Case Essay

An image-processing tool for the quantitative study of the developing *Drosophila* retina

Nicolas Jaccard*

Supervisors: Dr. Lewis Griffin** | Dr. Franck Pichaud***

Center for Mathematics and Physics in the Life Sciences and Experimental Biology, University College London, London*
UCL Department of Computer Sciences, University College London, London**
MCB, MRC Cell Biology Unit & Department of Developmental Biology, University College London, London***

The *Drosophila* retina is a highly ordered tissue in which cells adopt specific patterns with regularity reminiscent of inorganic crystal structures. Through morphogenesis, a coordination of cell growth and shape changes, unorganized epithelial sheets give rise to biologically active tri-dimensional structures, such as the *Drosophila* eye. Morphogenesis is driven by forces resulting from the subtle interplay of genetics, biochemistry and cellular mechanics. Most of the current knowledge is derived from fluorescence microscopy techniques that allow the imaging of cell patterns formation processes. However, it is only by adopting a system approach by integrating notions and tools borrowed from other fields such a fluid mechanics, material physics and computer sciences that it will be possible to fully understand the mechanisms underlying the formation of the *Drosophila* retina. More particularly, there is a need to make a transition from qualitative observations to robust and reproducible quantitative data.

The goal of this project was to develop an image-processing program that could be used to extract information about cells, borders and vertices from a previously delineated fluorescence microscopy picture of the developing *Drosophila* retina and organize this data in a query-able format, allowing various quantitative analyses to be performed.

The image-processing software was programmed using Mathematica and allowed the quantitative analysis of biologically relevant features through a user-friendly interface. The distribution of size and number of neighbors were determined in a region-specific manner. Cells in the patterned regions (morphogenetic furrow and post-furrow) tended to have narrow distributions both in size and in number of neighbors in opposition to the unorganized pre-furrow regions where both distributions were found to be much broader. This might indicate that these analyses could serve as a basis to quantify the degree of disorder of epithelial tissues. Moreover, it was shown the numbers of neighbors of adjacent cells were dependant and that this dependency varies from a region to another.

1. Introduction

Despite their formidable complexity and diversity, multicellular organisms originate from a single fertilized cell. Through ontogenesis, the succession of steps required for the development of an organism from the embryo to its adult form, cells follow a rigorous genetically encoded construction plan that will eventually lead to the emergence of complex structures with specific functions and shapes, such as tissues and organs. The repertoire of cellular mechanisms to achieve this gargantuan task is rather limited: a cell can divide, grow in size, migrate, die (apoptosis) or undergo differentiation, a genetic-driven process during which cells gain specialized phenotypes, often accompanied by a change in shape to accommodate specific functions, such as the disk-like shape of red blood cells, conferring them a maximized surface area for gas transfers.

Morphogenesis is a tightly-controlled coordination of cell growth and shape changes that allow unorganized groups of cells to adopt defined and biologically active tri-dimensional structures. Cells are subject to a complex physiological environment and must maintain homeostasis by exhibiting two seemingly antagonistic properties: robustness to resist to various external forces and plasticity to allow the remodeling in case adaptation to new conditions is necessary (Lecuit and Lenne, 2007). Cellular fate is not solely dependent on genetic programming but is also heavily affected by cellular mechanics. Understanding the interplay between molecular mechanisms and the intrinsic physical properties of the cells and tissues can be challenging. The extensive study of model organisms such as *Drosophila* combined with mathematical modeling and computational simulations are starting to reveal some biological and mechanical aspects of the processes involved (Lecuit and Le Goff, 2007).

1.1 The *Drosophila* retina and its development

This section of the report introduces the *Drosophila* retina and summarizes the current understanding of its formation during the development of the organism. If not otherwise specified, the information presented here is taken from the 'Pattern formation in the *Drosophila* retina' chapter of 'The development of *Drosophila melanogaster*' (Bate and Martínez Arias, 1993). In addition, complementary insights from the recent literature are mentioned.

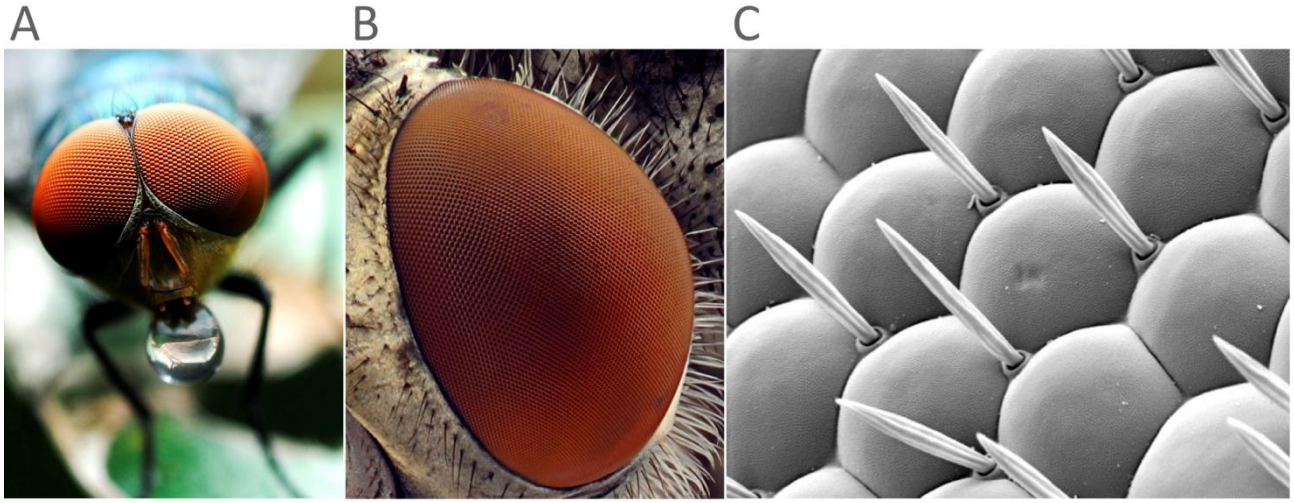

Figure 1 Pictures of the eye of a fly at different magnification levels. **A** Macro-photography of a fly **B** Detailed view of the eye of a fly **C** Electron micrograph of *Drosophila*'s eye (images taken from the Bored Panda website: http://www.boredpanda.com)

With the help of simple magnification devices or even with naked eyes, the highly patterned structure of the eye of a fly can easily be distinguished (Figure 1 A and B) but it is only by using high resolution techniques such as electron microscopy that is possible to appreciate the incredibly ordered nature of the *Drosophila*'s visual sensory organ (Figure 1 C). The compound (or unit) eye of *Drosophila melanogaster* is made up of approximately 750 facets or ommatidia, which are microstructures acting as lenses and containing photosensitive cells.

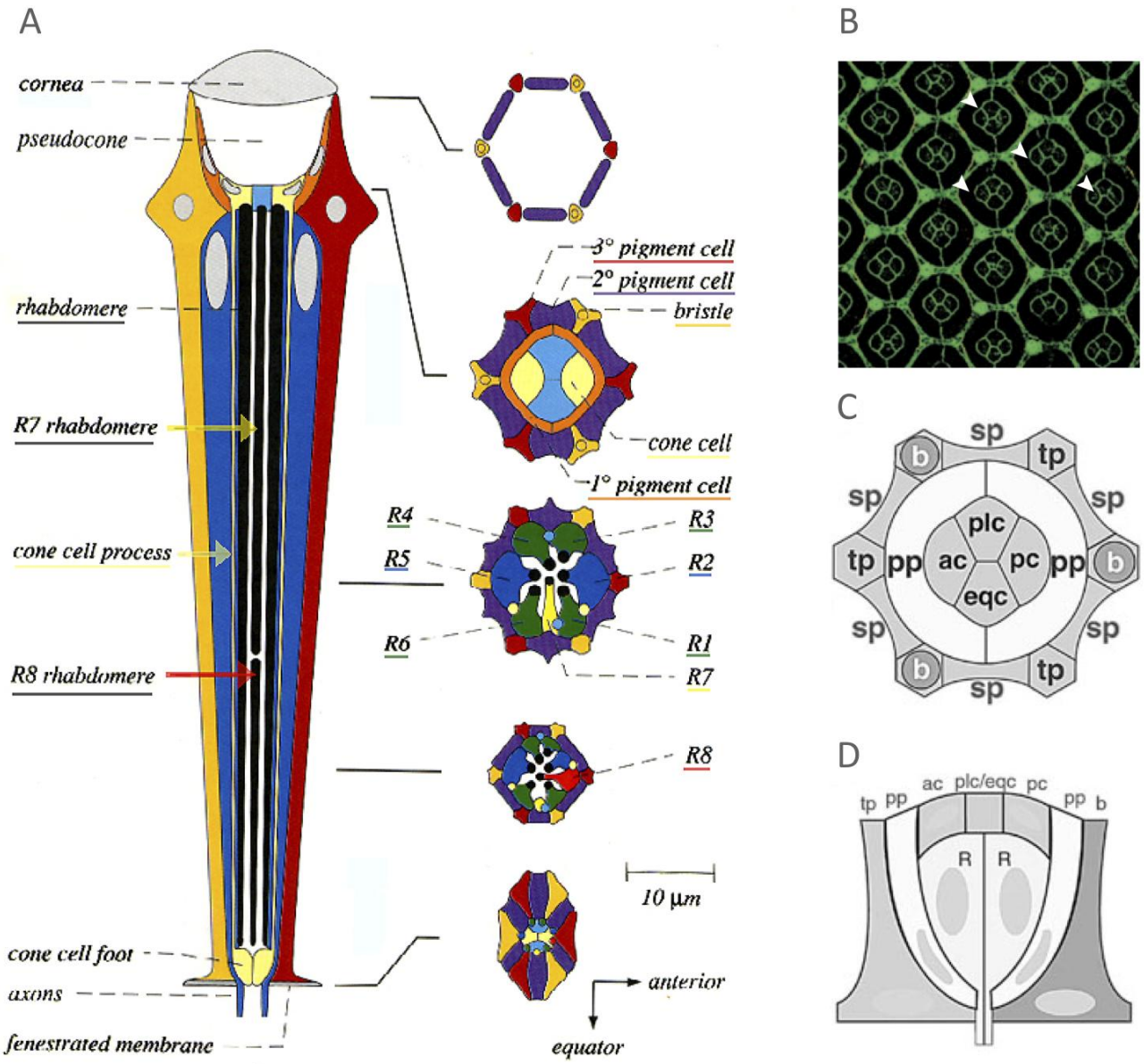

Figure 2 Cellular architecture of *Drosophila* ommatidia **A** Composition of the adult ommatidium **B** Retinal fly epithelium with E-cadherin labeled in green **C** and **D** Cross-section view (C) and side view (D) schematics of an ommatidium at 35% pupal life (plc = polar cone cells, ac = anterior cone cells, pc = posterior cone cells, eqc = equatorial cone cells, pp = primary pigment cells, tp = tertiary pigment cells, sp = secondary pigments cells, b = bristles). A is reproduced from (Bate and Martínez Arias, 1993), B from (Lecuit and Lenne, 2007), C and D from (Hayashi and Carthew, 2004)

The adult ommatidium is composed of 19 cells (Figure 2.A). The main part of the structure is made up of eight photoreceptor cells (R cells) that are elongated along the proximo-distal axis of the retina and bundled in a cylindrical fashion. Each of these cells posses a region rich in rhodopsin (a pigment part of the light detection mechanism) called rhadomere, which harbors some 60'000 microvilli that are projected towards the center of the cylindrical structure. This central part of an ommatidium, often referred as the rhabdom, is sealed at the top and bottom by four cone cells and wrapped into two primary pigment cells. Both cone and primary pigment cells secrete a chitinous material that forms a biconvex corneal lens at the top of the structure. Sheath of secondary pigment cells surround the ommatidium, optically insulating it from other facets. Each vertex of the hexagonal shape (Figure 2 B) is made up of either a bristle or a secondary pigment cell.

