily as the simulation progressed. At the end of the simulation (160 iterations), 85% of cells in the model were quiescent. Of these, 86% were contact inhibited, and 14% were prevented from leaving the G0 phase due to insufficient spreading.

In the case of the low calcium model, the first quiescent cell appeared at the 53rd iteration. The increase in the number of quiescent cells was extremely slow until the model appeared to reach confluence (approximately iteration 110), then expanded extremely rapidly as cells became squashed (see section 2.5). At the end of the simulation, 91% of the cells were quiescent. Of this number, only 7% were contact inhibited whereas 93% were insufficiently spread to continue through the G1 checkpoint.

4.3 Population growth

Figure 4 shows the population growth curves for the physiological and low calcium medium models.

There is little difference in the rate of growth of the models until approximately 70 iterations, where it appears that the degree of contact inhibition in the physiological calcium culture (approximately 10% of the total cell number) starts to have an effect. The time of visual confluence, where approximately 95% of the substrate is estimated to be covered by cells, is marked on the plot. This corresponds to a cell number of between 1300-1500, and is not identical for the two models due to difference in cell size, and the fact that in physiological calcium, cells are grouped in colonies hence covering the substrate in a more heterogeneous distribution. The low calcium model reaches this point approximately 10-15 iterations (20-30 hours) before the higher calcium model. After 160 iterations,

the total cell number in the low calcium model was 2652, but only 2087 in the physiological calcium model. This represents a 27% increase in the total cell population.

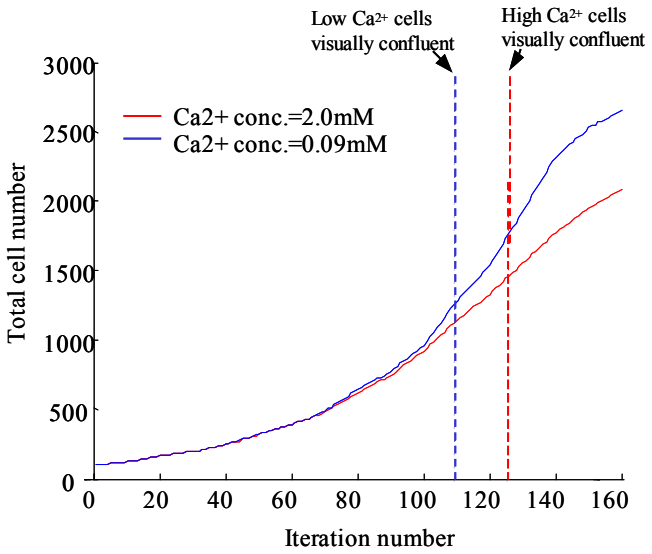

Figure 4 Growth curves produced by computational model (linear scale)

Figure 5 shows the optical density obtained from the *in vitro* MTT assay, which correlates with cell number. It can be seen that after an initial lag phase of approximately 1 day, the cell population in the physiological calcium medium begins to expand more rapidly. After approximately 5 days, a maximum cell number is reached and after 7 days, the absorbance is reduced, possibly indicating cell death. However, the expansion of the cell population in the low calcium medium appears to be continuing after the experiment was ended on the 10th day after plating.

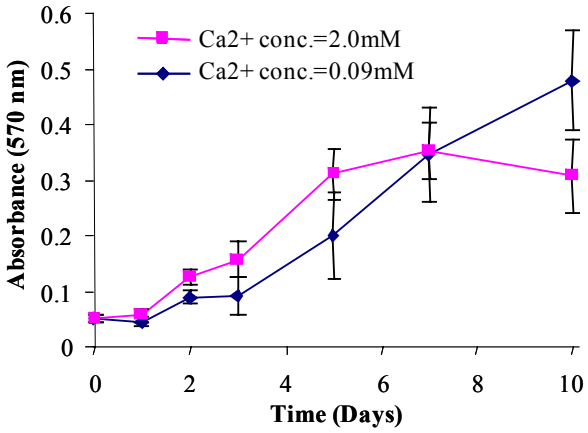

Figure 5 Results of MTT assay. Error bars represent standard deviation of 6 samples

5 DISCUSSION

When photomicrographs of NHU cell cultures were compared to simulations of the computational model at comparative cell densities, it was noted that in both biological and computational systems,

cells were fairly evenly distributed over the substrate in the low calcium conditions, but showed an increased tendency to form colonies when exposed to higher concentrations of calcium (Figure 2 and Figure 3). The computational model showed no significant difference in growth rate between populations in high and low calcium, but indicated that the maximal cell density at confluence would be higher in the low calcium conditions, presumably due to lack of contact inhibition (Figure 4). When the biological growth of NHU cells in low and high calcium concentrations was compared, two main points were noted. Firstly, in agreement with the computational model, the ultimate cell density attained was greatest in low calcium conditions and secondly, that there was a longer lag phase in cell cultures grown on a low calcium medium, that was not predicted from the computational model. This latter observation suggests that there are other factors that have not been taken account of in the computational model, such as juxtacrine signalling, which may stimulate cell proliferation through direct cell-cell contact (e.g. Nelson, 2002). Indeed, we have biological evidence for the existence of an epidermal growth factor receptor-stimulated autocrine loop that regulates proliferation in NHU cells and is mediated via a juxtacrine pathway (manuscript in preparation).