While eye cells specialization mainly occurs during the last two thirds of pupal life, the structural patterns are already final around 35% of the development (as shown on Figure 2 C and B compared to A). The unpatterned monolayer of cells that will give rise to the retina originates from the blastoderm of the embryo. A few cells of the

dorsolateral ectoderm will eventually form the eye-antennal imaginal disc after involution of the head. Through the two first instars to the beginning of the third, the disc grows by a succession of cell divisions. Once the disc attains an appropriate size (it is estimated that around 10'000 unpatterned cells are required for the formation of 750 facets), the proliferation process slows down. At this point, the morphogenetic furrow, an indentation in the imaginal disc, appears at the intersection of the posterior margin of the eye disc and the midline (Kumar, 2001), as shown in Figure 3.A and B.

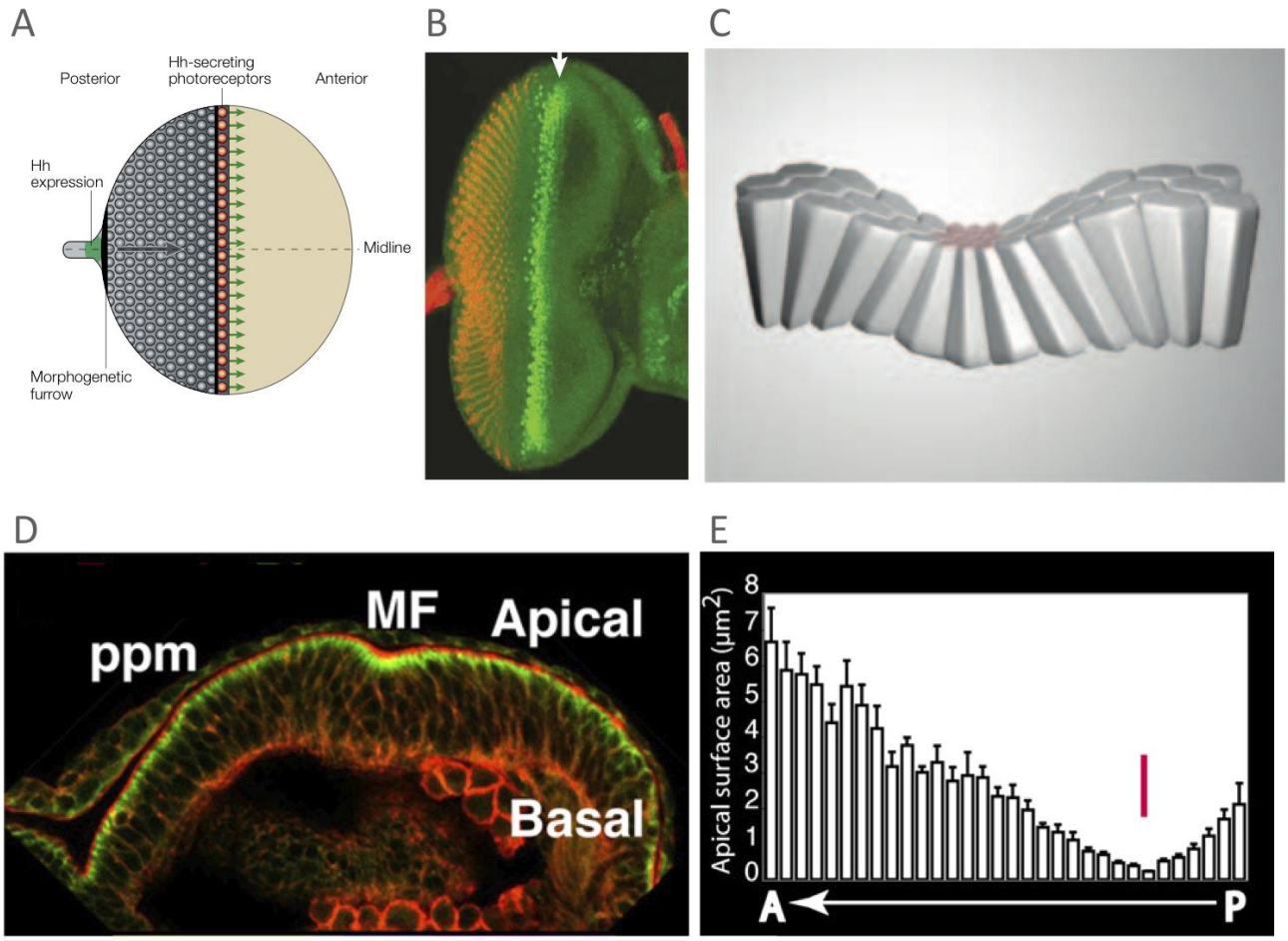

Figure 3 Different representations of the morphogenetic furrow. **A** Schematic of the *Drosophila* eye imaginal disc **B** Picture of the *Drosophila* imaginal disc stained for *atonal* (green), forming ommatidia are in red **C** 3D representation of the indentation of epithelial cells **D** Sagittal section of a developing eye labeled for F-actin (red) and Dig (green), ppm=peripodial membrane, MF shows the morphogenetic furrow **E** Apical surface area of the cells, the red line represents the position of the morphogenetic furrow. A and B are reproduced from (Kumar, 2001), C from (Lecuit and Lenne, 2007), D is adapted from (Corrigall, Walther, 2007) and E is reproduced from (Corrigall, Walther, 2007)

The furrow results from an apical constriction of the imaginal disc cells (Figure 3.C), which is a conserved and fundamental cell response in developmental biology (Lecuit and Lenne, 2007). Cell constriction involves the accumulation of F-actin in the apical surface of the cells and the action of non-muscle Myosin II which provides the mechanical force leading to a dramatic decrease in the surface area of the cells and eventually the bending of the structure as shown in Figure 3.D and E (Corrigall, Walther, 2007). The morphogenetic furrow can be pictured as a wave travelling across the unpatterned epithelium with ommatidia forming in its wake (Figure 4.A). The first discernable pattern to emerge is a rosette made up of 10 to 15 cells with a 4-5 cells core. After dissociation of the anterior cells, the structure elongates and forms an arc. Anterior cells join along the midline and leads to a five cells pre-cluster. New rows of ommatidia precursors appear every 1.5 hours and 20 minutes separate two clusters in a given row (Figure 4.A).

Cells that remains unpatterned behind the morphogenetic furrow divide, replenishing the reservoir of candidates that can subsequently be recruited in a stereotyped manner into the cluster to give rise to the adult structure. As they mature, the forming ommatidia undergo a rotation (Figure 4.B), result of the coordinated action of adhesion proteins (Mirkovic and Mlodzik, 2006).

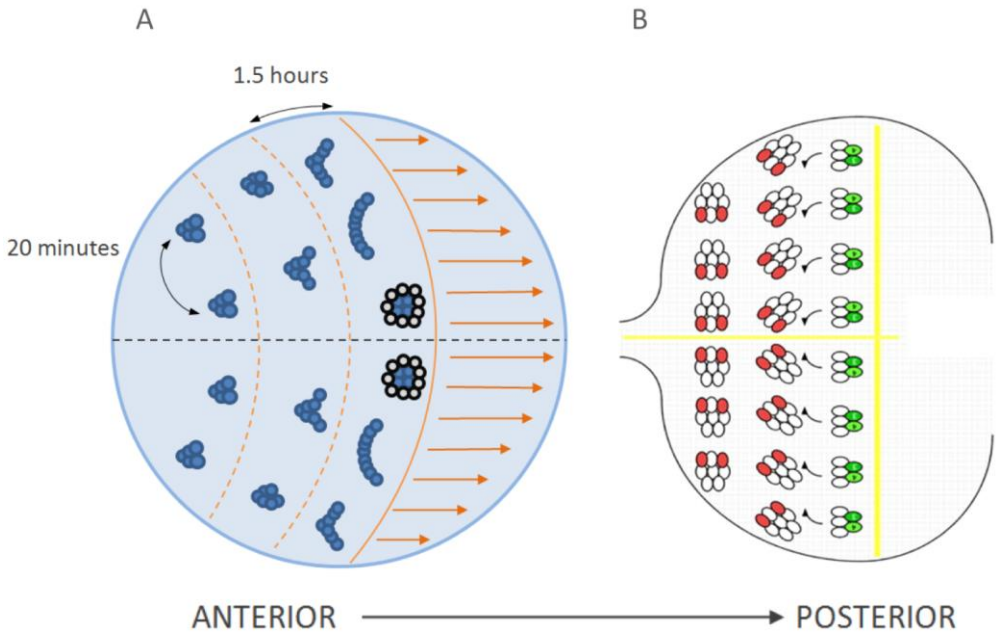

Figure 4 Formation of the pre-clusters and their maturation. **A** Progression of the morphogenetic furrow (orange) across the imaginal disc epithelium and formation of the five cells preclusters.The dashed grey line represents the equator (midline). **B** Maturation of the clusters and their rotation. B adapted from (Mirkovic and Mlodzik, 2006)

1.2 Understanding the mechanisms leading to the emergence of new patterns

Many models have been proposed to explain the emergence of complex cellular patterns in developing tissues. Exploring all of them in details would be a daunting task and is beyond the scope of this document. Instead, the mechanisms underlying the formation of patterns in the *Drosophila* retina, which are mainly controlled by adhesion, will be briefly introduced.

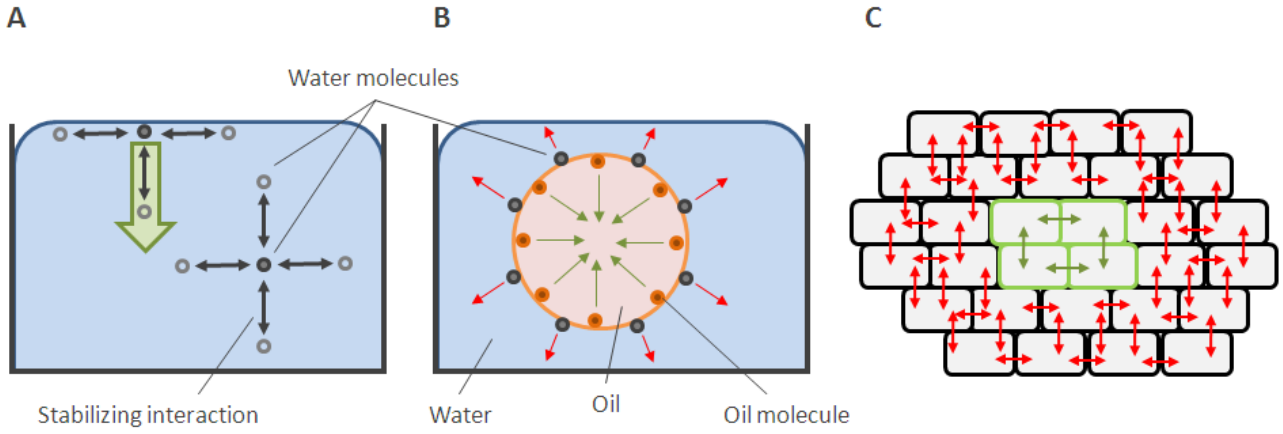

Figure 5 Surface tension **A** In the bulk, the molecule is stabilized by the interaction with its neighbors. The forces applied to the particle at the surface are unbalanced and result in a net force (green arrow) pushing the molecule towards the bulk. **B** Example of a drop of oil in water **C** Same analogy applied to cells expressing adherence proteins (green borders) and those that don't (black borders)

In the context of biological modeling, it is often interesting to find analogies relevant to a particular problem in other well-studied and understood fields. In this case, it seems appropriate to relate the question of epithelial patterning to that of viscous fluid dynamics, more especially to the notion of surface tension (Forgacs, Foty, 1998). In fluids, surface tension is generated when molecules at the surface are no longer stabilized by their neighbors and get pulled back towards the bulk, creating a surface film capable of exerting a force outward (Figure 5.A). It also implies the minimization of the area of the fluid's surface. This property explains the separation of oil and water where molecules only interact with those of the same fluid. As a result, water molecules cannot stabilize oil surface molecules, which will be pushed away and forced to minimize their exposed surface area, most likely by adopting a spherical conformation (Figure 5.B).

Similar phenomena can be observed when populations of cells with different adhesion properties are mixed. The simplest case involves the mixing of cells that express a specific adhesion protein and others that do not (Figure 5.C). Cells harboring the adhesion protein tend to bind together in a cluster, ending up surrounded by those that do not, much like an oil droplet would be surrounded by water. The idea that the difference in adhesion properties can lead to the emergence of complex patterns have been successfully tested both experimentally (Foty and Steinberg, 2005) and through computational simulations (Mombach, Glazier, 1995).

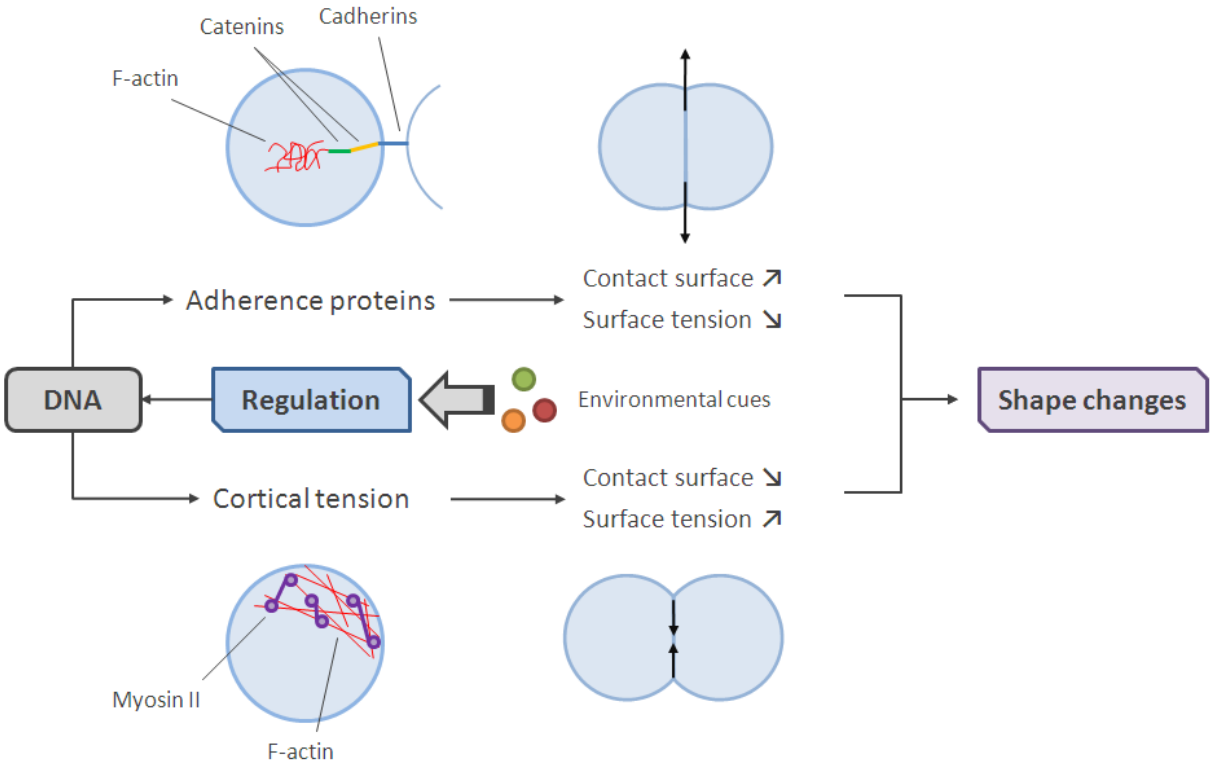

Figure 6 Possible mechanism for the regulation of cell shape in tissues. Environment cues (e.g. external force or morphogen signaling) regulate the expression of the proteins responsible for adherence or cortical tension and their spatial localization in the cell. The difference in surface tension leads to changes in cellular shape

To further push the fluid analogy, cells assemblies have been shown in particular cases to follow the set of rules proposed by Plateau to predict the shape of bubbles aggregates in soap films (Hayashi and Carthew, 2004). Unlike bubbles, complex structures involving cells do not form by membrane fusion but instead must rely on transmembrane proteins such as cadherins, which have one end connecting to a neighboring cell and the other one deeply anchored to the cell's actin network via intermediate proteins called catenins. E-cadherin in particular is actively involved in the adhesion of cells in tissues and is thought to play a significant role in the emergence of cellular patterns. For instance, the differential expression of N and E-cadherins was proposed as a possible mechanism for the regulation of the surface tension responsible for the complex structure adopted by cone cells in *Drosophila* retina (Hayashi and Carthew, 2004). Similarly, another mechanism linked with the rearrangement of cells during the pupal maturation of ommatidia was shown to depend on the expression of the

immunoglobulin domain cell adhesion molecules *Hibris* and *Roughest*, which can bind together *in-vivo*. Cells expressing the same molecule will tend to minimize their contact whereas cells expressing different molecules will adhere tightly and thus maximize their contact surface (Bao and Cagan, 2005).

It has also been suggested that, the adhesion properties of the cells are not solely responsible for the changes in cellular shape that eventually lead to the rearrangement of the cells in a tissue. Instead, a hypothetical model (Figure 6) is based on the competition between adhesion, which increases the contact surface and thus lowers the surface tension between cells, and cortical tension (e.g. generated by the contraction of actomyosin networks) that, in the contrary, decrease the contact surface, resulting in an increased surface tension (Lecuit and Lenne, 2007). An interesting point is that F-actin is involved in both the adhesion and cortical tension mechanisms, which could suggest that its distribution in the cell and the structure of the network it forms could be the main regulator of cellular shape changes, modulated by physical (e.g. shear or compressive forces) or molecular (e.g. morphogen gradient) environmental cues.

Some other aspects of pattern formation still need to be completely understood, such as the mechanisms that allow cells to maintain junctions' integrity during the addition of new cells to the cluster, either following the division of a member of the cluster or after the recruitment of an external cell. Interestingly, these mechanisms seem to involve the same key players than what was described for the regulation of cellular shape (Figure 6). Indeed, cells intercalation, which requires the remodeling of junctions (Figure 7.A), has been shown to be driven by the polarized expression of Myosin-II and PAR3, a promoter of cell adhesion through the stabilization of E-cadherins (Zallen and Wieschaus, 2004), as shown in Figure 7.B.

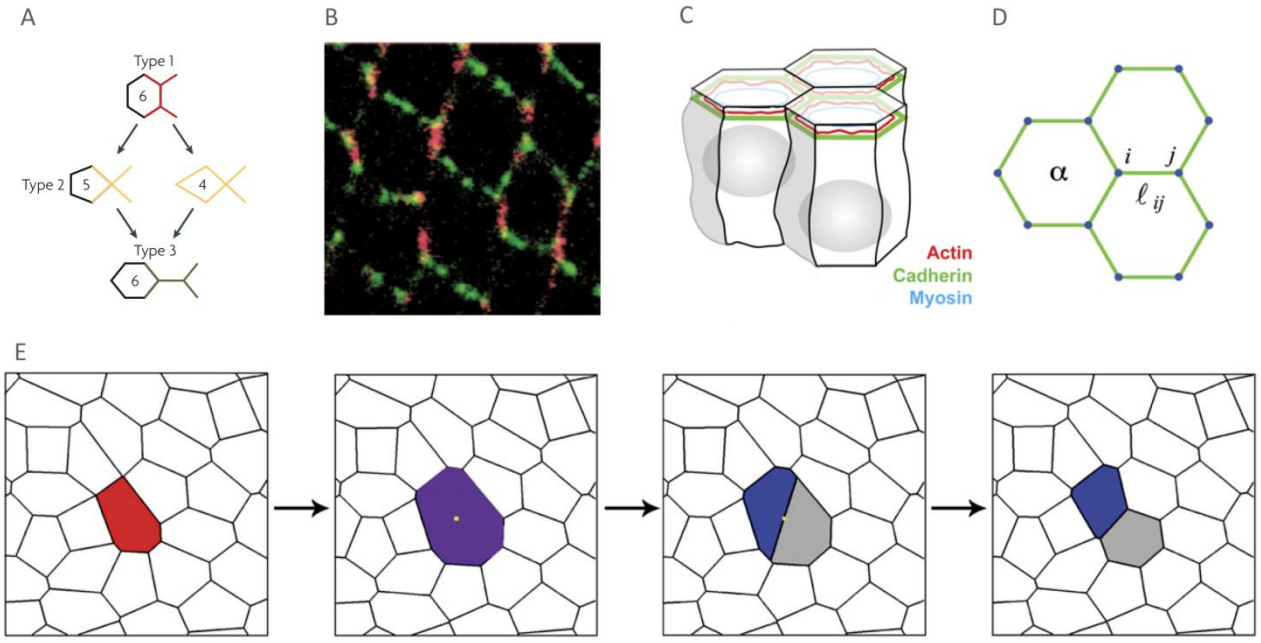

Figure 7 Cell intercalation. **A** Different shapes adopted during cell intercalation. **B** Differential expression of Myosin-II (red) and PAR3 (green). **C** & **D** Integrating the effects of actin, cadherins and myosin in a vertices-based model of cells packaging **E** Cell intercalation as simulated using a vertices-based model. A et B adapted from (Lecuit and Lenne, 2007). C,D and E adapted from (Farhadifar, Röper, 2007)

The ability to label these different molecules allows to image cell patterns formation processes and to gain a better understanding of the underlying mechanisms. However, it is only by adopting a system approach by integrating notions and tools borrowed from other fields such a fluid mechanics, material physics and computer sciences that it will be possible to fully understand the interplay between the genetics, biochemisty and cellular mechanics that lead to the emergence of complex patterns in tissues. Indeed, a vertices-based model taking into account cell elasticity, adhesion molecules and the action of actomyosin networks (Figure 7.C and D), was used to

successfully predict cell packing during *Drosophila* wing development and the response of cells to laser ablation (Farhadifar, Röper, 2007).