6 CONCLUSION

We have developed a rule-based model of monolayer cell culture growth, considering individual cells as autonomous agents executing a set of rules according to differentiation state, position in the cell cycle, and the immediate environment. We have coupled our agent-based model to a physical model, which governs cell-cell separation using a numerical technique based on the application of simple repulsive and attractive forces.

Comparison of the growth characteristics of simulations incorporating low and physiological environmental calcium concentrations yielded a good qualitative fit with the results of *in vitro* culture growth under similar conditions. This comparison between *in silico* and *in vitro* systems demonstrates that even a model based on simple rules governing cell cycle, intercellular bonding and basic physical relationships between neighbouring cells can successfully reproduce the behaviour of a real biological system. The model will be extended to include more complex intercellular signalling and three dimensional physical interactions between neighbouring cells, allowing model cells to leave the proliferative compartment, differentiate and migrate upwards. We will identify a subset of genetic markers that are closely correlated with cell growth and differentiation for inclusion in the model. These developments, which will involve close co-operation between computational scientists, engineers and cell biologists, will allow us to build a comprehensive model of epithelial morphogenesis, homeostasis and potentially pathogenesis incorporating elements from the scale of expression of individual genes, to macroscopic tissue properties and behaviour.

7 REFERENCES

Bassingthwaighte, J. B. (2000). "Strategies for the physiome project." Annals of Biomedical Engineering **28**(8): 1043-1058.

Baumgarter W, Hinterdorfer. P., Ness W, Raab A, Vestweber D, Schindler H, Drenckhahn D (2000). "Cadherin interaction probed by atomic force microscopy." Proceedings of the National Academy of Science USA **97**(8): 4005-4010.

Dover R, Potten. C. S. (1983). "Cell Cycle Kinetics of Cultured Human Epidermal Keratinocytes." Journal of Investigative Dermatology **80**: 423-429.

Gheorghe M, Holcombe. M., Kefalas P (2001). "Computational models of collective foraging." BioSystems **61**: 133-141.

Holcombe M, Bell. A. (1998). Computational models of immunological systems. Information Processing in Cells and Tissues. P. R. Holcombe M, Plenum Press**:** 213-226.

Huang S, Ingber. D. E. (1999). "The structural and mechanical complexity of cell-growth control." Nature Cell Biology **1**: E131-E138.

Jensen U B, Lowell. S., Watt F M (1999). "The spatial relationship between stem cells and their progeny in the basal layer of human epidermis: a new view based on whole-mount labelling and lineage analysis." Development **126**: 2409-2418.

Jones D M, R. H. Smallwood, D. R. Hose, B. H. Brown (2003) "Finite Element Analysis of Electrical Properties of Glandular Mucosa." Physiological Measurement Submitted

Morel D, Marcelpoil. R., Brugal G (2001). "A proliferation control network model:the simulation of two-dimensional epithelial homeostasis." Acta Biotheoretica **49**: 219-234.

Noble, D. (2002). "Modeling the Heart - from Genes to Cells to the Whole Organ." Science **295**: 1678-1682.

Orford K, Orford. C. C., Byers S W (1999). "Exogenous expresion of Beta-Catenin regulates contact inhibition, anchorage-independent growth, anoikis and radiation-induced cell cycle arrest." Journal of cell biology **146**(4): 855-867.

Potten C S, Booth. C. (2002). "Keratinocyte Stem Cells: a Commentary." Journal of Investigative Dermatology **119**: 888-899.

Potten C S, Morris. R. J. (1988). "Epithelial stem cells in vivo." Journal of Cell Science Suppl. **10**: 45-62.

Smallwood R, Holcombe M, Walker D, Southgate J, Mac Neil S (2003) "Computational modelling of the social behaviour of cells." BioSystems (submitted)

Southgate J, Hutton K. A. R., Thomas D F M, Trejdosiewicz L K (1994). "Normal human urothelial cells in vitro: proliferation and induction of stratification." Laboratory Investigation **71**(4): 583-594.

Southgate, J., Masters J R, Trejdosiewicz L K, . (2002). Culture of Human Urothelium. Culture of Epithelial Cells. M. G. Freshney. New York, J Wiley and Sons, Inc.**:** 381-400.

St Croix B, Sheehan. C., Rak J W, Florenes V A, Slingerland J M, Kerbel R S (1998). "E-Cadherin-dependent growth suppression is mediated by the cyclin-dependent kinase inhibitor p27 KIPI." Journal of cell biology **142**(2): 557-571.

Stekel D, Rashbass. J., Williams E D (1995). "A computer graphic simulation of squamous epithelium." Journal of theoretical biology **175**: 283-293.

Venter, C. *et. al.* (2001). "The sequence of the human genome." Science **291**: 1304-1351.

Walker D C, Brown B. H., Smallwood R H, Hose D R, Jones D M (2002). "Modelled current distribution in cervical squamous tissue." Physiological Measurement **23**(1): 159-168.