1.3 Project motivation and objectives

The typical workflow for the analysis of microscopy pictures involves several semi-automated and manual operations, rendering the generation of large amount of data at best tedious and sometime simply impracticable. The current understanding of the patterning in the developing *Drosophila* retina is thus mostly based on static snapshots at different points in time due to the low output rate of current methods. Furthermore, multipurpose image processing packages rarely offer all the functionalities that would be required for the analysis of cellular patterns. The highly dynamic nature of the studied processes means that a full understanding of the underlying mechanisms entails the development of a high throughput analysis workflow capable of processing multiple frames of a movie to generate robust and quantitative experimental data that could then eventually be used to feed a computational model.

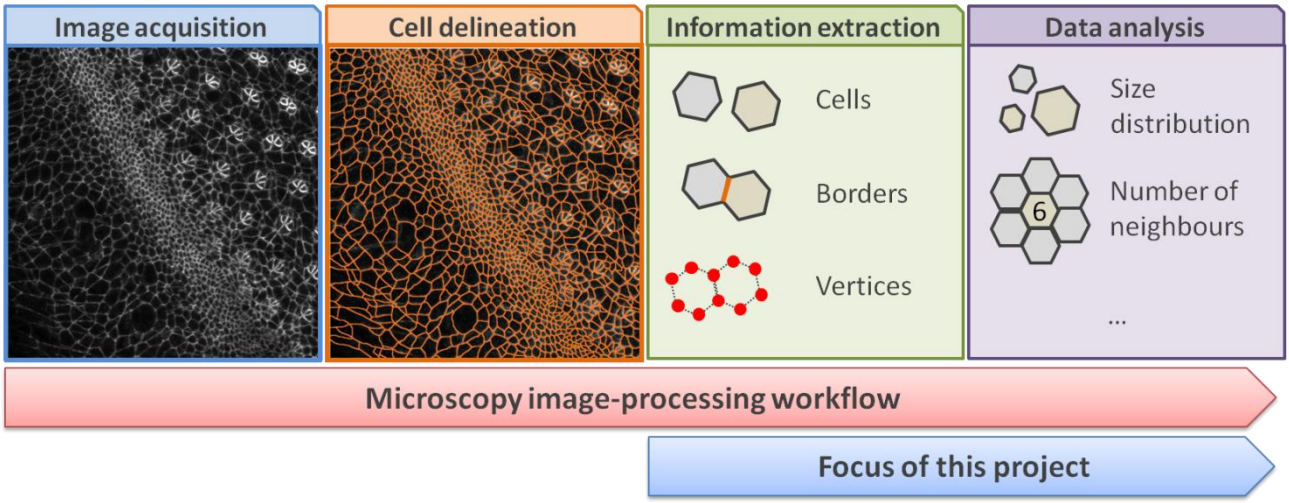

Figure 8 Microscopy image-processing workflow and the focus of this particular project

The goal of this project was to develop an image-processing program that could be used to extract information about cells, borders and vertices from a previously delineated microscopy picture and organize this data in a query-able format, allowing various quantitative analyses to be performed (Figure 8).

2. Methods

The image-processing tool was programmed using Mathematica (Version 7.0.1.0, Wolfram Research Inc.). The sample data used for the development and validation of the method comprised of a confocal microscopy image of the developing retina in *Drosophila* (taken during the third instar, approximately 75 hours after egg laying) and its manually delineated equivalent (Figure 9). The user interface is shown in annex 6.1.

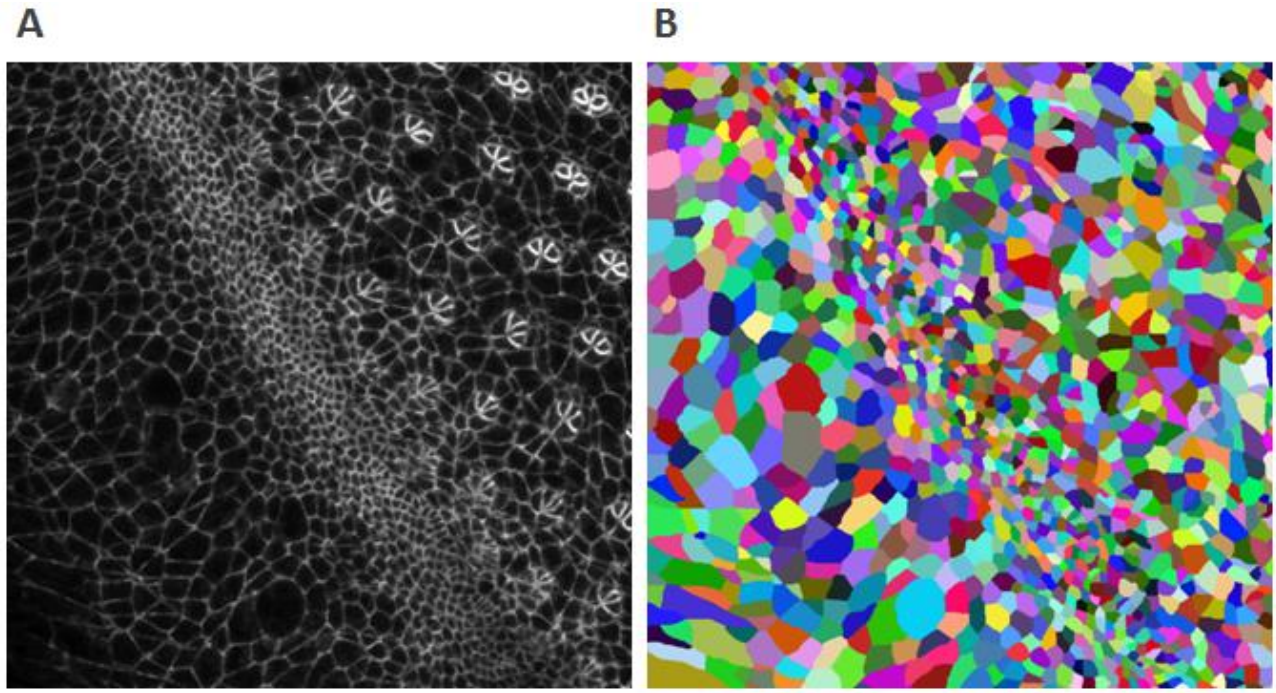

Figure 9 Sample data for the development and validation of the image-processing program: picture of the eye imaginal disc during the third instar (approximately 75 hours after egg laying) **A** Confocal microscopy image showing labeled borders (by fluorescently tagging cadherins with GFP). **B** Manually delineated version of A (cells are randomly colored)

2.1 Data structure

Three different types of features were of interest when processing the picture: cells as a whole, borders (defined as a set of pixel shared by two cells) and vertices (defined as points that are shared by three or more cells). Data related to each of these features was stored in multidimensional arrays as following (Figure 10):

- **Cells**: coordinates of the points were stored in a two-dimensional array, the first dimension being the identification of the cell (ranging from 1 to the maximum number of cells) and the second dimension representing a given point of the cell (ranging from 1 to the maximum number of points in a given cell).
- **Borders**: a given border entry (defined as the set of pixels of a cell that had neighbors belonging to another cell, see Figure 11.B) contained the identification of the cells it was associated with in addition to the coordinates of the points part of the border (separate list for each of the two cells).
- **Vertices**: vertices were defined as points that would be at the intersection of the contours of cells if these were polygonal. A given vertex entry had its coordinates and a list containing the identification of the cells it was associated with.

This structure allowed relying on heavily optimized array functions that are part of Mathematica and using functional programming whenever possible.

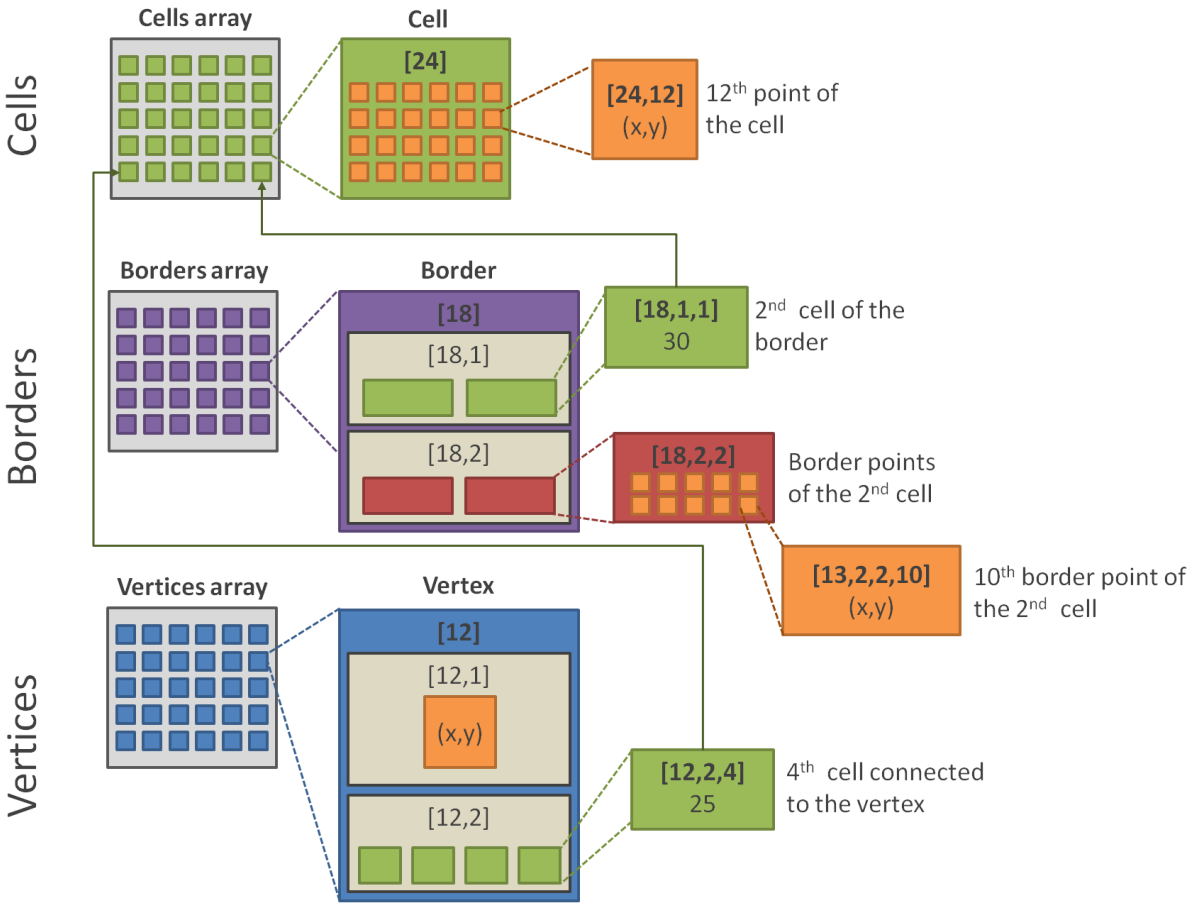

Figure 10 Data structure for cells, borders and vertices

2.2 Data extraction

Due to its computation intensive nature, data extraction was performed in a separate part of the program that had to be run once per microscopy picture to generate the query-able data structure discussed above. The delineated picture (Figure 9) came in the form of a binary file containing a list of integers, which were used to associate each pixel with the cell it represents. The array containing cell data was built by traversing the file and appending the coordinate of each pixel to the corresponding entry in the array (Figure 11.A). Borders were determined by scanning the neighbors of each pixel (Figure 11.B). If the scanned neighbor belonged to another cell, its coordinate were added to the corresponding borders entry. Vertices were detected using a similar approach but a single pixel needed to be neighbor with at least two distinct cells (Figure 11.C).

These raw results are then filtered in order to keep only unique and relevant information. Duplicate points coordinates generated by the borders determination method were deleted and vertices were combined if they happened to be within a certain distance (as chosen by the user) of each other and had at least three cells in common.

This process only had to be run once per picture and data would be saved so it can be analyzed multiple times without having to go through this time consuming step again. The file format used was internal Mathematica input format, with a single file containing the definitions for cells, borders and vertices.

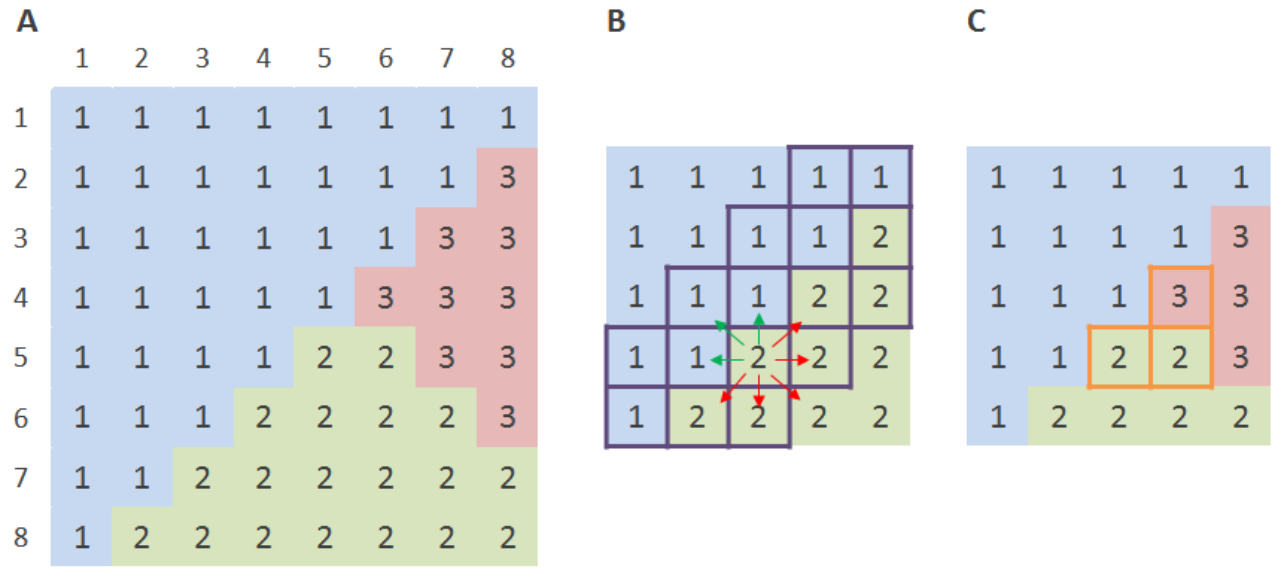

Figure 11 Methods to extract cells, borders and vertices data. **A** Representation of the input image, each pixel is represented by an integer that relates it to its corresponding cell. **B** Borders are determined by scanning the neighbors of each pixel. Green arrows indicate that a border has been detected (the target pixel belongs to another cell), red arrows indicate that the neighbor is associated with the same cell than the pixel of interest. Purple borders indicate pixels that would be detected as being part of the border. **C** Vertices are determined similarly to borders but a pixel must have at least two neighbors belonging to distinct cells. Orange borders indicate pixels that would be detected as vertices

2.3 Data manipulation and helper functions

Multiple helper functions were written to facilitate access to the data and to perform various analyses. Most of these were regular array manipulation and plotting functions but a few should be described in more details. The program allowed the user to define regions on the original picture to conduct location-specific analyses. It was decided to use the center of mass as a basis for the localization of cells. The centers of mass were determined by averaging the coordinates of the points on the border of the cells. The next step was to work out in which region a cell was situated. To determine whether a center of mass lies inside a region (defined as a polygon), a discrete form of a Gauss integral was used (original Mathematica implementation by Rafael Torres Cart[1]). If V is the set of N vertices forming the polygon and p the center of mass, the integral can be computed as shown in (Equation **1**.

$$I = \sum_{1}^{N-1} Angle(V_{i+1} - p, V_i - p) + Angle(V_N - p, V_1 - p) \qquad \text{(Equation 1)}$$

If the result of the integral is not zero, the center of mass lies inside the polygon. Another point of interest concerned the criteria that define the neighbor relationship between two cells. In the simplest case it was relatively straight forward, two cells would be flagged as neighbors if they shared a common border. But some limit cases made it clear that this definition wasn't stringent enough (as shown in the results). It was decided that in order to qualify as a neighbor relationship, the border shared by a cell and a potential neighbor had to be longer than a minimal size proportional to the perimeter of the cell of interest.

$$dist_{min,i} = \alpha \cdot \sqrt{perimeter(cell_i)} \qquad \text{(Equation 2)}$$

Where α is a proportionality coefficient that could be changed by the user. In addition, it is important to note that cells on the edge of the picture were not taken into account for the various calculations.

[1] Mathematica-users Wiki Homepage (http://www.mathematica-users.org/)

3. Results and discussion

The following results were obtained using the segmentation data presented in Figure 9. The borders were correctly determined (Figure 12.A) and vertices detection method showed promising preliminary results (Figure 12.B). Although vertices were correctly associated with the cells they delimitate, more work will be required to determine the branching between them.

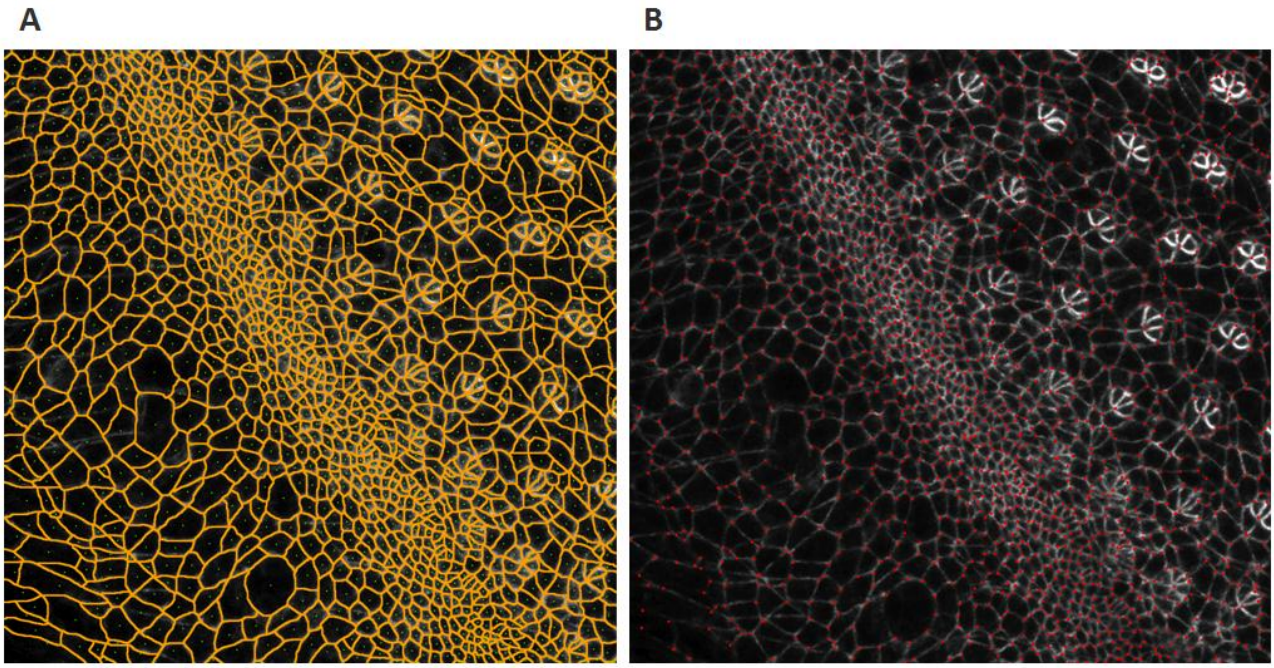

Figure 12 Features of the data as determined by the software **A** Borders of the cells **B** Preliminary results for vertices

Three regions of interest were defined using the software's user interface: region 1 corresponded to the unpatterned cells (in front of the morphogenetic furrow), region 2 to cells in the furrow and region 3 to cells that were behind the furrow, where patterns were already present or being formed (Figure 13.A).

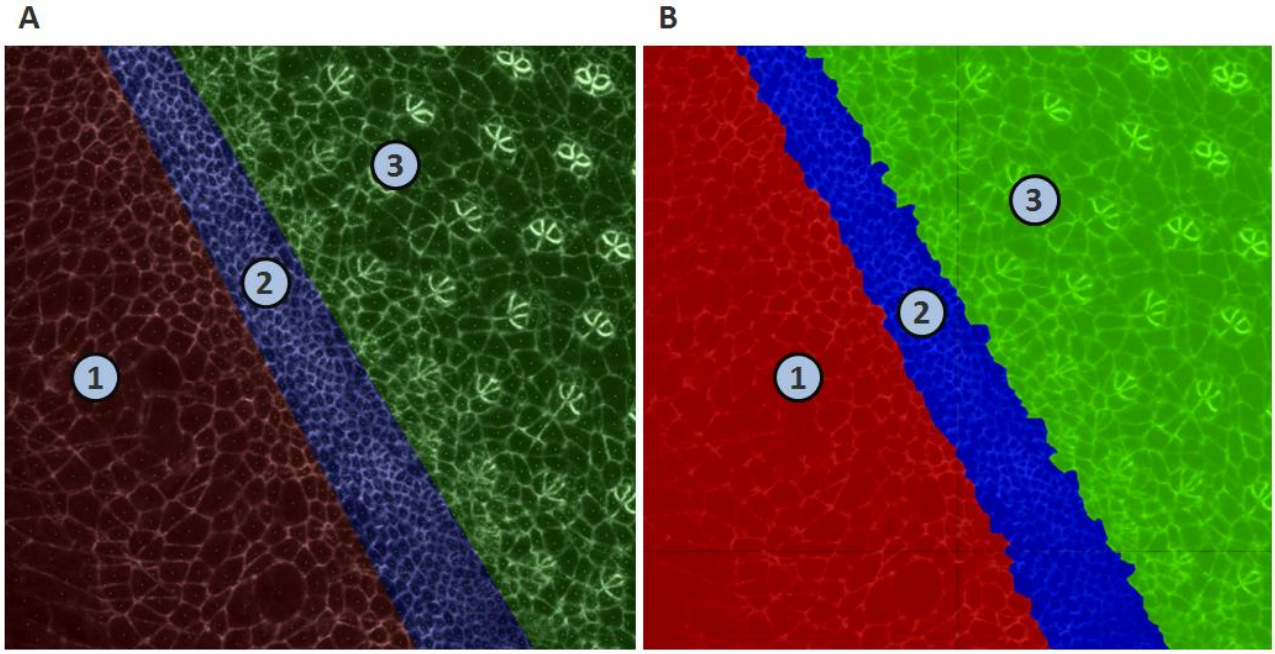

Figure 13 Regions of interest: (1) pre-furrow, (2) furrow and (3) post-furrow **A** Regions defined using the user interface. **B** Cells colored according to their location

The cells were successfully associated with their respective region (Figure 13.B). Further analyses were conducted in a location-specific fashion according to the regions as defined in Figure 13.

3.1 Neighbors count distribution

Cellular patterns can be seen as sets of cells whose mutual arrangements possess some stable properties, such as the number of neighbours of a given cell in the structure. The distribution of the numbers of neighbours across an epithelium can thus potentially be used as a measure of its degree of organization. As discussed in the methods section, determining the number of neighbors of a cell was not trivial. The main difficulty came from the fact that the definition of neighboring cells (encoded in the software in the form of (Equation **1**) depends on both the context of the analysis and the operator.

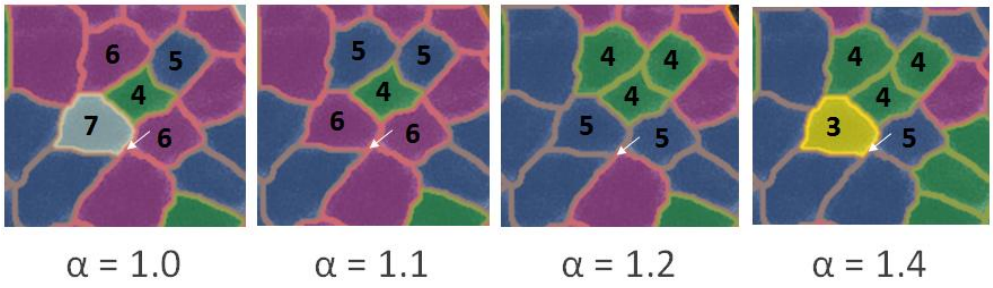

Figure 14 Effects of the proportionality coefficient on the number of neighbours as determined by the software

Slight increases of the value of the proportionality coefficient α were resulting in drastically different distributions of neighbors' counts in a given part of the epithelium (Figure 14), highlighting the need to define rigorous rules for the determination of neighboring cells. For the purpose of this study, the coefficient was set to 1.1 as it gave results comparable to what was found in the literature where the number of neighbors was determined by manually counting cells.

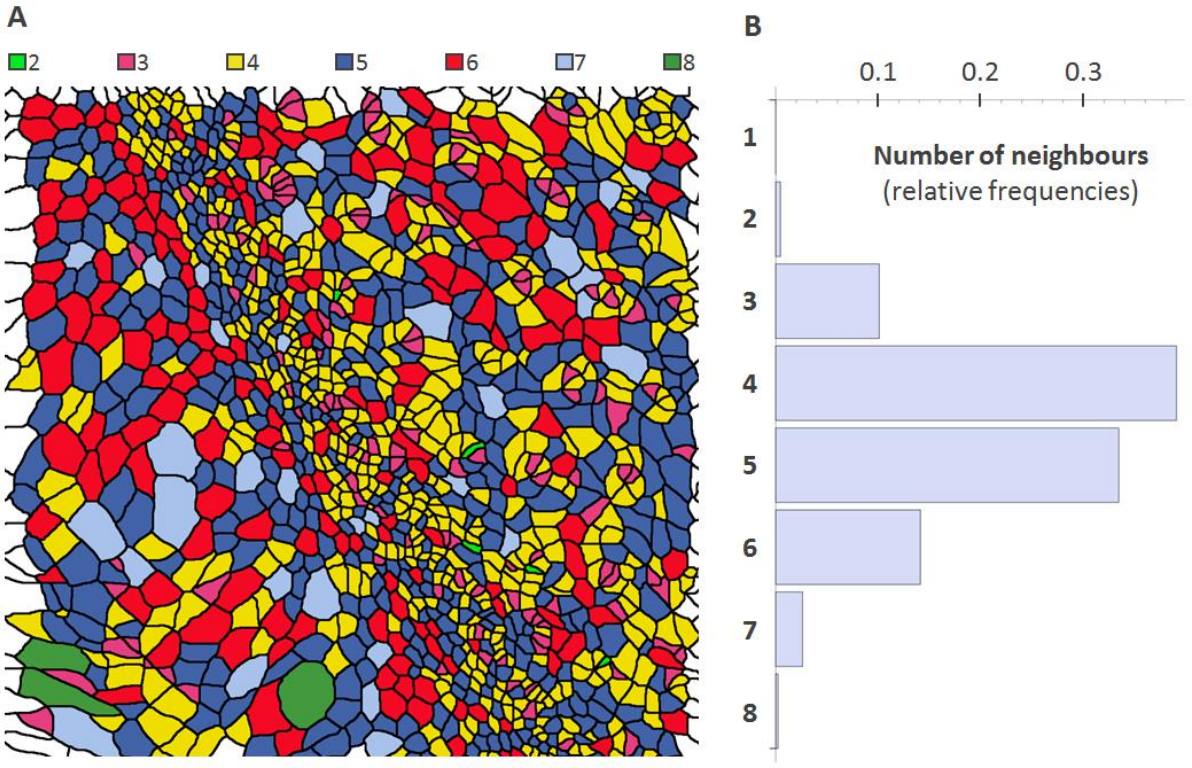

Figure 15 A Cells colored according to their number of neighbors. **B** Distribution of the number of neighbors across the whole picture (cells on the edge are not taken into account)

The number of neighbors was determined for every cell of the picture (Figure 15.A). The cells with 4 and 5 neighbors were largely predominant (Figure 15.B) and the average of the distribution was 4.59 +/- 0.027 (standard error of the mean).

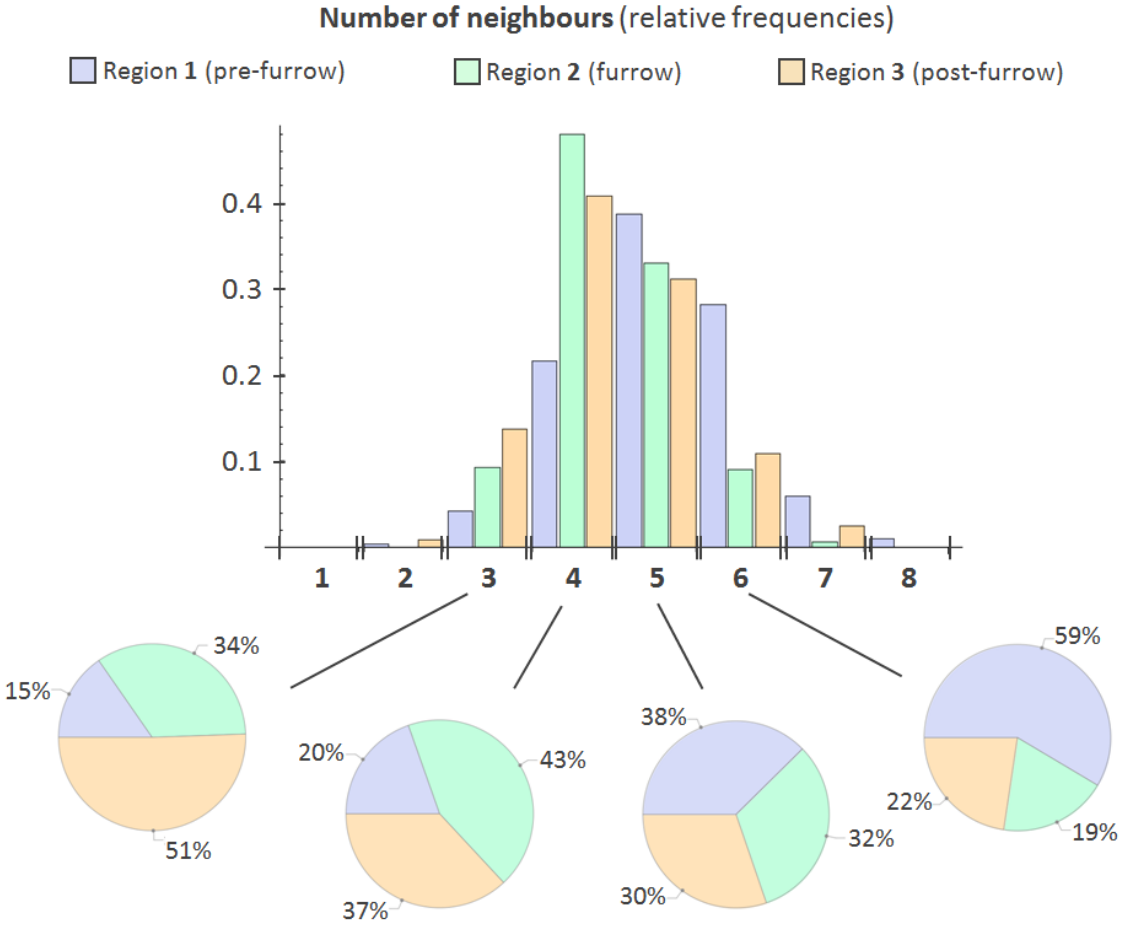

Figure 16 Region-specific distributions of the number of neighbors

Perhaps more interesting and information-rich results were obtained by determining the distributions of the number of neighbors in the three regions of the picture (Figure 16). The averages of the distributions were 5.12, 4.44 and 4.45 for regions 1,2 and 3 respectively. Cells with six or more neighbors were predominantly found in the unpatterned region. This might be explained by a higher amount of divisions in this region, resulting in cells having temporarily more neighbors as they grow in surface area right before dividing. The majority of cells with three neighbors were located in the post-furrow region, some of them contributing to the forming rosette structures that include three-sided cells. Cells with four neighbors were mainly present in the morphogenetic furrow itself and the patterned region, hinting at a possible optimal configuration found in ordered epithelium. Overall, the pre-furrow region showed a broad distribution of number of neighbors whereas that of cells in the furrow or in the post-furrow regions was narrower.

A more in-depth analysis of the neighboring relationship was performed in order to understand whether the number of neighbors of adjacent cells was independent or not. This relationship was represented as a matrix plot showing both the observed and an indicator of the deviation from the state of independence (Figure 17). Results showed that the most represented combination in both region 2 and 3 (morphogenetic furrow and post-furrow regions respectively) was borders associated with a 4 neighbors and a 5 neighbors cells whereas adjacent cells with 5 and 6 neighbors were predominant in region 1 (pre-furrow region). In all cases, there was a deficit of adjacent cells having five neighbors each.

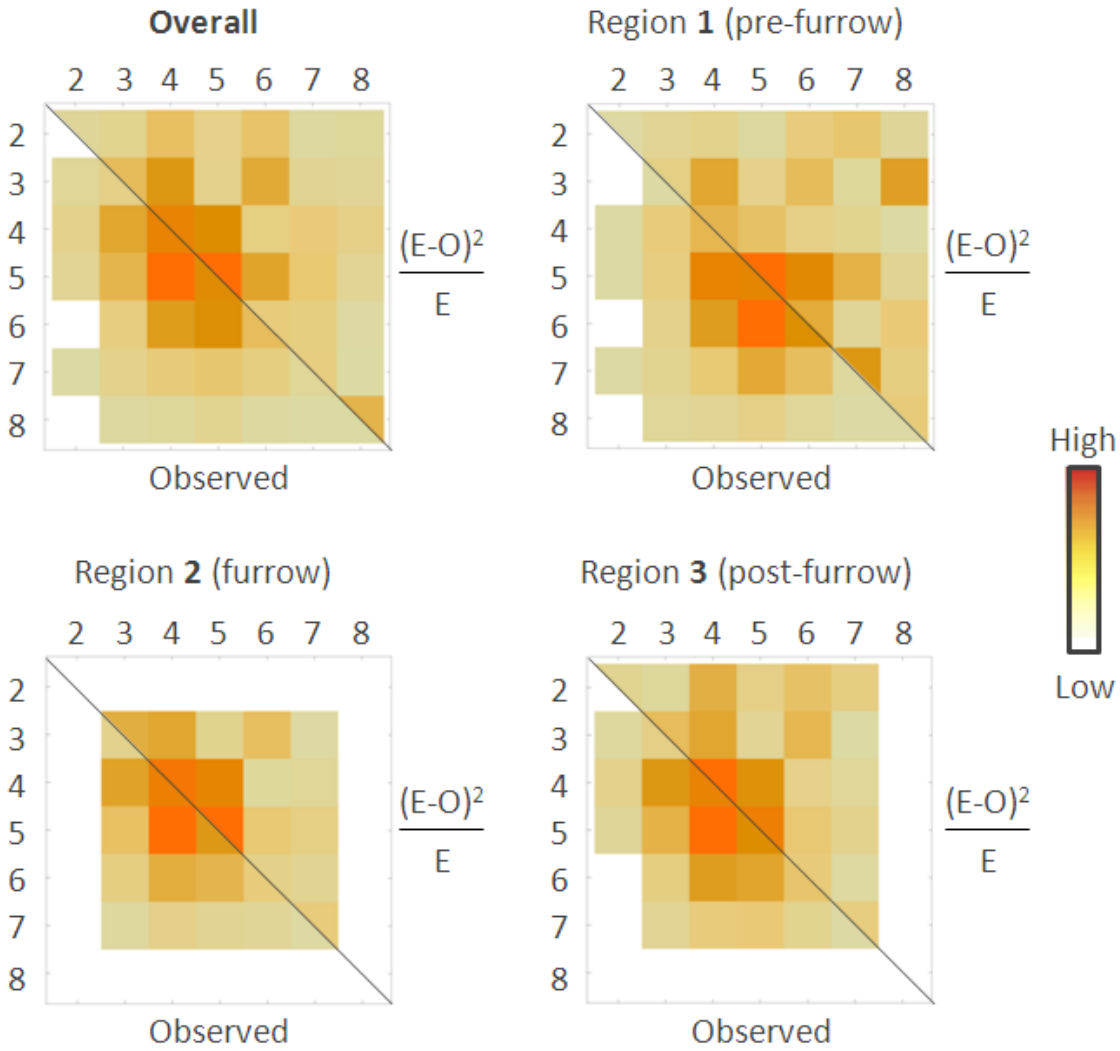

Figure 17 Relationship between the numbers of neighbors of adjacent cells. The lower part of the matrix plots represents the observed frequencies of the combinations of number of neighbors and the upper part is an indicator of the deviation from the state of independance

The fact the expected frequencies were different than the observed frequencies already showed that there was a dependency between the numbers of neighbors of adjacent cells. This relationship was further characterized by determining the χ^2 value associated with each case (Table 1).

Table 1 Chi-square values as computed for the observed and expected frequencies shown in **Figure 17**

Region	Number of borders	Chi-Square value	Pearson test p-value
Overall	3788	328.91	< 0.0001
Pre-furrow (**1**)	779	81.84	< 0.0001
Furrow (**2**)	1246	134.35	< 0.0001
Post furrow (**3**)	1580	129.1	< 0.0001

Pearson's chi-square test was applied with the null hypothesis that the two variables (that is, the two numbers of neighbors of adjacent cells) were independent. In all cases, the null hypothesis was rejected (p-value < 0.0001); confirming that the two variables were, in fact, dependant. In addition, pre-furrow cells presented the lowest Chi-Square value and thus the number of neighbors of adjacent cells in this region tended to be more independent, which could be linked with their unorganized state.

3.2 Size distribution

Another interesting feature of the picture to analyze was the size distribution of the cells. Much like the number of neighbors, it can be a potential indication of cells organization. The distributions were computed using both the surface area of the cells and their perimeter (Figure 18).

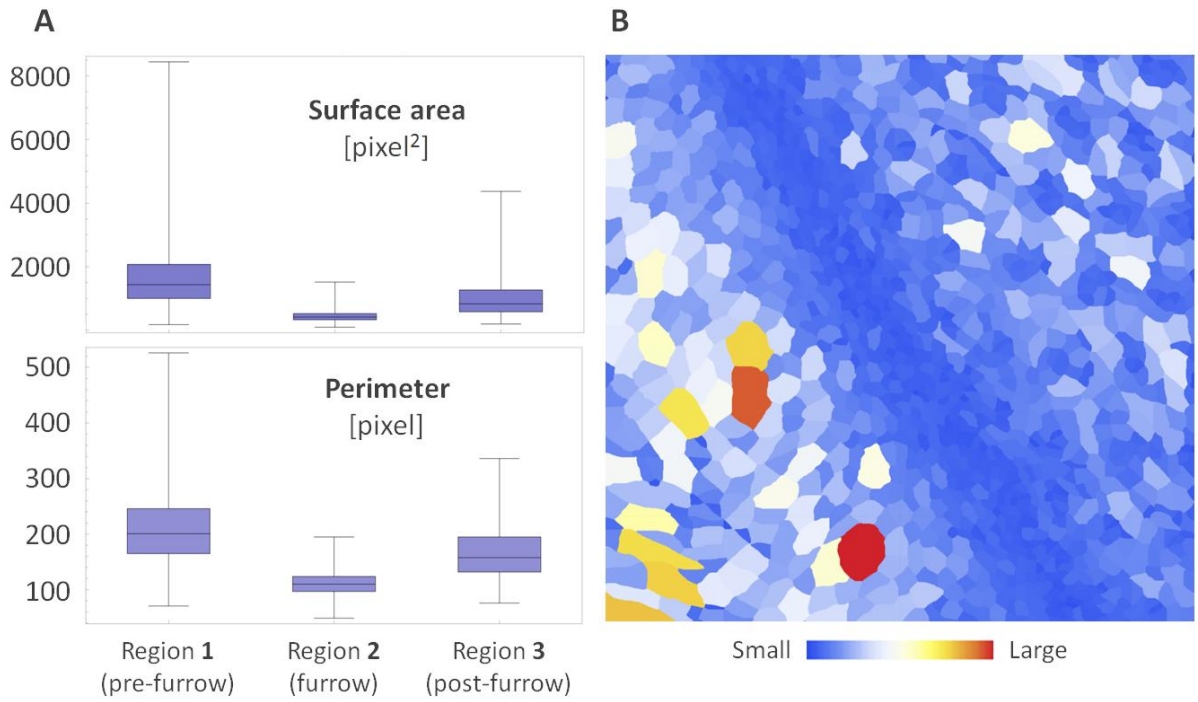

Figure 18 Size distribution of the cells **A** Box and whisker plot of the size distributions (the line is the median, box spans quartile 1 and 3, whiskers represent the extrema) **B** Cells colored according to their size (surface area)

The information given by the surface area and the perimeter of the cells were very similar. As expected due to their undergoing apical constriction, cells in the furrow were found to be the smallest and the most 'packed' (small spread in sizes). Region 1 and 3 were found to have similar medians but the former had a much broader size distribution, reflecting the facts that these cells are unorganized and not yet patterned.

4. Conclusion and outlooks

A proof of concept image-processing software programmed using Mathematica allowed the quantitative analysis of biologically relevant features of a confocal microscopy picture of the developing *Drosophila* retina through a user-friendly interface. Analyses were performed to determine the number of neighbors and size distributions in a region-specific manner. Cells in the patterned regions (morphogenetic furrow and post-furrow) tended to have narrow distributions both in size and in number of neighbors in opposition to the unorganized pre-furrow regions where both distributions were found to be much broader. This might indicate that these analyses could serve as a basis to quantify the degree of disorder of epithelial tissues. In addition, this metric could be further refined by integrating statistical tests such as the one applied in this study to determine if the numbers of neighbors of adjacent cells are dependant. Indeed, it was shown that they are not only dependant but also that the unorganized region had the lowest Chi-Square value, indicating that cells in this region are the most likely to have adjacent cells with independent number of neighbors, a sign of an increasing degree of disorder.

Tools such as the one described in this case essay are most likely going to be critical to fully understand the subtle interplay between biological and physical mechanisms that underlie morphogenesis. Further work should focus on integrating quantitative analyses into a specialized, flexible and expandable software framework that

allows the complete processing (covering the whole workflow described in Figure 8) of not only pictures but also of time-lapse movie frames,. The current vertices detection method should also be improved in order to be able to detect cells shape, which may be better at qualifying tissue organization than the number of neighbors.

Such a software package should go beyond relying on the data originating from a single microscopy channel and must exploit other fluorescent channels that were so far excluded from such quantitative analyses. For instance, by fluorescently tagging F-actin in addition to E-cadherins, polarization of the expression of these molecules could be tracked and quantified, producing data that could be in turn be used to create novel computational models of the mechanisms implicated in adhesion and cortical tension regulation and more generally in mechanisms leading to the emergence of cellular patterns in unorganized epithelial tissues.

5. References

1. Lecuit, T. and P.-F. Lenne, *Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis.* Nat Rev Mol Cell Biol, 2007. **8**(8): p. 633-44.
2. Lecuit, T. and L. Le Goff, *Orchestrating size and shape during morphogenesis.* Nature, 2007. **450**(7167): p. 189-92.
3. Bate, M. and A. Martínez Arias, *The Development of Drosophila melanogaster, Volume 1*. 1993: Cold Spring Harbor Laboratory Press. 1558.
4. Hayashi, T. and R.W. Carthew, *Surface mechanics mediate pattern formation in the developing retina.* Nature, 2004. **431**(7009): p. 647-52.
5. Kumar, J.P., *Signalling pathways in Drosophila and vertebrate retinal development.* Nat Rev Genet, 2001. **2**(11): p. 846-57.
6. Corrigall, D., et al., *Hedgehog signaling is a principal inducer of Myosin-II-driven cell ingression in Drosophila epithelia.* Dev Cell, 2007. **13**(5): p. 730-42.
7. Mirkovic, I. and M. Mlodzik, *Cooperative activities of drosophila DE-cadherin and DN-cadherin regulate the cell motility process of ommatidial rotation.* Development, 2006. **133**(17): p. 3283-93.
8. Forgacs, G., et al., *Viscoelastic properties of living embryonic tissues: a quantitative study.* Biophys J, 1998. **74**(5): p. 2227-34.
9. Foty, R.A. and M.S. Steinberg, *The differential adhesion hypothesis: a direct evaluation.* Dev Biol, 2005. **278**(1): p. 255-63.
10. Mombach, J., et al., *Quantitative comparison between differential adhesion models and cell sorting in the presence and absence of fluctuations.* Phys Rev Lett, 1995. **75**(11): p. 2244-2247.
11. Bao, S. and R. Cagan, *Preferential adhesion mediated by Hibris and roughest regulates morphogenesis and patterning in the Drosophila eye.* Dev Cell, 2005. **8**(6): p. 925-935.
12. Zallen, J.A. and E. Wieschaus, *Patterned gene expression directs bipolar planar polarity in Drosophila.* Dev Cell, 2004. **6**(3): p. 343-55.
13. Farhadifar, R., et al., *The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing.* Curr Biol, 2007. **17**(24): p. 2095-104.

6. Annexes

6.1 User Interface

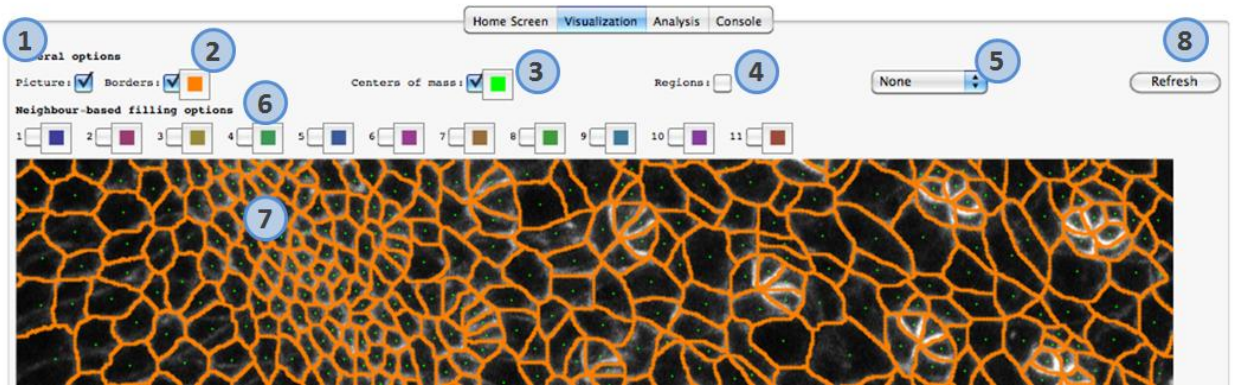

Figure 19 Visualization panel. (1) Show the microscopy picture in the background (2) Show borders and select the drawing color (3) Show the center of masses and select the drawing color (4) Show the regions selection tool (5) Filling options (6) Colors for the filling based on the number of neighbors of the cells (7) Main vizualisation area (8) Refresh the picture to reflect the chosen options

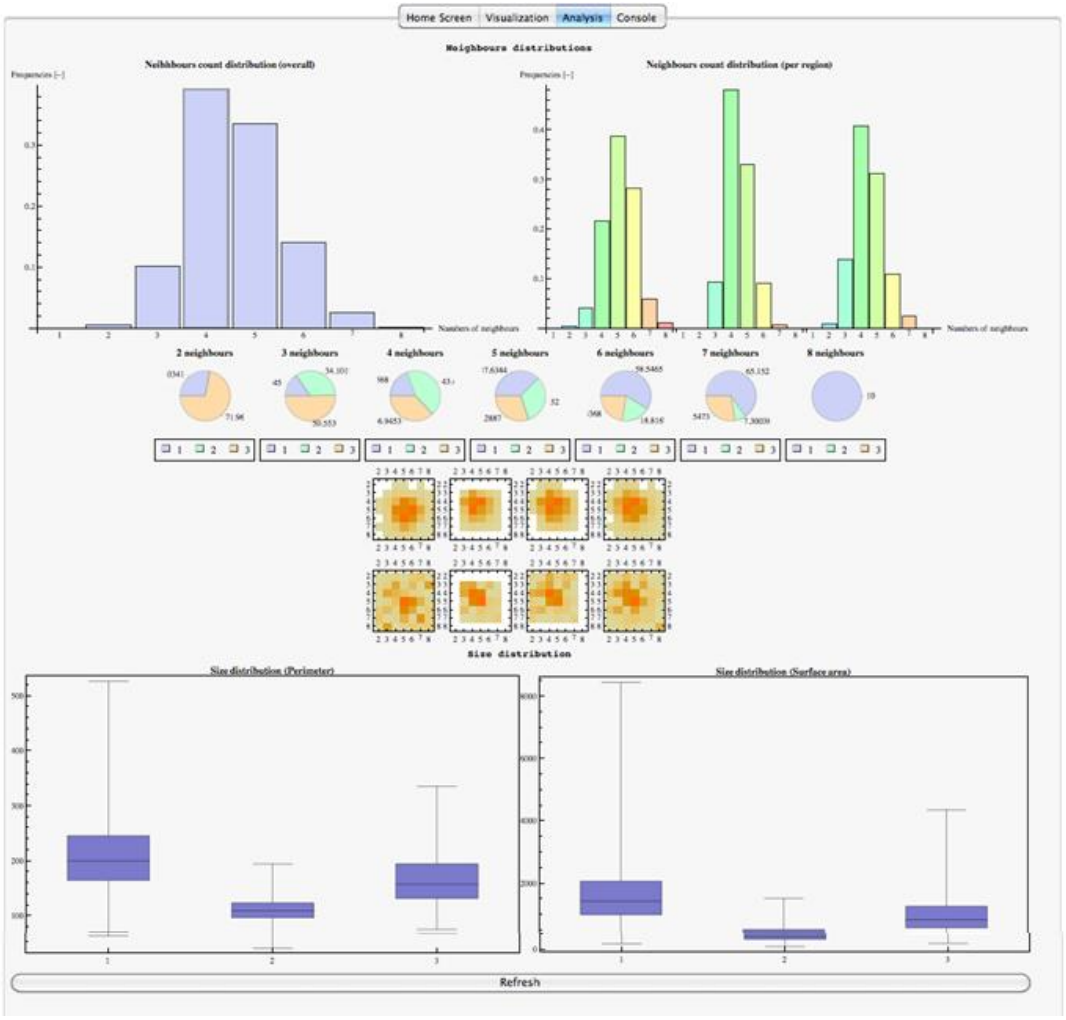

Figure 20 The analysis panel

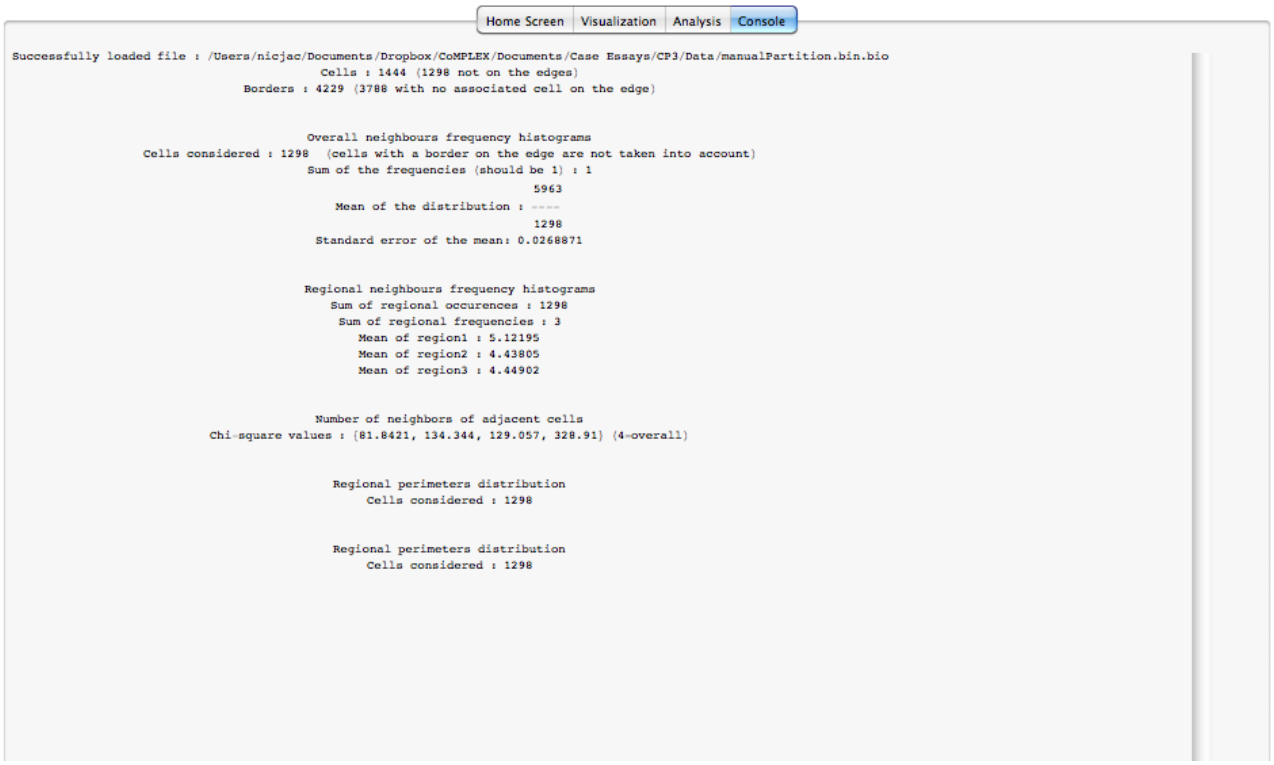

Figure 21 The console panel